

K-Means clustering-based semi-supervised for DDoS attacks classification

Mahdi Nsaif Jasim¹, Methaq Talib Gaata²

¹Department of Business Informatics, University of Information Technology and Communications, Baghdad, Iraq

²Department of Computer Science, Mustansiriyah University, Baghdad, Iraq

Article Info

Article history:

Received Jul 4, 2022

Revised Aug 5, 2022

Accepted Aug 16, 2022

Keywords:

CICIDS2017

Clustering

Distributed denial of service

Feature selection

K-Means algorithm

Network security

ABSTRACT

Network attacks of the distributed denial of service (DDoS) form are used to disrupt server replies and services. It is popular because it is easy to set up and challenging to detect. We can identify DDoS attacks on network traffic in a variety of ways. However, the most effective methods for detecting and identifying a DDoS attack are machine learning approaches. This attack is considered to be among the most dangerous internet threats. In order for supervised machine learning algorithms to function, there needs to be tagged network traffic data sets. On the other hand, an unsupervised method uses network traffic analysis to find assaults. In this research, the K-Means clustering algorithm was developed as a semi-supervised approach for DDoS classification. The proposed algorithm is trained and tested with the CICIDS2017 dataset. After using the proposed hybrid feature selection methods and applying multiple training, testing, and carefully sorting DDoS traffic through a series of experiments, the optimum 2 centroids were found to be DDoS and normal. The generated centroids can be used to classify network traffic. So the proposed method succeeded to cluster the network traffic to safe and threat.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Mahdi Nsaif Jasim

Department of Business Informatics, University of Information Technology and Communications

Baghdad, Iraq

Email: mahdinsaif@uoitc.edu.iq

1. INTRODUCTION

A distributed denial of service (DDoS) attack is a form of a denial of service (DoS) attack in which the attacker targets the victim by utilizing the IP address of an authorized user. The numerous DDoS assaults consist of SYN-flood, ACK-flood, UDP-flood, connection DDoS, DNS reflect, and ICMP flood, among others [1]. An attack's primary goal is to prevent its intended recipients from making use of its intended services by overloading those resources. One tactic attacker uses to accomplish this is to send a barrage of fake requests through the network. DDoS is launched from multiple computers simultaneously. By overwhelming the infrastructure that surrounds the internet traffic flow, a DoS attack is a malicious technique that interferes with the regular traffic and networking operations of a targeted server. The rate and volume of network traffic sent to the target closely correlate with the attack's severity [2].

Since the 1990s, sophisticated intrusion detection systems have been made with the help of data mining. Data mining techniques in general, and machine learning techniques in particular, must be applied in five steps: selection, preprocessing, transformation, mining, and interpretation [3], [4]. Out of all the ways to find intrusions using data mining, these three important steps are the hardest. There are three types of machine learning-based DDoS detection methods that are already in use. Supervised ML approaches that build the detection model from datasets of network traffic that have been generated and labeled. The supervised

approaches have to deal with two big problems. First, making labeled network traffic datasets takes a lot of time and computing power. Without constant model updates, supervised machine learning techniques cannot predict novel actions that are simultaneously safe and risky. Second, supervised ML classifiers don't work as well when there is a lot of abnormal data in the traffic of the network. This is called noise. In the second group, there is no need for a labeled dataset to build the detection model, which is different from the first group. The main problem with the unsupervised methods is that they give out a lot of false positives. The curse of the dimensionality problem [5] makes it hard for unsupervised methods to find attacks accurately [6]. By being able to work on both labeled and unlabeled datasets, semi-supervised ML concepts take advantage of both supervised and unsupervised techniques. Also, using both supervised and unsupervised methods together can improve accuracy and reduce the number of false positives. But the problems with both approaches also make it hard for semi-supervised approaches to work. So, semi-supervised approaches need to have their parts put together in a smart way to make up for the problems with supervised and unsupervised approaches.

A group of machine learning tasks and techniques known as "semi-supervised learning" combine labeled with unlabeled samples for training, frequently combining a little amount of labeled samples with a large number of unlabeled samples. Semi-supervised learning way lies in the middle between supervised and unsupervised learning. Numerous machine learning researchers have demonstrated that integrating small amounts of labeled data with unlabeled data can dramatically improve learning accuracy compared to unsupervised learning without the time and expense of supervised learning. The general rule is first explored using labeled data in a semi-supervised learning process, and then the rule is applied to infer unmarked data. The machine learning algorithm that is enhanced for intrusion detection [7].

The primary goal of this work is to locate an appropriate method for classifying DDoS attacks by making use of semi-supervised learning and basing it on a global DDoS dataset. In addition to locating the most effective centroids for application in the offensive classification. The following are some of the benefits of our proposed algorithm over earlier detection solutions using supervised learning and unsupervised learning approaches: i) fewer labeled samples are needed to train detection models with our proposed method than with supervised learning detection algorithms, ii) proposed hybrid feature selection method using both low variance filter and information gain ration techniques, iii) present DDoS and regular centroids to assist in the implementation of them online for traffic classification. Following is a summary of the remaining sections of this paper. The related works in DDoS attack detection are introduced and their limitations. Our detection model, built on a semi-supervised clustering algorithm, is presented in section 2. Following the results and analyses of the experiments and a discussion of their significance, the paper concludes with recommendations for further research. The detection of DDoS attacks has been proposed using a variety of different methods such as [8]–[10]. Techniques based on machine learning are the ones that appear most frequently in published works of research. Table 1 (in Appendix) provides a brief overview of some recent research and developments in DDoS detection.

2. THE PROPOSED METHOD

In the beginning of this part, the dataset utilized in this study is described. Then, the proposed method used for intrusion detection and proposed centroids clustering, are present as shown in Figure 1. Finally, the results are analyzed and discussed.

2.1. Description of the dataset

Sharafaldin *et al.* [21] suggested the CICIDS2017 to get around the fact that there aren't enough IDS datasets that satisfy criteria of real-world network traffic [22]. The valid and widely used dataset CICIDS2017 [23], which is the largest and most used dataset [24]. 20% from the CICIDS2017 dataset is used in current work to train the machine learning algorithm. This set of data includes 84 features, as well as both unattack traffic and attack traffic. The CICIDS2017 dataset has a lot of information with a high-class imbalance.

2.2. K-Means clustering algorithm

A vector quantization technique known as "k means" try to group n observations in order to create k clusters, where every one observation belongs to one cluster that has the nearest mean (also known as the cluster centroid or cluster centers), which acts as the cluster's prototype [25]. The both algorithms (Hierarchical clustering and K-Means) frequently use canopy method as a preprocessing step in their respective processes [26]. Its purpose is to increase the speed at which clustering operations are performed on large data sets, where it may be impractical to use another algorithm directly due to the volume of the dataset.

2.3. Feature selection methods

One of the most common problems researchers' encounters is choosing which features are most important and thus relevant for use in detecting attacks. Feature selection is critical because it affects how well the system works. Too few features may be guide to subpar detection accuracy, while too many may lead to excellent detection

accuracy at the expense of an overly complex system that eats up more resources. This work employed two attractive features selection techniques; Figure 1 represent the main diagram of proposed framework.

2.3.1. Variance filter feature selection technique

The low variance filter method [27] was used to choose the features that were used in this paper, since all of the attributes were numbers. The method was used to exclude features with low variances that contributed slight or nil to the model's overall performance. Calculating the variance of each characteristic is involved (1).

$$\text{Variance } (\sigma^2) = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N} \quad (1)$$

where μ is the average of all the values that are associated with the attribute. The attribute values, denoted by X_i , are taken from a collection of data, where N is the total number of samples.

2.3.2. Information gain

Due to its usefulness and importance in detecting a class type, the IGR [28] is also employed as a weight for attributes in this work (2).

$$\text{IGR}(Y, A_j) = \frac{H(Y) - H(Y|A_j)}{H(A_j)} \quad (2)$$

where Y represents the class and A_j the index of j^{th} attribute. The entropy function, $H(\cdot)$, is defined as follows:

$$H(X) = -\sum_{i=1}^n p(x_i) \log p(x_i) \quad (3)$$

Given an input, the probabilities can be expressed as where $P(\cdot)$ represents the probability operator and i represents an index of the probabilities.

2.4. Proposed centroids clustering

The proposed method is the use of semi-supervised K-Means Clustering to generate multiple centroids that can be used to classify traffic as either safe or malicious. Starting with the selected CICIDS2017 dataset, we use the K-Means algorithm to produce semi-supervised centroids for detecting DDoS attacks. The idea of semi-supervised involves the use of small number of labelled data for the purpose of labeling larger data sets. Figure 1 shows semi-supervised framework diagram.

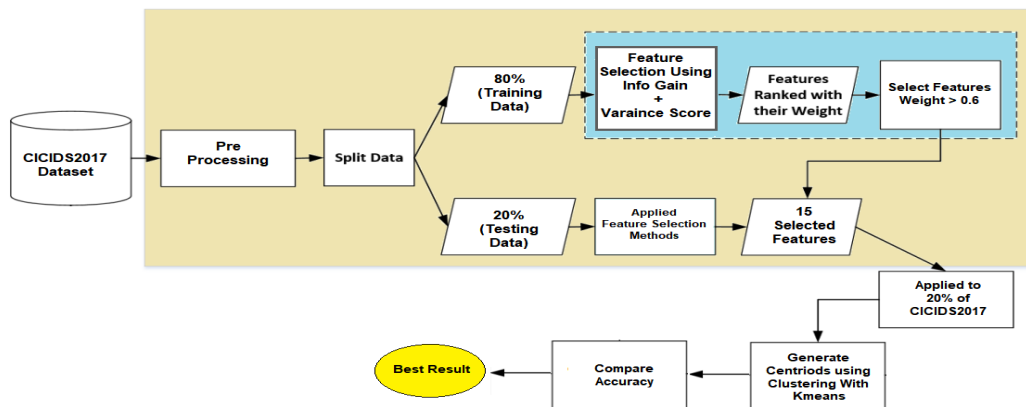


Figure 1. Proposed framework

The main processes in proposed framework are illustrated as follows:

- The features that were chosen using hybrid the feature selection algorithms. In this work, the variance scores and the information gain were used to discover the perfect list of features. By applying variance to exclude useless features with a variance score less than 3. In addition, discarding features with a minimum weight of 0.6 from the information gain, then 15 selected features are produced, as shown and listed in Table 2. Note the variance values for all the data ranges (0 to 9.99E+14) for (Bwd PSH Flags and Fwd IAT Total) features respectively.
- Utilize the K-Means algorithm to generate the appropriate centroids. 20% of the CICIDS2017 dataset was used to train the proposed method to generate centroids, and the remaining 80% of the dataset was used to test generated centroids.

- c. Compare the results with the accuracy scores and select the best result.

Table 2. Features scores using info.gain

Feat No.	Feat Name	Feat Score
1	SubflowFw Bytes	0.939343
2	TotalLength of FwdPackets	0.939343
3	AveragePacketSize	0.80995
4	TotalLength of BwdPackets	0.782456
5	SubflowBwdBytes	0.782456
6	BwdPacketLengthMean	0.781841
7	AvgBwdSegmentSize	0.781841
8	FwdHeaderLength	0.778016
9	DestinationPort	0.77582
10	BwdPacketLengthMax	0.760317
11	InitWinbytesforward	0.708411
12	AvgFwdSegmentSize	0.706064
13	FwdPacketLengthMean	0.706064
14	FwdPacketLengthMax	0.701009
15	BwdHeaderLength	0.682524

3. RESULTS AND EVALUATION

The detection performance of the semi-supervised K-Means algorithm was measured in this experiment. WEKA's performance of clustering and feature selection by information gain. Accuracy measures the algorithm's ability to detect attacks in both unattack and attack traffic. The accuracy computed according (4).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

The performance of the detection engine can also be measured by its accuracy. The machine's ability to predict traffic based on its actual conditions is indicated by its accuracy. In other words, the capacity of a machine to precisely classify a class. Figure 2 and Table 3 present values of generated centroids of the proposed method. It is providing two optimum centroids to classify traffic into normal and DDoS attack. Table 4 displays K-Means accuracy performance. The results shown in Table 4 illustrate that the test 1 was the best choice to achieved accuracy with 2 centroids that labeled into normal and another with DDoS. Figure 3 present performance comparison between the proposed K-Means and Canopy.

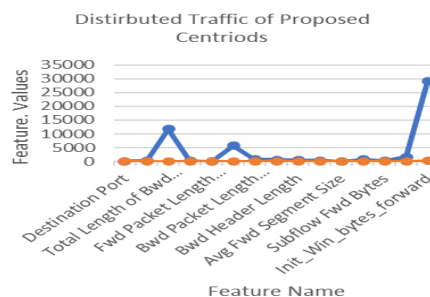


Figure 2. Distributed traffic of proposed centroids

Table 3. Values of generated centroids

No.	Feat. Name	Centroid1 (DDoS)	Centroid2 (Normal)
1	Destinatio Port	80	80
2	TotalLength of FwdPackets	288	30
3	TotalLength of Bwd Packets	11724	0
4	FwdPacketLengthMax	288	6
5	FwdPacketLengt Mean	13.714286	6
6	BwdPacketLengthMax	5792	0
7	BwdPacketLengthMean	732.75	0
8	FwdHeaderLength	680	100
9	BwdHeaderLength	520	0
10	AveragePacketSize	324.648649	7.2
11	AvgFwdSegmentSize	13.714286	6
12	AvgBwdSegmentSize	732.75	0
13	SubflowFwdBytes	288	30
14	SubflowBwdBytes	1724	0
15	Init_Win_bytes_forward	29200	256

Table 4. Accuracy of K-Means and canopy algorithms

Test No.	Keans Accuracy (%)	Canopy Accuracy (%)
Test1 (2 centroids)	79.60	72.30
Test2 (4 centroids)	68.90	65.70
Test3 (6 centroids)	42.10	55.90

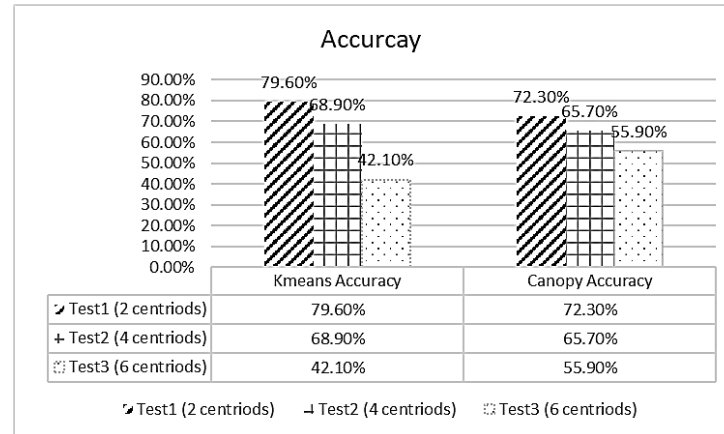


Figure 3. Performance of proposed method and canopy algorithm

4. CONCLUSION AND FUTURE WORK

This paper presents the algorithm to classify DDoS attacks using a semi-supervised machine learning method. It starts with traffic statistics that aren't labeled that are gathered from three parts of the victim-end defense, which is the web server. Proposed hybrid feature selection techniques to reduction dataset feature from 84 to 15 of the features are used to final labeling of traffic flows in proposed framework. K-Means clustering algorithm group the data that doesn't have labels. The scheme used a representative part of the benchmark CICIDS2017 dataset with new normal and attack centroids to test how well labels were given. In the future, we want to find better ways to voting based label traffic online, add more ML algorithms to the clustering and classification processes, and put the proposed four centroids into the online detection framework.

ACKNOWLEDGEMENTS

The Authors would like to thank, University of Information Technology and Communications and Mustansiriyah University (<https://uomustansiriyah.edu.iq/>), Baghdad-Iraq for its support in the present work.

APPENDIX

Table 1. Recent related work

No	References	Technique Name	Results Discussion
1	[11]	Fuzzy c-means clustering	The research is based on network traffic characteristics retrieved from the network that might indicate the presence of DDoS botnets in the network. According to the findings of the experiments, the detection rate is around 95%, with only 6% of false positives.
2	[12]	Co-clustering, Information Gain Ratio, and the Extra-Trees technique and estimating entropy	The entropy estimator examines the entropy of network traffic data over a sliding time-based frame. Co-clustering divides incoming network traffic into three groups when entropy exceeds thresholds. The information gain ratio (IGR) is calculated using the average network header entropy between each cluster and the current time frame subset. Extra-Trees ensemble classifiers are used for preprocessing and classification of high-gain anomalous network traffic data clusters.
3	[13]	Clustering Using Representative (CURE), Entropy	The intrusion detection method described in this article combines several unsupervised data mining techniques. Entropy theory in terms of packet windowing and data mining are integrated to identify the DDoS attack in network flow. As a cluster analysis, clustering using representative (CURE).

Table 1. Recent related work (continue)

No	References	Technique Name	Results Discussion
4	[14]	Random Forest, Bagging, and AdaboostM1	This study proposes a semi-supervised multi-layered clustering (SMLC) model for detecting and preventing network intrusion. SMLC may learn from partially labeled data and achieve detection performance comparable to IDPS based on supervised machine learning. The performance of SMLC on two datasets of the benchmark network-intrusion, NSL and Kyoto 2006, is compared to one of a well-known semi-supervised approach (tri-training) and the supervised ensemble ML models, particularly Random Forest, Bagging, and Adaboost.
5	[15]	K-Means algorithm and Hybrid Feature Selection	This study proposes an enhanced density-based initial cluster centers selection method after a Hadoop-based hybrid feature selection technique to find the most useful feature sets, in order to address the problem of outliers and local optimums.
6	[16]	Verification approach	The researchers present a new semi-supervised intrusion detection model that utilizes a verification strategy to produce consistent classifications across time, even when model updates are not available. Use semi-supervised learning to update the underlying machine learning models without the requirement for human interaction. The pool verifier, depending on the conclusion of the pool of classifiers, uses the classifications recognized by the verifier to determine whether it is reliable or not.
7	[17]	K-Nearest Neighbor (K-NN) and Artificial Neural Network (ANN)	This study proposes FloodDetector, an effective architecture for detecting known and unknown flooding assaults in SDN. It is a controller-agnostic SDN application that employs two machine learning classifiers to detect both known and unknown flooding attacks: K-nearest neighbor (K-NN) and artificial neural network (ANN).
8	[18]	Deep neural networks	To detect intelligent systems, this study proposes the use of machine learning frameworks. The study uses deep learning to distinguish between benign data exchange and harmful data traffic attacks.
9	[19]	The N-Gram line generation, feature selection algorithm, and SVM algorithm	This paper offers network traffic flow-based approach for mobile malware detection that assumes each HTTP flow as a document and analyzes HTTP flow requests using natural language processing string analysis. An effective malware detection model is created using the N-Gram line generation, feature selection method, and SVM algorithm.
10	[20]	DBSCAN, SVM, and Random Forest	In this paper, a hybrid supervised/unsupervised strategy is proposed. First, the clustering algorithm separates the anomalous traffic from the regular data by using numerous flow-based criteria. After determining the statistical characteristics each cluster shares, they can be assigned names using a categorization method. The authors conduct an evaluation of the proposed method by processing vast amounts of data.
#	Our Proposed	K-Means algorithm	1- Training and testing with CICIDS2017 dataset. 2- Proposed hybrid feature selection techniques 3- Produce DDoS and Normal centroids. 4- Evaluation.




REFERENCES

- [1] C. Douligeris and A. Mitrokotsa, "DDoS attacks and defense mechanisms: a classification," *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (IEEE Cat. No.03EX795)*, 2003, pp. 190-193, doi: 10.1109/ISSPIT.2003.1341092.
- [2] F. Zola, L. Seguro-Gil, J. L. Bruse, M. Galar, and R. Orduna-Urrutia, "Network traffic analysis through node behaviour classification: a graph-based approach with temporal dissection and data-level preprocessing," *Computers & Security*, vol. 115, p. 102632, April 2022, doi: 10.1016/j.cose.2022.102632.
- [3] K. Kalegele, K. Sasai, H. Takahashi, G. Kitagata and T. Kinoshita, "Four Decades of Data Mining in Network and Systems Management," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 10, pp. 2700-2716, 1 Oct. 2015, doi: 10.1109/TKDE.2015.2426713.
- [4] J. Han, J. Pei, and M. Kamber, "Data mining: Concepts and techniques," *Morgan Kaufmann*, vol. 10, pp. 559-569, 2006.
- [5] S. Karanam, "Curse of Dimensionality—A 'Curse' to Machine Learning," *Towards Data Science*, [Online], Available: <https://towardsdatascience.com/curse-of-dimensionality-a-curse-to-machine-learning-c122ee33bfeb> (Accessed Jun. 19, 2022).
- [6] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping multidimensional data*, pp. 25-71, 2006, doi: 10.1007/3-540-28349-8_2.
- [7] R. A. I. Alhayali, M. Aljanabi, A. H. Ali, M. A. Mohammed, and T. Sutikno "Optimized machine learning algorithm for intrusion detection," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 1, pp. 590-599, 2021, doi: 10.11591/ijeecs.v24.i1.pp590-599.
- [8] M. I. Kareem and M. N. Jasim, "DDoS Attack Detection Using Lightweight Partial Decision Tree algorithm," *2022 International Conference on Computer Science and Software Engineering (CSASE)*, 2022, pp. 362-367, doi: 10.1109/CSASE51777.2022.9759824.
- [9] M. I. Kareem and M. N. Jasim, "Fast and accurate classifying model for denial-of-service attacks by using machine learning," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 3, pp. 1742-1751, 2022, doi: 10.11591/eei.v11i3.3688.
- [10] M. I. Kareem and M. N. Jasim, "The Current Trends of DDoS Detection in SDN Environment," *2021 2nd Information Technology To Enhance e-learning and Other Application (IT-ELA)*, 2021, pp. 29-34, doi: 10.1109/IT-ELA52201.2021.9773744.
- [11] S. Lysenko, O. Savenko, and K. B-kova, "DDoS Botnet Detection Technique Based on the Use of the Semi-Supervised Fuzzy C-Means Clustering," in *ICTERI Workshops*, 2018, pp. 688-695.




- [12] M. Idhammad, K. Afdel, and M. Belouch, "Semi-supervised machine learning approach for DDoS detection," *Applied Intelligence*, vol. 48, no. 10, pp. 3193–3208, 2018, doi: 10.1007/s10489-018-1141-2.
- [13] W. Bhaya and M. EbadyManaa, "DDoS attack detection approach using an efficient cluster analysis in large data scale," *2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT)*, 2017, pp. 168-173, doi: 10.1109/NTICT.2017.7976110.
- [14] O. Y. A. -Jarrah, Y. A. -Hammdi, P. D. Yoo, S. Muhaidat, and M. Al-Qutayri, "Semi-supervised multi-layered clustering model for intrusion detection," *Digital Communications and Networks*, vol. 4, no. 4, pp. 277–286, 2018, doi: 10.1016/j.dcan.2017.09.009.
- [15] Y. Gu, K. Li, Z. Guo and Y. Wang, "Semi-Supervised K-Means DDoS Detection Method Using Hybrid Feature Selection Algorithm," in *IEEE Access*, vol. 7, pp. 64351-64365, 2019, doi: 10.1109/ACCESS.2019.2917532.
- [16] E. K. Viegas, A. O. Santin, V. V. Cogo and V. Abreu, "A Reliable Semi-Supervised Intrusion Detection Model: One Year of Network Traffic Anomalies," *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1-6, doi: 10.1109/ICC40277.2020.9148916.
- [17] S. Y. Khamaiseh, A. Al-Alaj and A. Warner, "FloodDetector: Detecting Unknown DoS Flooding Attacks in SDN," *2020 International Conference on Internet of Things and Intelligent Applications (ITIA)*, 2020, pp. 1-5, doi: 10.1109/ITIA50152.2020.9312310.
- [18] A. A. Ojugo and R. E. Yoro, "Forging a deep learning neural network intrusion detection framework to curb the distributed denial of service attack," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 2, p. 1498, 2021, doi: 10.11591/ijece.v11i2.pp1498-1509.
- [19] S. R. Akula, "Semi supervised machine learning approach for DDOS detection," *International Journal of Innovative Research in Education*, vol. 8, no. 1, pp. 27–35, 2021, doi: 10.18844/ijire.v8i1.6445.
- [20] M. Najafimehr, S. Zarifzadeh, and S. Mostafavi, "A hybrid machine learning approach for detecting unprecedented DDoS attacks," *Journal Supercomputing*, vol. 78, pp. 8106-8136, 2022, doi: 10.1007/s11227-021-04253-x.
- [21] I. Sharafaldin, A. Gharib, A. H. Lashkari, and A. Ghorbani, "Towards a reliable intrusion detection benchmark dataset," *Computer Science*, vol. 2018, no. 1, pp. 177–200, 2017, doi: 10.13052/JSN2445-9739.2017.009.
- [22] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," *4th International Conference on Information Systems Security and Privacy*, vol. 1, pp. 108–116, 2018, doi: 10.5220/0006639801080116.
- [23] "Canadian Institute for Cybersecurity," UNB, [Online], Available: <https://www.unb.ca/cic/datasets/ids-2017.html> (Accessed Jun. 16, 2022).
- [24] A. Ahmim, L. Maglaras, M. A. Ferrag, M. Derdour and H. Janicke, "A Novel Hierarchical Intrusion Detection System Based on Decision Tree and Rules-Based Models," *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2019, pp. 228-233, doi: 10.1109/DCOSS.2019.00059.
- [25] M. Alian and G. Al-Naymat, "Questions clustering using canopy-K-means and hierarchical-K-means clustering," *International Journal of Information Technology*, pp. 1–10, 2022, doi: 10.1007/s41870-022-01012-w.
- [26] U. Kumar, A. Dasgupta, L. S. N. V. V. Krishna and P. K. Chintakunta, "Towards Semi-supervised Tree Canopy Detection and Extraction from UAV Images," *Communications in Computer and Information Science*, vol. 1568, doi: 10.1007/978-3-031-11349-9_26.
- [27] R. Silipo, I. Adae, A. Hart, and M. Berthold, "Seven Techniques for Dimensionality Reduction," *Open for Innovation KNIME*, 2014.
- [28] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986, doi: 10.1007/BF00116251.

BIOGRAPHIES OF AUTHORS



Mahdi Nsaif Jasim    is Assistant 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 interest 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.



Methaq Talib Gaata    is Assistant professor at College of Science, Mustansiriyah University, Iraq. He Holds a PhD degree in Computer Science with specialization in information security. His research areas are information hiding, multimedia processing, biometrics, pattern recognition and computer networks. He worked as the Head of the Computer Science Department from 2016 to 2020. Currently, he is the Deputy Dean for Scientific Affairs and Postgraduate Studies at the College of Science. He received many scientific awards from the Iraqi Ministry of Higher Education and Scientific Research. He has supervised and co-supervised more than 25 masters and 5 Ph.D. students. He has authored or coauthored more than 30 publications in proceedings and journals, with 5 H-index. He can be contacted at email: dr.methaq@uomustansiriya.edu.iq.