Fast and accurate classifying model for denial-of-service attacks by using machine learning

Mohammed Ibrahim Kareem[1], Mahdi Nsaif Jasim[2]
[1]Department of Information Networks, University of Babylon, Hillah, Iraq
[2]Department of Business Informatics, University of Information Technology and Communications, Baghdad, Iraq

Article Info

Article history:

Received Feb 10, 2022
Revised May 9, 2022
Accepted May 24, 2022

Keywords:

CICIDS2017
Denial of service
Fast DDoS detection
J48 Classifier
Machine learning

ABSTRACT

A denial of service (DoS) attack is one of the dangerous threats to networks that Internet resources and services will be less available, as they are easily operated and difficult to detect. As a result, identifying these intrusions is a hot issue in cybersecurity. Intrusion detection systems that use classic machine learning algorithms have a long testing period and high computational complexity. Therefore, it is critical to develop or improve techniques for detecting such an attack as quickly as possible to reduce the impact of the attack. As a result, we evaluate the effectiveness of rapid machine learning methods for model testing and generation in communication networks to identify denial of service attacks. In WEKA tools, the CICIDS2017 dataset is used to train and test multiple machine learning algorithms. The wide learning system and its expansions and the REP tree (REPT), random tree (RT), random forest (RF), decision stump (DS), and J48 were all evaluated. Experiments have shown that J48 takes less testing time and performs better, whereases it is performed by using 4-8 features. An accuracy result of 99.51% and 99.96% was achieved using 4 and 8 features, respectively.

Corresponding Author:

Mohammed Ibrahim Kareem
Department of Information Networks, University of Babylon
Hillah, Iraq
Email: Mohamed.ibrahim@uobabylon.edu.iq

1. INTRODUCTION

Cybercriminals employ communication network and system weaknesses to launch denial of service (DoS) and distributed denial of service (DDoS) attacks, which flood the network and overwhelm servers with a huge request, compromising the availability of resources for legitimate users DDoS/Dos attack is not hard to implement that it can be running with limited resources to a big target network called "asymmetric attack". Most of the current research focuses on how to effectively send information from source to destination over the lack of security protocol designed to separate malicious intent [1]. These security gaps can be exploited for DDoS attacks [2]. DoS attacks are carried out by one system, whereas DDoS attacks are coordinated and carried out by numerous systems. Floods, fragmentation, TCP state exhaustion, and application-layer fatigue are some of the categories [3]. Floods are a type of attack that overwhelms machines with a large number of packets. A botnet, which is a group of malware-infected devices controlled by the online attacker, floods a victim's bandwidth with user datagram protocol (UDP) or internet control message protocol (ICMP) packets.

Fragmentation attacks use modified packets of networks that cannot be reassembled due to excessive packet headers. TCP state exhaustion attacks (protocol assaults) transmit huge internet protocol (IP) or TCP synchronize (SYN) packets to attack firewalls, load balancers, and servers. Simple mail transfer protocol (SMTP), hypertext transfer protocol secure (HTTPS), and domain name system (DNS) services are all monopolized by application-layer attacks. Because the requests appear legitimate, these assaults are the most difficult to detect.

Activity profiling, change-point detection, and wavelet analysis are some of the detection approaches for DoS and DDoS attacks [4]. In activity profiling, the headers of packets are monitored to estimate the average packet rate of inbound and outgoing flows by creating an activity profile based on packet field similarity and analysis of successive packets. Change-point detection creates a time series by clustering traffic data based on the address, port, or protocol. In order to explain network traffic, wavelet analysis is utilized to extract spectral components and distinguish anomalous events from typical network activity. Techniques for detecting intrusions recently developed [5], [6] that are based on machine learning (ML) algorithms [7].

Machine learning techniques to detecting DoS and DDoS cyber assaults necessitate a quicker discovery model that able to distinguish normal traffic from attack. As a result, when it comes to preventing cyberattacks on servers and avoiding denial of service to legitimate users when infractions first arise, a vetting process is essential to making the right decision. As a result, we analyze trustworthy machine learning methods such as REP tree (REPT), random tree (RT), random forest (RF), decision stump (DS), and J48.

The Canadian institute for cybersecurity (CIC) has created datasets [8] that capture DoS and DDoS attacks using a testbed infrastructure [9]. These databases are intentionally constructed by analyzing both normal and malevolent user behavior. Approaches that are routinely perform DoS and DDoS assaults are employed to represent malicious behavior. The CICIDS2017 and CSE-CIC-IDS2018 datasets cover application-layer DoS assaults, and the CICDDoS2019 datasets include TCP, UDP, and TCP/UDP DDoS attacks. Using the CICIDS2017 dataset, we compare the performance of several algorithm models.

This study compares the performance of REPT, RT, RF, DS, and J48 approaches in terms of testing duration, accuracy, F Score, precision, and recall. The major goal of this research is to develop a DDoS detection classifier that is both fast and accurate. As a result, employ feature selection techniques to decrease features before incorporating them into a quick and effective DDoS attacks detection model. Removed useless of features in IDS enhances its speed, lowers its memory requirements, and allows it to be used in real-time.

2. RELATED WORK

The surge in DDoS attacks, as well as the inadequacies of standard network-based detection processes, mandates the creation of novel attack detection systems. To detect and prevent various types of DDoS assaults, some machine learning-based data mining methodologies and algorithms have been used [10]. The authors developed a DRL-BWO algorithm for intrusion detection in UAV networks. Primarily, the networking data, fed as input, undergoes preprocessing to remove the unwanted data and transform it into a compatible format. Besides, the DRL involves improved reinforcement learning-based DBN for intrusion detection. Then, the DBN model has applied the determination of the existence of intrusions in UAV networks. At last, the BWO algorithm is employed to determine the optimal hyperparameter values involved in the presented model [11].

Instead of actual DDoS assault packets, utlize statistical traffic taken from an SNMP protocol monitoring IB is used. Its version of the support vector machine (SVM) algorithm was successful in detecting a real DDoS attack when tested with real DDoS attack traces, the solution proposed in [12]. A semi-supervised, federated-learning-based intrusion detection system is provided in this work. This model was trained using both labeled and unlabeled data in a semi-supervised manner. A semi-supervised FL that combines client-side unsupervised learning with server-side supervised learning. The untrained and supervised models are then automatically combined to produce a unified learning and classification solution for IDS [13].

The approach proposed in this study combines supervised and unsupervised methods. Using multiple flow-based criteria, a clustering method first isolates the abnormal traffic from the usual data. A classification technique is then used to label the clusters using particular statistical parameters. The authors analyze the suggested strategy using a large data processing framework, training on the CICIDS2017 dataset and testing on a different set of assaults from the more recent CICDDoS2019 [14].

The work presents a clustering-based method for distinguishing data representing network traffic flows, including both conventional and DDoS activity. Two distinct clustering methods cluster the unlabeled data, and a vote procedure determines the final classification of traffic flows. After labeling, the trained models for future classification are obtained using supervised machine learning techniques such as k-nearest neighbors (kNN), SVM, and RF [15].

Barati *et al.* [16] propose a DDoS attack detection mixed machine learning algorithms. The Multilayer Perceptron technique was used to detect threats using tenfold validation. This method has a good detection of over 99% throughout the article, but it was only evaluated on a single data set, thus it should be validated on a wider and more recent set of data sets. One of the flaws of DDoS attack systems based on ML techniques is the use of insufficient domain knowledge in conjunction with ML detection algorithms [17]. Importantly, ML algorithms have issues from the 'dimensional curse,' which means that as many of the not relevant features grow, the learned models fail to generalize effectively, implying that most ML techniques perform poorly with new types of DDoS. Using variance as a features identification method is used to extract characteristics associated with DDoS attacks [17].

3. ML ALGORITHMS

ML is a popular and effective method for detecting DDoS attacks. It offers a number of popular ways for creating a model that is trained using data, and then the model may be used to identify an attack. This work employs five classifiers to perform the binary classification task.

3.1. REPT

REP Tree is a rapid decision tree approach that is based on the C4.5 technique and can use either classification or regression trees (continuous result). The acquisition of information/variance is utilized to construct a regression/decision tree, which is then pruned using low error pruning (with back fitting). This approach has been used in the identification of anomalies in research such as [18].

3.2. RT

A dataset is divided into sub-spaces using RT, which then fits a constant to each sub-space. A single tree model has a high proclivity for instability and poor forecast accuracy. By bagging RT as a decision tree technique, however, it can produce extremely accurate results. RT features a significant degree of versatility as well as the capacity to train quickly [19]. This approach has been used in the identification of anomalies in research such as [18].

3.3. DS

DS is one-level decision tree-based machine learning paradigm [20]. That is, it is a decision tree with a single internal node (the root) that is connected to the terminal nodes instantaneously (its leaves). A decision stump makes a forecast based on just one input feature's value.

3.4. J48 or C4.5

Classifier J48. C4 generates a decision tree and uses this technique to generate it (an extension of ID3). A statistical classifier is another name for it. It is quite useful for categorizing and continually examining data [21]. J48 algorithm has been used in identify of DDoS by [22]-[24].

3.5. RF

The RF classifier has great prediction performance for classification issues because it uses an ensemble of Decision Trees. The many decision trees assist in categorizing in such a way that each tree in the forest determines which class the new instance should be assigned to. The new class's classification will be determined by a majority vote. The precision of decision-making improves as the number of trees involved grows. The number of trees must be given before the classifier can be applied to the datasets [20].

4. DATSETS

Test rules consisting of the victim and attacker networks were used to obtain CIC datasets [8]. B-profiles that mimic ordinary user activity were used to generate regular (beneficial) traffic. M-profiles were used to produce malicious traffic using standard technologies such as brute-force file transfer protocol (FTP), botnets, secure shell (SSH), DoS, DDoS, heartbeats, hacking, and web attacks. A network traffic flow analyzer was used to extract dataset features from aggregated TCP and UDP network streams. The destination IP address, port, protocol type, stream duration, and maximum/minimum packet size are all included in each dataset. Table 1 show lists the attacks that were considered in the CIC datasets. This paper focuses on the CICIDS2017 dataset to build a fast and reliable model and it is validated via use the CICDDoS2019 dataset.

Table 1. Types of attacks in CIC datasets

Dataset	Attact	No of data points
	GoldenEdye	10,293
CICIDS2017	Hulk	230,124
July 05, 2017	SlowHTTPTest	5,499
	Slowloris	5,796
CSE-CIC-IDS2018	GoldenEye	41,508
February 15, 2018	Slowloris	10,990
	Domaun name system	5,071,011
CICDDoS2019	Lightweight directory	2,179,930
December 01, 2019	Accces protocol network time protocol	1,202,642

5. DATA PRE-PROCESSING AND FEATURE SELECTION

The CICIDS2017 dataset with all eighty-four features was downloaded. As the first phase in data cleaning, NaN values, and duplicate columns were removed. The standard scaling method was used. The univariate feature selection approach was used to determine the relevance of each feature. WEKA was used to implement information gain and select relevant features. To achieve more accurate results while utilizing WEKA, features were developed based on multiple trials. Based on the feature score, the features were chosen. Table 2 lists the features that have a score for them. Table 3 lists the feature groupings.

Table 2. Ranks of features with previous work

S. No	Name of feature	F. score	Previous work
1	SourcePort	0.396	[25]
2	DestinationPort	0.776	[25]-[29]
3	Protocol used	0.2	[30], [31]
4	FlowBytes/s	0.309	[26], [29]-[30]
5	FlowPackets/s	0.287	[26], [30]
6	PacketLengthMean	0.568	[26], [28], [29]
7	PacketLengthVariance	0.532	[26], [28], [29], [32], [33]
8	AveragePacketSize	0.81	[25], [26], [28], [29], [33], [34]

Table 3. Feature groups

S. No	Number of features group	Features
1	Feature group 1	Source port, destination port, protocol, flow bytes/s, flow packets/s, packet length mean, packet length variance, average packet size
2	Feature group 2	Source and destination ports, protocol, flow bytes/s, flow packets/s, packet length mean, packet length variance
3	Feature group 3	Destination port, protocol, flow bytes/s, packet length mean, packet length variance
4	Feature group 4	Destination port, flow bytes/s, packet length mean, packet length variance

6. DISCUSSION OF THE RESULTS

The four major performance measures were used to evaluate the performance of classifiers and feature selection via information gain in WEKA (evaluate the worth of an attribute by evaluating the information obtained about the class). There are four of them: accuracy, recall, precision, and F1 scores. Accuracy represents the algorithm's accuracy in identifying attacks over both regular and attack traffic. Recall reflects the percentage of real assault traffic that is accurately detected. Precision refers to the detection of an assault over projected positive cases. The F1 score, which combines recall and accuracy, is also computed, and the metrics are defined according to (1)-(5):

$$\text{Accuracy } \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$\text{Recall } \frac{TP}{TP+FN} \tag{2}$$

$$\text{Precision } \frac{TP}{TP+FP} \tag{3}$$

$$\text{Precision } \frac{\text{Precision} * \text{Recall}}{\text{Precision} * \text{Recall}} \tag{4}$$

This work used three test modes: i) the first test-mode is split 50% train, remainder test; ii) the second test mode is divided into the 10.0% train and the 90% test; iii) the third test mode is split 8.0% train, remainder test. The remaining four evaluation criteria are used to calculate fit time, which shows the classifier's fitting time during the testing stage. The tables below show how well classifiers perform in terms of group concern. The performance of classifiers is shown in Table 4. (Feature group 1). In this experiment, the first test mode of split data was used. In this scenario, all eight attributes are considered while categorizing the items. When the data is analyzed, four classifiers, REPT, RT, RF, and DT-J48, are shown to have an accuracy of better than 99 percent. The DT-J48 classifier has a 99.9539 percent accuracy. The RFclassifier able know all fake packets with a 99.955 percent accuracy score, but it takes longer to test than DT. DS classifier has the lowest accuracy score of 97.7%, yet it is still respectable.

Table 5 shows the classifier's performance with feature group 2. The first test mode of split data was employed in this experiment. The DT classifier obtains the greatest accuracy of 99.81% when using feature group 2 with 7 specified features. The RF, RT, and REPT classifiers all have the same accuracy score for feature groups 2 and 1. Except for one exception, reducing the number of features from eight to seven has no discernible effect on classifier performance. It's also worth noting that with feature group 2, the accuracy of the J48, RF, and REPT classifiers reduces marginally.

Table 4. Performance of classifiers on feature group 1 with 8 selected features

Algorithm	Accuracy%	Recall%	Precision%	F1 score%	Fit time
REP tree	99.9513	1.00	1.00	1.00	0.36
Decision stump	81.5791	1.00	0.76	0.86	0.21
Random tree	99.961	1.00	1.00	1.00	0.13
Random forest	99.969	1.00	1.00	1.00	3.99
J48	99.9566	1.00	1.00	1.00	0.13

Table 5. Classifier performance on feature group 2 with 7 selected features

Algorithm	Accuracy%	Recall%	Precision%	F1 score%	Fit time
REP tree	99.8219	0.999	0.998	0.998	0.11
Decision stump	81.5791	1.00	0.76	0.86	0.08
Random tree	99.7821	0.998	0.998	0.998	0.15
Random forest	99.8379	0.999	0.998	0.999	2
J48	99.81%	0.998	0.998	0.998	0.21

Table 6 shows the performance of classifiers using feature groups 3 and 5 chosen features. A third test mode with split data was employed in this experiment. The classifiers REPT and J48 attain the greatest accuracy of 99.44% and 99.44%, respectively. About feature group 2, the accuracy of RF and RT classifiers falls.

Table 6. Classifier performance on feature group 3 with 5 selected features

Algorithm	Accuracy%	Recall%	Precision%	F1 score%	Fit time
REP tree	99.4467	0.998	0.992	0.995	0.26
Decision stump	81.1982	1.000	0.858	0.751	0.51
Random tree	99.2596	1.00	0.992	0.993	0.33
Random forest	99.2855	0.994	0.993	0.994	2.58
J48	99.4423	0.998	0.992	0.995	0.23

Table 7 shows the performance of classifiers using feature groups 3 and 5 chosen features. In this case, the REPT and DT classifiers attain the maximum accuracy of 99.44% and 99.44%, respectively. About feature group 2, the accuracy of RF and RT classifiers falls. Classifier fitting time for the RF classifier is the longest for all three feature sets, followed by the REPT classifier.

Table 7. Classifier performance on feature group 4 with 5 selected features

Algorithm	Accuracy%	Recall%	Precision%	F1 score%	Fit time
REP tree	99.48	0.993	0.995	0.994	0.38
Decision stump	81.6164	1.00	0.755	0.86	0.27
Random tree	99.4631	0.996	0.995	0.995	0.67
Random forest	99.467	0.996	0.995	0.995	2.78
Decsion tree-J48	99.5127	0.999	0.993	0.996	0.23

For eight attributes, Figures 1–4 provides a comparison of classifier accuracy, precision, recall, and F1 measure. Figure 1 depicts the evaluation of classifiers for the eight-feature group 1. The RF classifier achieves the best accuracy of 99.995% for feature group 1 with all 8 features. With 8 features, the REPT, DT, and RT classifiers get the maximum accuracy. Algorithm accuracy on average (excluding DS) with feature groups 1, 2, 3, and 4 is 99.96%, 99.81%, 99.36%, and 99.48%, respectively. Feature group 1 with all eight features has the best overall classifier accuracy.

Figure 2 depicts classifier recall for the three feature groupings. With eight feature groups, the recall of classifiers J48, DS, REPT, and RF remains the same. With four groups, the DS classifier obtains higher recall. Across eight feature groups, the recall of all classifiers is nearly constant. RT classifiers achieve the greatest recall for feature groups 1 and 3. Feature groups do not affect the DS classifier's recall. The average recall of algorithms with feature group 1, feature group 2, feature, and feature group 4 are 100%, 99.88%, 99.70%, and 99.68%, respectively. Feature group 1 achieves the highest overall recall.

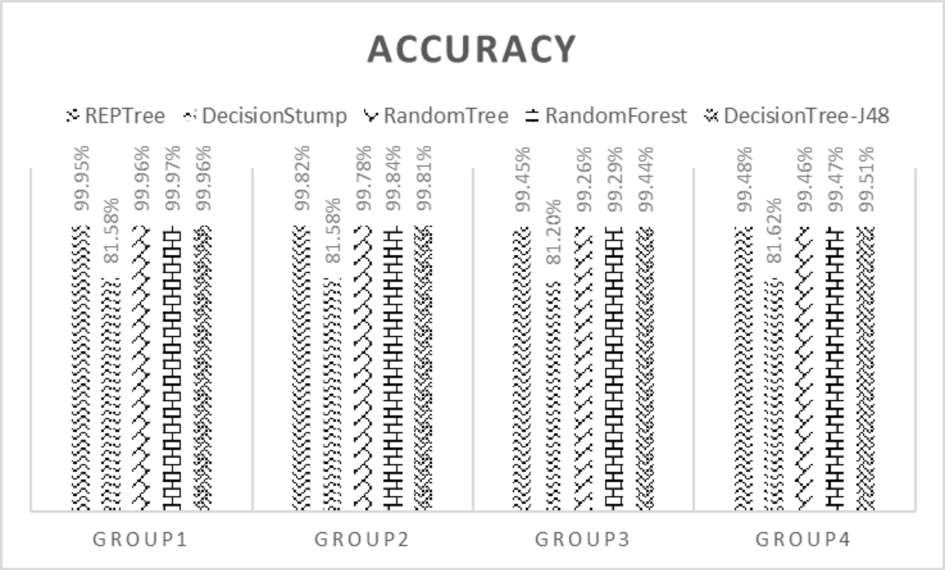

Figure 1. Classifier's accuracy for eight feature groups

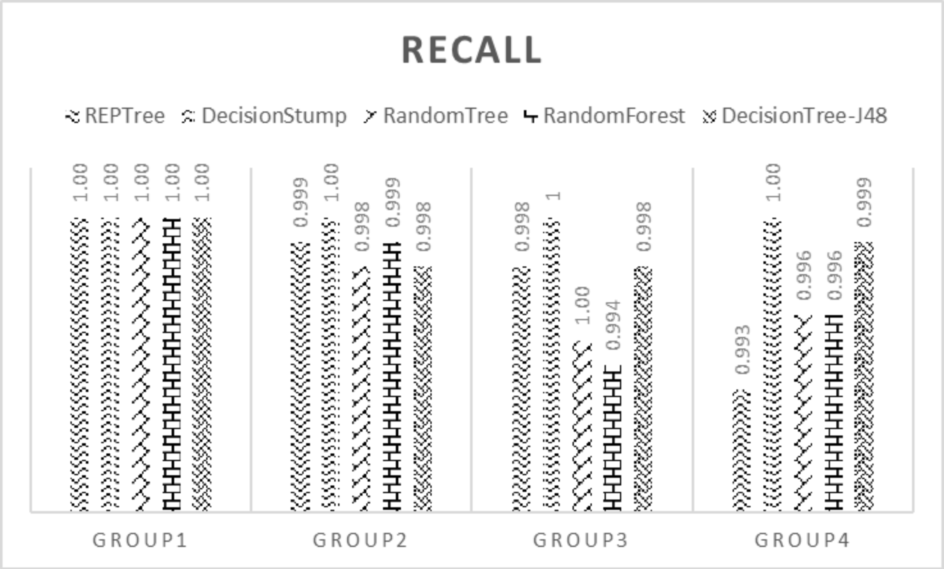

Figure 2. Classifier recall for eight feature groups

In terms of accuracy, as shown in Figure 3, all classifiers except DS score 100 percent with feature group 1. Except for DS, all classifiers achieve higher precision with feature group 2. The majority of algorithms are capable of identifying DDoS assaults with great accuracy. Except for the DS, the average precision of algorithms with feature groups 1, 2, 3, and 4 is 100%, 99.80%, 99.2%, and 99.4%, respectively. Feature group 3 achieves the highest overall recall.

Figure 4 depicts the F-measure scores of classifiers for the four feature groups. With feature group 1, RT, RF, and Decision Tree classifiers acquire a higher F measure. Feature groups have an impact on these algorithms. Average F-measure for all classifiers for feature group 1, feature group 2, feature group 3, and feature group 4 are 99.98%, 99.83%, 99.43%, and 99.50% respectively. Feature group 1 has the greatest overall classifier F-measure score.

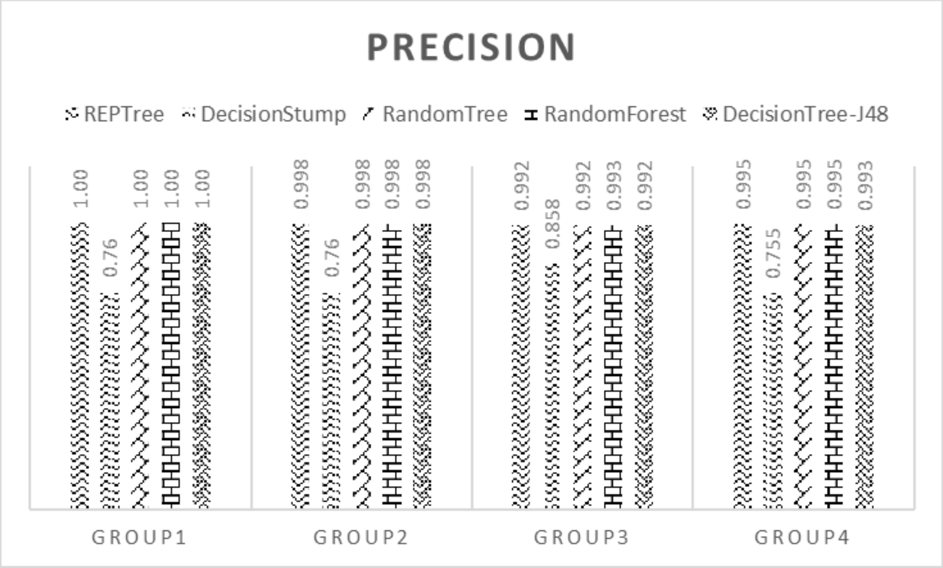

Figure 3. Precision of classifiers for three feature groups

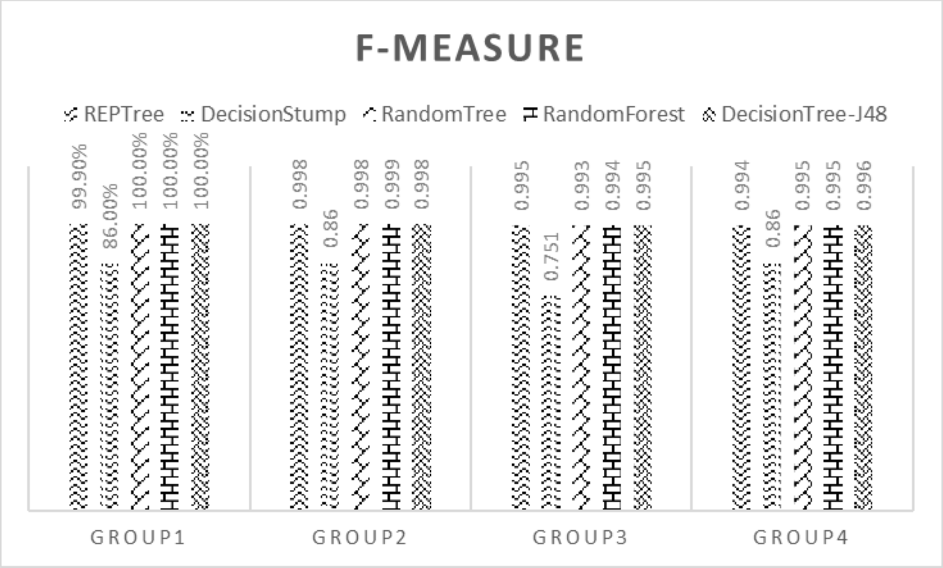

Figure 4. F-Measure for four different feature groups

After examining the four performance measures obtained by the five classifiers for the four feature groups, it is discovered that feature group 1 with all features achieves the highest overall classifier performance. According to the experimental results, the CICIDS2017 dataset characteristics chosen and employed in this work are capable of identifying DDoS assaults with an average accuracy of 99.65%, a recall rate of 99.82% on average, the average accuracy was 99.62%, while the average F1 score was 98.68%. Table 8 summarizes the best accuracy score for each feature group. The RF classifier achieved the maximum accuracy of 99.96% when all eight characteristics were used.

Table 8. Best feature group accuracy score

S. No	Number of features group	Features	The best classifier based on the rate of accuracy
1	Feature group 1	Source port, destination port, protocol, flow bytes/s, flow packets/s, packet length means, packet length variance, average packet size	99.96% with RF classifier
2	Feature group 2	Source and destination ports, protocol, flow bytes/s, flow packets/s, packet length mean, packet length variance	99.94% with RF classifier
3	Feature group 3	Destination port, protocol, flow bytes/s, packet length means, packet length variance	99.44% with decision tree-J48 and REPT classifiers
4	Feature group 4	Destination port, flowbytes/s, packetlengthmean, packet length variance	99.51% with decision tree-J48 classifier

While conducting the classification with six features, the J48 and REPT classifiers detected attacks with 99.44% and 99.45% accuracy, respectively. Using only four significant characteristics for classification, the J48 and REPT classifiers obtained 99.51% and 99.48% accuracy, respectively. With the four best-ranked characteristics – 'destination port, Flow Bytes/s, Packet Length Mean, Packet Length Variance', J48 is capable of identifying DDoS assaults - HTTP request flooding attacks, TCP SYN flooding attacks, UDP flooding attacks, and ICMP flooding attacks. These characteristics may be applied to the development of lightweight models, it is used in multi-stage classification systems for first-stage classification. Table 9 highlights the accuracy of categorization acquired in previous experiments.

Table 9. Previous work on classification is compared

S. No	Authors	Dataset	Accuracy of the classifier (%)	No. Features
1	Kurniabudi *et al.* [28], 2020	CICIDS2017	99.79	15
			96.47	4
2	Kurniabudi *et al.* [26], 2021	CICIDS2017	99.79	22
3	Çakmakçı *et al.* [34], 2020	CICIDS2017	99.55	10
4	Swe *et al.* [32], 2021	CICIDS2017	99.50	unknown
5	The current study	CICIDS2017	99.51	4
			99.96	8

In comparison to previous research, the suggested model using the top four features based on feature score derived by the InfoGainAttributeEval test was able to accurately detect DDoS attacks in real-time. Based on tests, the suggested J48 model takes short time to test and good perfom. As indicated in Table 9 for prior efforts, group 4 performed well with a tree size of 21 and 11 leaves, with an accuracy of 99.51%.

7. EVALUATION THE PROPOSED CLASSIFIER ON CICDDOS2019 DATASET

The validated models were not built with CICDDOS 2019, which is a current dataset, but it was used to confirm that the conclusion was correct. SYN and UDP assaults have both reached great accuracy, with 99.9914% and 99.7334%, respectively. When four features were employed to validate the efficacy of the suggested J48 model framework on the CICDDOS2019 with SYN attaches, the test time was 1.74 seconds, but the test time for model: was 0.53 seconds with UDP attack.

8. CONCLUSION

Using two CIC datasets, we compared the performance of five supervised ML methods. In this work used four metrics as a fellow (accuracy, F-score, precision, and recall) were employed as performance measurements. The number of mapped features, groups of mapped features, and enhancement nodes affected testing time for five algorithm models, whereas the number of estimators, learning rate, maximum depth, and a number of leaves in the J48 affected testing time. The accuracy of REPT, RT, RF, and J48 was good. For fast and accurate detection models, the smallest testing time was essential. The results illustrated the advantages of the J48 algorithm as the best choice when real-time detecting DoS and DDoS attacks. The proposed classifier was also tested on the CICDDoS2019 dataset, which showed a short test time and good accuracy for both UDP and SYN attacks. So, the proposed model is lightweight and useful for real-time implementation.

REFERENCES

[1] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *Computer Communication Review*, vol. 34, no. 2, pp. 39–53, 2004, doi: 10.1145/997150.997156.

[2] B. Chinnaiah, "Protection of DDoS Attacks at the application layer: HyperLogLog (HLL) cardinality estimation," in *Cognitive Informatics and Soft Computing*, Springer, 2021, pp. 595–604.

[3] A. Bhardwaj, V. Mangat, R. Vig, S. Halder, and M. Conti, "Distributed denial of service attacks in cloud: State-of-the-art of scientific and commercial solutions," *Computer Science Review*, vol. 39, p. 100332, Feb. 2021, doi: 10.1016/j.cosrev.2020.100332.

[4] G. Carl, G. Kesidis, R. R. Brooks, and S. Rai, "Denial-of-service attack-detection techniques," *IEEE Internet Computing*, vol. 10, no. 1, pp. 82–89, Jan. 2006, doi: 10.1109/MIC.2006.5.

[5] J. P. A. Maranhao, J. P. C. L. da Costa, E. P. de Freitas, E. Javidi, and R. T. de Sousa, "Noise-robust multilayer perceptron architecture for distributed denial of service attack detection," *IEEE Communications Letters*, vol. 25, no. 2, pp. 402–406, Feb. 2021, doi: 10.1109/LCOMM.2020.3032170.

[6] A. Iman and T. Ahmad, "Data reduction for optimizing feature selection in modeling intrusion detection system," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 6, pp. 199–207, Dec. 2020, doi: 10.22266/ijies2020.1231.18.

[7] P. Maniriho, L. Mahoro, E. Niyigaba, Z. Bizimana, and T. Ahmad, "Detecting Intrusions in computer network traffic with machine learning approaches," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 3, pp. 433–445, Jun. 2020, doi: 10.22266/ijies2020.0630.39.

[8] "Canadian Institute for Cybersecurity." https://www.unb.ca/cic/datasets/index.html (accessed Jan. 04, 2022).

[9] I. Sharafaldin, A. Gharib, A. H. Lashkari, and A. A. Ghorbani, "Towards a Reliable intrusion detection benchmark dataset," *Software Networking*, vol. 2017, no. 1, pp. 177–200, Jan. 2017, doi: 10.13052/jsn2445-9739.2017.009.

[10] S. Sen, K. D. Gupta, and M. M. Ahsan, "Leveraging Machine learning approach to setup software-defined network (SDN) controller rules during DDoS attack," in *Proceedings of International Joint Conference on Computational Intelligence. Algorithms for Intelligent Systems.*, Springer, 2020, pp. 49–60.

[11] V. Praveena *et al.*, "Optimal deep reinforcement learning for intrusion detection in UAVs," *Computers, Materials & Continua*, vol. 70, no. 2, pp. 2639–2653, 2022, doi: 10.32604/cmc.2022.020066.

[12] J.-H. Yu, J.-S. Park, H.-S. Lee, M.-S. Kim, and D.-H. Park, "Traffic flooding attack detection on SNMP MIB using SVM," *The KIPS Transactions:PartC*, vol. 15C, no. 5, pp. 351–358, Oct. 2008, doi: 10.3745/KIPSTC.2008.15-C.5.351.

[13] O. Aouedi, K. Piamrat, G. Muller, and K. Singh, "Intrusion detection for softwarized networks with semi-supervised federated learning," in *ICC 2022-IEEE International Conference on Communications*, 2022, pp. 1–6.

[14] M. Najafimehr, S. Zarifzadeh, and S. Mostafavi, "A hybrid machine learning approach for detecting unprecedented DDoS attacks," *The Journal of Supercomputing*, vol. 78, no. 6, pp. 8106–8136, Apr. 2022, doi: 10.1007/s11227-021-04253-x.

[15] M. Aamir and S. M. A. Zaidi, "Clustering based semi-supervised machine learning for DDoS attack classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 4, pp. 436–446, May 2021, doi: 10.1016/j.jksuci.2019.02.003.

[16] M. Barati, A. Abdullah, N. I. Udzir, R. Mahmod, and N. Mustapha, "Distributed denial of service detection using hybrid machine learning technique," in *2014 International Symposium on Biometrics and Security Technologies (ISBAST)*, Aug. 2014, pp. 268–273, doi: 10.1109/ISBAST.2014.7013133.

[17] B. A. Tama and K. H. Rhee, "Data mining techniques in DoS/DDoS attack detection: A literature review," *Information (Japan)*, vol. 18, no. 8, pp. 3739–3747, 2015.

[18] M. J. Awan *et al.*, "Real-time DDoS Attack detection system using big data approach," *Sustainability*, vol. 13, no. 19, p. 10743, Sep. 2021, doi: 10.3390/su131910743.

[19] A. Niranjan, D. H. Nutan, A. Nitish, P. D. Shenoy, and K. R. Venugopal, "ERCR TV: Ensemble of Random committee and random tree for efficient anomaly classification using voting," in *2018 3rd International Conference for Convergence in Technology (I2CT)*, Apr. 2018, pp. 1–5, doi: 10.1109/I2CT.2018.8529797.

[20] R. Khairy, A. Hussein, and H. ALRikabi, "The Detection of counterfeit banknotes using ensemble learning techniques of adaboost and voting," *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 1, pp. 326–339, Feb. 2021, doi: 10.22266/ijies2021.0228.31.

[21] A. L. Buczak and E. Guven, "A Survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016, doi: 10.1109/COMST.2015.2494502.

[22] A. B. Dehkordi, M. Soltanaghaei, and F. Zamani, "The DDoS attacks detection through machine learning and statistical methods in SDN," *The Journal of Supercomputing*, vol. 77, no. 3, pp. 2383–2415, Mar. 2021, doi: 10.1007/s11227-020-03323-w.

[23] O. Rahman, M. A. G. Quraishi, and C.-H. Lung, "DDoS attacks detection and mitigation in SDN using machine learning," in *2019 IEEE World Congress on Services (SERVICES)*, Jul. 2019, vol. 2642, pp. 184–189, doi: 10.1109/SERVICES.2019.00051.

[24] K. M. Sudar, M. Beulah, P. Deepalakshmi, P. Nagaraj, and P. Chinnasamy, "Detection of distributed denial of service attacks in SDN using machine learning techniques," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2021, pp. 1–5, doi: 10.1109/ICCCI50826.2021.9402517.

[25] T. Mourouzis and A. Avgousti, "Intrusion detection with machine learning using open-sourced datasets," *arXiv Computer Science*, pp. 1–18, Jul. 2021, [Online]. Available: http://arxiv.org/abs/2107.12621?utm_source=researcher_app&utm_medium=referral&utm_campaign=RESR_MRKT_Researcher_inbound%0Ahttp://arxiv.org/abs/2107.12621.

[26] Kurniabudi, D. Stiawan, Darmawijoyo, M. Y. Bin Idris, A. M. Bamhdi, and R. Budiarto, "CICIDS-2017 dataset feature analysis with information gain for anomaly detection," *IEEE Access*, vol. 8, pp. 132911–132921, 2020, doi: 10.1109/ACCESS.2020.3009843.

[27] E. M. Zeleke, H. M. Melaku, and F. G. Mengistu, "Efficient Intrusion detection system for SDN orchestrated internet of things," *Journal of Computer Networks and Communications*, vol. 2021, pp. 1–14, Nov. 2021, doi: 10.1155/2021/5593214.

[28] K. Kurniabudi, D. Stiawan, D. Darmawijoyo, M. Y. Bin Idris, B. Kerim, and R. Budiarto, "Important features of CICIDS-2017 Dataset for anomaly detection in high dimension and imbalanced class dataset," *Indonesian Journal of Electrical Engineering and*

Informatics (IJEEI), vol. 9, no. 2, pp. 498–511, May 2021, doi: 10.52549/ijeei.v9i2.3028.

[29] A. Abbas, M. A. Khan, S. Latif, M. Ajaz, A. A. Shah, and J. Ahmad, “A New ensemble-based intrusion detection system for internet of things,” *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 1805–1819, Feb. 2022, doi: 10.1007/s13369-021-06086-5.

[30] S. Sarraf, “Analysis and detection of DDoS attacks using machine learning techniques,” *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, vol. 66, no. 1, pp. 95–104, 2020.

[31] R. B. Adhao and V. K. Pachghare, “Performance-based feature selection using decision tree,” in *2019 International Conference on Innovative Trends and Advances in Engineering and Technology (ICITAET)*, Dec. 2019, pp. 135–138, doi: 10.1109/ICITAET47105.2019.9170235.

[32] Y. M. Swe, P. P. Aung, and A. S. Hlaing, “A slow ddos attack detection mechanism using feature weighing and ranking,” *Proceedings of the International Conference on Industrial Engineering and Operations Management*, pp. 4500–4509, 2021.

[33] A. Yulianto, P. Sukarno, and N. A. Suwastika, “Improving AdaBoost-based intrusion detection system (IDS) performance on CIC IDS 2017 Dataset,” *Journal of Physics: Conference Series*, vol. 1192, no. 1, p. 012018, Mar. 2019, doi: 10.1088/1742-6596/1192/1/012018.

[34] S. D. Çakmakçı, T. Kemmerich, T. Ahmed, and N. Baykal, “Online DDoS attack detection using Mahalanobis distance and Kernel-based learning algorithm,” *Journal of Network and Computer Applications*, vol. 168, p. 102756, Oct. 2020, doi: 10.1016/j.jnca.2020.102756.

BIOGRAPHIES OF AUTHORS

Mohammed Ibrahim Kareem his Bachelor’s degree in Information Technology from the University of Babylon, Iraq, in 2013, and his Master’s degree in Data Mining from the University of Babylon, Iraq, in 2019. He currently holds a Ph.D. Information security student at Babylon University, Iraq. Worked for several years as a software engineer at OMNNEA Telecom from 2014 to 2020. His research interests include computer networking, network security, and data mining. He can be contacted at email: mohamed.ibrahim@uobabylon.edu.iq.

Mahdi Nsaif Jasim is Assit. Prof. Dr. Mahdi Nsaif Jasim, University of Information Technology and Communications, College of Business Informatics Dept. of Management Information Systems. Born in Babylon, Iraq, lives in Baghdad. Interest: information systems, data and information security, mining in vector data, GIS, database systems. The researcher has interrest in SDN data acquisition and data processing. He also Supervised a number of PhD and MSc. Students in different Iraqi universities. Dr. Mahdi has been supervised 10 MSc students and 5 PhD students. He taught number os BSc. and MSc. courses a number of Iraqi universities. He can be contacted at email: mahdimnsaif@uoitc.edu.iq.