TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 19, No. 2, April 2021, pp. 664~671
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v19i2.18325

An ensemble based approach for effective intrusion detection using majority voting

Alwi M. Bamhdi[1], Iram Abrar[2], Faheem Masoodi[3]
[1]Department of Computer Science, Umm Al-Qura University, Saudi Arabia
[2, 3]Department of Computer Science, University of Kashmir, India

Article Info

Article history:

Received Aug 26, 2020
Revised Sep 26, 2020
Accepted Oct 15, 2020

Keywords:

DoS
Ensemble
Intrusion detection system
Majority voting
Multi-layer perceptron
Particle swarm optimization

ABSTRACT

Of late, Network Security Research is taking center stage given the vulnerability of computing ecosystem with networking systems increasingly falling to hackers. On the network security canvas, Intrusion detection system (IDS) is an essential tool used for timely detection of cyber-attacks. A designated set of reliable safety has been put in place to check any severe damage to the network and the user base. Machine learning (ML) is being frequently used to detect intrusion owing to their understanding of intrusion detection systems in minimizing security threats. However, several single classifiers have their limitation and pose challenges to the development of effective IDS. In this backdrop, an ensemble approach has been proposed in current work to tackle the issues of single classifiers and accordingly, a highly scalable and constructive majority voting-based ensemble model was proposed which can be employed in real-time for successfully scrutinizing the network traffic to proactively warn about the possibility of attacks. By taking into consideration the properties of existing machine learning algorithms, an effective model was developed and accordingly, an accuracy of 99%, 97.2%, 97.2%, and 93.2% were obtained for DoS, Probe, R2L, and U2R attacks and thus, the proposed model is effective for identifying intrusion.

Corresponding Author:

Faheem Masoodi
Department of Computer Science
University of Kashmir, India
Email: masoodifahim@uok.edu.in

1. INTRODUCTION

Computer networks are ubiquitous, a modern-day reality with their extensive application in transferring sensitive or classified information between computing devices/networks. Data is the virtual gold and thus, attracts attention of illegitimate interest groups, who try to break in with attacks, sabotages. The technology becomes the victim of technology; as the internet magnifies the potential vulnerability and consequent theft of the data. Such a scenario warrants robust security mechanisms to ensure data is securely transferred between several intended entities while simultaneously protected from unauthorized access. Even though several security mechanisms such as encryption techniques and firewalls have been employed to avoid the attacks yet, security breaches are witnessing a spike. To counter such threats, it is crucial to have an IDS that adapts to the network security policies, allowing for an in-time remedial mechanism with no serious damage to the users as well as the system [1]. IDS is broadly grouped as misuse and anomaly detection [2]. In the former method, the gathered information from the network is compared with the database that contains the attack signatures, and if a match is found, the corresponding data is labeled as an attack. They have a high detection

rate, but cannot be used to detect new attack classes. Conversely, in anomaly detection, variation from the normal behavior is labeled as intrusion [3]. Nowadays several machine learning classifiers have been used to detect network intrusion yet individually they suffer from limitations that need to be addressed for the sake of an efficient IDS; which at the same the time is effective in terms of computations and as such a number of hybrid approaches have been proposed [4]. The rationale of the proposed research work was to build an ensemble-based IDS, which is efficient, stealth, robust, and has less training time.

Some of the latest research works in the field of IDS have been investigated, and this section presents an overview of the same. Recently, ML algorithms such as support vector machine (SVM) [5-7], decision trees [8, 9], multi-layer Perceptron [10], K-Nearest neighbor [11, 12], random forest [13, 14] has been employed to classify the data as normal or intrusive. Although single classifiers are extensively employed for intrusion detection, recent research works have revealed that ensemble-based approaches yield a better performance due to which ensemble is a preferred choice.

Clustering-based ensemble models are quite popular in the current works owing to their effectiveness in the detection of attacks. In [15], spectral clustering and deep neural network-based (SCDNN) were used to identify intrusion, and performance of the model was tested using KDD-99 and NSL-KDD. Although SCDNN outperformed other algorithms yet its parameters need to be optimized so that it could be efficiently used on a vast network for intrusion detection. Similarly, [16] proposed a clustering-based ensemble for enhancing the detection rate of IDS, where a filter-based information gain method was used to identify significant features. However, in [16], KDD-99 dataset was used in which redundant records are present which led to biased results whereas, in the current work, unbiased results were obtained as NSL-KDD was used. Furthermore, fewer features were used and nearly the same accuracy was attained. Likewise, a clustering-based ensemble was presented and accuracy of 91.0333% was obtained on the NSL-KDD dataset in [17]. Compared to it, our model is more diverse, and reasonably better results were obtained in terms of accuracy.

Random forest is another widely used ensemble technique. In [18], random forest (RF) has been used to detect probe attacks on KDD-99 dataset. The authors claimed that the proposed model has an average detection and false positive rate (FPR) of 99.85 percent and 0.052 percent, respectively, for the probe attack, which outperformed the other classification techniques. Conversely, their work lacks completeness as they did not consider other attack types present in the dataset, and hence, it cannot be ascertained that the model could be effectively used for IDS whereas these problems have been addressed in the current work.

Feature selection plays a vital role as far as intrusion detection is concerned. To optimize the same, an ensemble-based wrapper feature selection method was presented in [19]. For classification purposes, three classifiers, namely Bayesian network, Naïve Bayes, and J4.8, were used in combination, and based on majority voting final result were obtained. Despite promising results, their model was more focused on feature selection strategy rather than on a novel intrusion detection framework whereas the current work is directed towards building a competent IDS. Similarly, an ensemble of filter-based approaches was proposed in [20], which was specially designed for a cloud platform to detect DoS attack. An accuracy of 99.67% was achieved using majority voting on NSL-KDD. In [20], only DDOS attack in the network was addressed whereas in case of IDS it is imperative to detect various attacks so that a coherent model can be formed.

Optimizing the system performance is the key; and with that in mind, particle swarm optimization (PSO) has been widely used in recent research works. In [21], efficient IDS where time-varying chaotic PSO (TVCPSO) was presented used to identify key features. The system performance was evaluated on the NSL-KDD dataset using two classifiers, namely SVM and multiple criteria linear programming (MCLP). An accuracy of 96.88% and 97.84% was obtained in the case of TVCPSO-MCLP and TVCPSO-SVM, respectively. Further, an accuracy of 97.23% and 97.03% and a false alarm of 2.41% and 0.87% was attained for TVCPSO-MCLP and TVCPSO-SVM, respectively. However, their model was tested/trained on seventeen features extracted from the dataset whereas in our model only ten features were taken into consideration.

Even though several researches work [22-24] have been conducted in the domain of IDS yet there were several issues that need to be addressed. To tackle the same, in the current work a more diverse model was formed where several classifiers that were not used in the prior research works were used in combination to optimize the model performance. As such, a more robust model was obtained that could be effectively used for intrusion detection.

2. RESEARCH METHOD

In the current work, the results from several classifiers were combined using majority voting-based ensemble method to improve the detection rate. Decision tree was used to identify significant attributes so that the training time and the computational complexity of the data could be minimized. The data was then classified as normal or intrusive using one is to many approaches, and PSO was used for optimization. NSL-KDD dataset was employed to assess model performance. A general framework for ensemble-based IDS has been presented in Figure1.

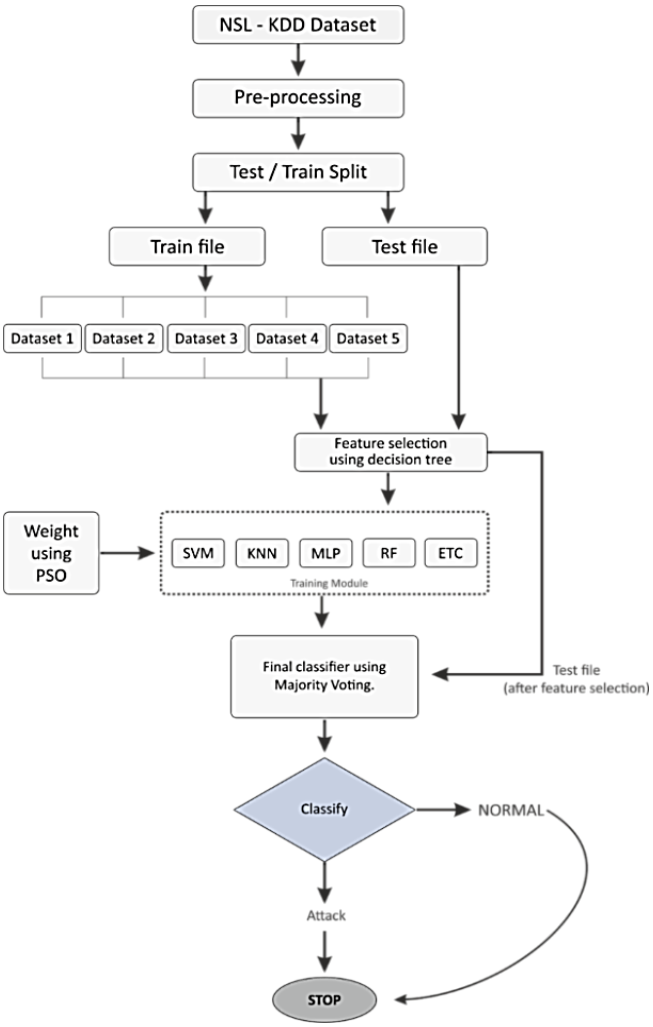

Figure 1. System framework

2.1. Dataset and pre-processing

In the proposed model, NSL-KDD has been employed, consists of 148517 records, which is partitioned as training (125973 records) and test dataset (22544 records) [25]. In the current work, the dataset was partitioned into five groups, and each one of them was trained/tested using different classifiers. Final results were obtained using majority voting. Pre-processing is a crucial step as it helps in the removal of statistical irregularities and the selection of important features from the data. Those features that are incredibly correlated to a given class distribution are considered to be significant, and in the current work, decision tree has been used for feature selection. Prior studies in this field indicate that a proper set of attributes can enhance system performance by reducing the training time [26]. The dataset was initially partitioned into subsets, and the irrelevant features present in NSL-KDD were removed. Among 41 attributes, only ten most significant features were taken into account depending on which the model was further trained/ tested. The decision tree selects the best attributes from a given set, according to which the data is partitioned into classes. Moreover, generalization accuracy and excellent real-time performance of decision trees can be used to detect new attacks [27].

2.2. Classifiers

Individual classifiers that are used to assess the model performance have their flaws that lead to compromise the performance of the system. These issues need to be addressed so that the security of the system is not compromised, and more reliable services are provided to the users. For this purpose, the ensemble-based approach can be employed, which is a combination of individual classifiers and is a more flexible and powerful tool as far as detection of network intrusion is concerned. An excellent ensemble is the one that maximizes the difference between the base classifier, and this can be achieved by dividing the dataset into subsets [28]. The process used in the model can be outlined in the subsequent steps.

- Divide the dataset into several datasets so that the difference between the various base classifiers is maximized, and thus, a better ensemble can be formed.
- Combine the results of base classifiers after completion of the training phase so that the final classifier is obtained.
- The output obtained in the training phase is then used to test the data, and accordingly, the data can be classified as normal or intrusive.

2.2.1. Support vector machine (SVM)

SVM was proposed by Vapnik (1995) [29]. It is widely used to solve regression and classification problems by creating a number of hyper-planes, and among them, the optimal hyper-plane is chosen based on the maximum separation between the classes. SVM can effectively be used for the classification of both linear and non-linear data [30]. In current work, radial linear basis kernel (RBF) function was used as it works well for the non-linear data. RBF uses gamma parameter to classify the data which in the current work is calculated as the reciprocal of the features.

2.2.2. Random decision forests (RF)

RF is a combination of decision trees that is used to boost the effectiveness of the model. In RF, bagging is employed to split the data into various subsets, and then decision trees are constructed using these subsets [31]. RF does not suffer from over-fitting problem and has low classification errors [32]. Bootstrap samples from the dataset were used to construct individual trees in the forest. Gini impurity measure was applied to choose the best node for the split and maximum number of trees were set to 25 based on which the performance was evaluated.

2.2.3. K-nearest neighbor (KNN)

KNN is a supervised classifier. Initially, the value of K is selected, and then Euclidean distance is computed among the various data points based on which the data is divided into K-number of clusters. The data points that have a minimum distance between them are placed in one cluster as they have similar characteristics [33]. KNN is easy to implement and works well for large datasets [34]. In the current work, the model performance was evaluated on K=5 and the Euclidean distance measure was used to find the nearest neighbors. Prior to this, the dataset was normalized to reduce variations among various attributes.

2.2.4. Particle swarm optimization (PSO)

PSO was presented by Kennedy and Eberhart [35]. It mimics the behavior of a flock of birds and uses the same principle to direct the particles to explore the optimal global solution. PSO are comparatively easier to implement as compared to genetic algorithms as they do not have evolutionary operators [11, 36]. In the current work, one of training data subset comprising of 4455 instances was used as an input to the PSO and the weights in form of an array were obtained and accordingly the data was classified.

2.2.5. Extra-tree classifier (ETC)

ETC is an ensemble-based approach which is primarily based on decision trees and are used to induce variations by changing the approach in which trees are built. Owing to its randomized characteristics, they are computationally less expensive compared to that of random forest and, thus, can be efficiently used for intrusion detection. In ETC, square root of the total number of features was taken into consideration to choose the best node for split whereas Gini impurity measure was used to determine the quality of the same.

2.2.6. Multilayer perceptron (MLP)

It is a neural network comprising of interconnected nodes or neurons that are used to map the inputs with the corresponding output vectors. The nodes have weights assigned with them, and the output is computed as the function of the summation of inputs to a node, subjected to a non-linear function [37]. As neural networks have flexible time-delaying abilities, they can be efficiently used for network intrusion detection as they decrease false alarm rate. To achieve the desired results, a rectified linear unit activation function was used in the hidden layer. Moreover, a regularization value of 0.0001 was added to loss function to prevent model overfitting.

2.2.7. Majority voting

In majority voting, the dataset is partition into several sub-sets and various classifiers are used to train the same, thereby accomplishing the learning phase. The final results of allocating a label to the sample depends on the maximum number of votes received in favor of a particular class obtained from different classifiers. In the current work, weighted majority voting was used to classify the data where PSO was employed for assigning weights to various classifiers. Moreover, prior studies on IDS [38, 39] have revealed

that voting can be used to improve the performance of an IDS considerably, and it generally outperforms other ensemble techniques.

3. RESULTS AND ANALYSIS

In the proposed model, several classifiers viz, SVM, ETC, KNN, MLP, and RF were used for evaluating the model on NSL-KDD dataset. Accuracy of the model is predicted using precision, F1 score, support, and recall. In the current work, as NSL-KDD dataset was divided into five subsets, the result of the same has been depicted Tables 1 and 2. From the experimental results depicted in Figure 2, it was concluded that the performance of the model was 99%, 97.2%, 97.2%, and 93.2% for DoS, Probe, R2L, and U2R, respectively. Likewise, model accuracy for five data subsets were 98.76%, 98.80%, 98.82%, 98.77%, and 98.79%, respectively yielding an average model accuracy of 98.788%. The results were least promising for U2L attack owing to their lack of entries in NSL-KDD dataset.

Table 1. Experimental results of training using majority voting

Expert	Parameters	Normal	DoS	Probe	R2L	U2R
Dataset 1	Accuracy	99%	99%	98%	97%	85%
	Recall	100%	99%	95%	95%	47%
	F1- score	99%	99%	97%	96%	61%
	Support	8082	11547	2087	525	36
Dataset 2	Accuracy	100%	99%	98%	98%	100%
	Recall	100%	99%	96%	97%	44%
	F1- score	99%	100%	97%	97%	61%
	Support	11427	8063	2141	607	39
Dataset 3	Accuracy	99%	99%	98%	98%	92%
	Recall	100%	100%	95%	96%	30%
	F1- score	99%	100%	97%	97%	45%
	Support	11700	7950	2034	556	37
Dataset 4	Accuracy	99%	99%	98%	97%	93%
	Recall	100%	100%	96%	96%	31%
	F1- score	99%	99%	97%	96%	46%
	Support	11704	7886	2065	580	42
Dataset 5	Accuracy	99%	99%	97%	98%	100%
	Recall	99%	100%	95%	96%	34%
	F1- score	99%	99%	96%	97%	51%
	Support	11505	8066	2104	568	35

Table 2. Experimental results of testing using majority voting

Expert	Parameters	Normal	DoS	Probe	R2L	U2R
Dataset 1	Accuracy	99%	98%	96%	94%	73%
	Recall	99%	99%	93%	92%	38%
	F1- score	99%	99%	94%	93%	50%
	Support	13337	19170	3646	913	63
Dataset 2	Accuracy	99%	98%	96%	95%	100%
	Recall	99%	99%	92%	94%	27%
	F1- score	99%	99%	94%	94%	42%
	Support	19170	13337	3646	913	63
Dataset 3	Accuracy	99%	98%	96%	95%	100%
	Recall	99%	99%	92%	93%	29%
	F1- score	99%	99%	94%	94%	44%
	Support	19170	13337	3646	913	63
Dataset 4	Accuracy	99%	98%	96%	95%	86%
	Recall	99%	99%	92%	93%	29%
	F1- score	99%	99%	94%	94%	43%
	Support	19170	13337	3646	913	63
Dataset 5	Accuracy	99%	98%	97%	96%	95%
	Recall	99%	99%	93%	94%	30%
	F1- score	99%	99%	95%	95%	46%
	Support	19170	13337	3646	913	63

Analyzing the results, it can be concluded that:

- Firstly, promising results are obtained for these algorithms owing to their capability of handling non-uniformly distributed data. Exceptional generalization property of these classifiers, their ability to inspect the outcomes, result in minimizing the cost, thereby, making the proposed model a positive candidate for network intrusion detection as they are able to reduce the misclassified data.

- Secondly, additional resources are used by IDS as they continuously monitor the network to detect the presence of intrusive activities leading to wastage of resources. However, feature selection method was employed in the current work to minimize resource consumption, thereby reducing the complexity of data.
- Thirdly, IDS cannot rely on the results of a single classifier as the intruders can examine its loopholes making it prone to tampering thus, resulting in probable immobilizing of the system. However, the current model does not rely on a single classifier but combines the opinion of several classifiers, therefore ensuring that a robust model for intrusion detection is formed.

Overall performance of the proposed ensemble-based majority voting was good as the final result depends on the agreement of several classifiers. Predominantly promising results were obtained for DoS attacks. Thus, the proposed model is reliable and can be effectively used for network intrusion detection.

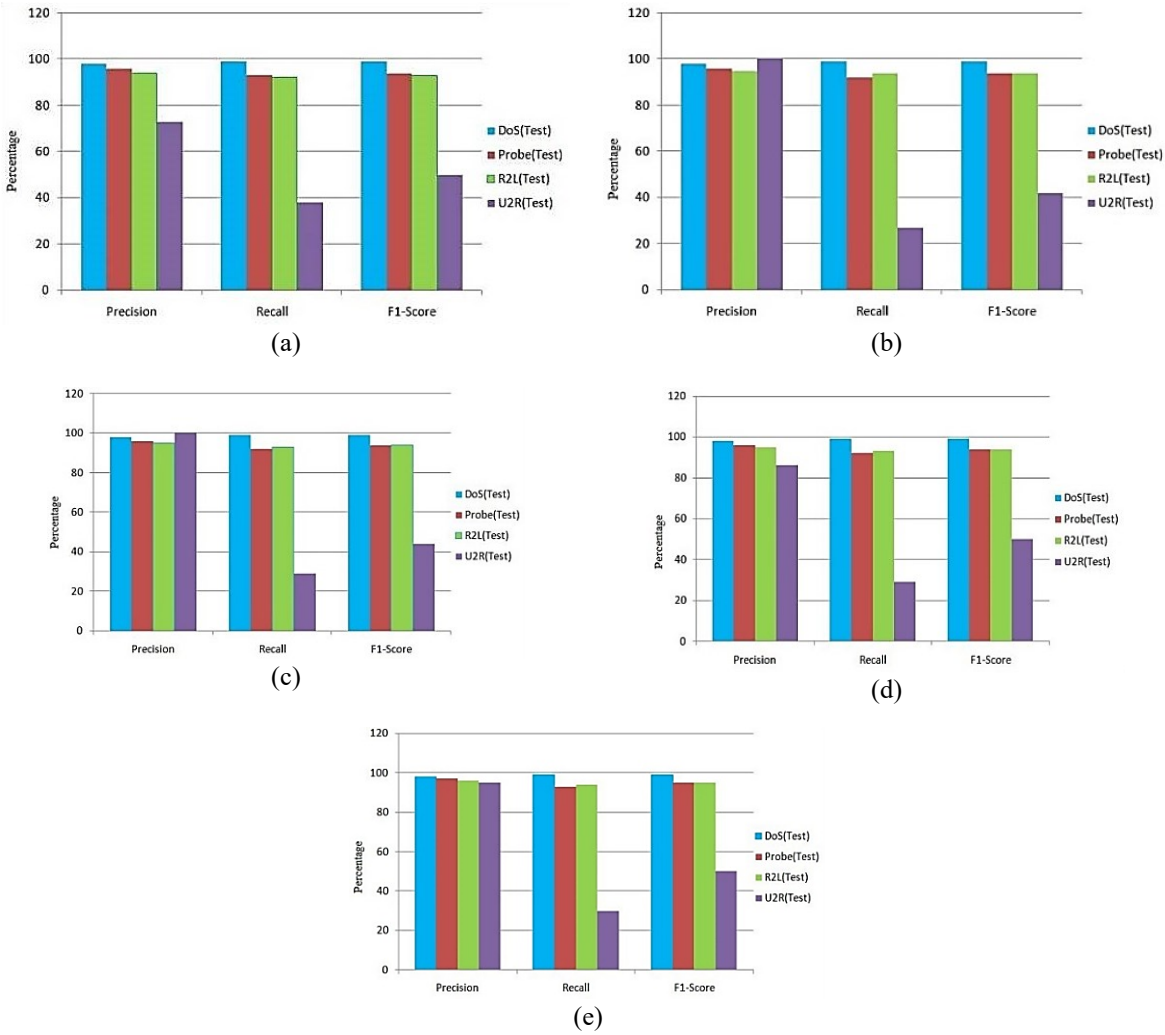

Figure 2. (a) Dataset 1, (b) Dataset 2, (c) Dataset 3, (d) Dataset 4, (e) Dataset 5, graphical representation of results for 5 datasets during testing phase

4. CONCLUSION

The latest ensemble-based ML algorithms were analyzed to present a summary of the work conducted in the domain of IDS. The rationale of current work is to offer a robust IDS and for the same ensemble-based technique was employed. Several classifiers were used for training/testing, and final results were achieved using the voting approach. PSO was employed for improving the model performance. The final results reflect excellent performance. The proposed model, as such, could thus be effectively used for network intrusion detection. There is scope to improve the feature selection, i.e., ensemble techniques can be employed to select significant attributes so that the feature selection technique can be optimized. Moreover, as machine learning algorithms have certain vulnerabilities, attackers can use the same to launch the attacks, and thus, there is a need to focus in this direction and design a classifier that can withstand the relevant security issues.

REFERENCES

[1] Y. Zhou, G. Cheng, S. Jiang and M. Dai, "Building an efficient intrusion detection system based on features selection and ensemble classifier," *Computer networks*, vol. 174, pp. 1-17, 2020.

[2] L. Lv, W. Wang, Z. Zhang and X. Liu, "A novel intrusion detection system based on optimal hybrid kernel extreme learning machine," *Knowledge- based systems*, vol. 195, pp. 1-17, 2020.

[3] H. Alazzam, A. Sharieh and K.E. Sabri, "A feature selection algorithm for intrusion detection system based on pigeon inspired optimizer," *Expert systems with applications*, vol. 148, pp.1-14, 2020.

[4] H. Rajadurai and U.D. Gandhi, "A stacked ensemble learning model for intrusion detection in wireless network," *Neural computing and applications*, pp.1-9, 2020.

[5] D. G. Kalita, V. P. Singh and V. Kumar, "SVM Hyper-parameters optimization using multi-PSO for intrusion detection," *Social networking and computational intelligence. Lecture notes in networks and systems*, vol. 100, pp. 227-242, 2020.

[6] F. Salo, A. B. Nassif and A. Essex, "Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection," *Computer Networks*, vol. 148, pp. 164-175, 2019.

[7] M. Safaldin, M. Otair and L. Abualigah, "Improved binary gray wolf optimizer and SVM for intrusion detection system in wireless sensor networks," *Journal of ambient intelligence and humanized computing*, pp. 1-18, 2020.

[8] M.A. Ferrag, L. Maglaras, A. Ahmim, M. Derdour and H. Janicke, "RDTIDS: Rules and decision tree- based intrusion detection system for internet of things networks," *Future internet*, pp.1-14, 2020.

[9] I.H. Sarker, Y.B. Abushark, F. Alsohami and A.I. Khan, "IntruDTree: A machine learning based cyber-security intrusion detection model," *Symmetry*, pp. 1-15, 2020.

[10] K. Chisholm, C. Yakopcic, M. S. Alam and T. M. Taha, "Multilayer perceptron algorithms for network intrusion detection on portable low power hardware," in *computing and communication workshop and conference, 2020. CCWC 2020.Tenth Annual IEEE*, Las Vegas, USA, 2020, pp. 901-906.

[11] T.B. Prasad, P.S. Prasad and K.P. Kumar, "An intrusion detection system software program using KNN nearest neighbours approach," *International journal of science research and innovation engineering*, pp.1-6, 2020.

[12] S. Rajagopal, P. P. Kundapur and K.S. Hareesha, "A stacking ensemble for network intrusion detection using heterogeneous datasets," *Secure Communication Networks*, pp.1-9, 2020.

[13] W. Wang, Y. Li, X. Wang, J. Liu and X. Zhang, "Detecting android malicious apps and categorizing benign apps with ensemble of classifiers," *Future Generation Computing System*, vol. 78, pp. 987- 94, 2018.

[14] C. Ambikavathi and S.K. Srivatsa, "Predictor selection and attack classification using random forest for intrusion detection," Journal *of scientific and industrial research*, vol. 79, pp. 365-68, 2020.

[15] T. Ma, F. Wang, J. Cheng, Y. Yu and X. Chen, "A hybrid spectral clustering and deep neural network ensemble algorithm for intrusion detection in sensor networks," *Sensors*, vol. 16, pp. 1-23, 2016.

[16] M. Gudadhe, *et al.*, "A new data mining-based network intrusion detection model," in *Computing Commununication technology*, *2010. ICCCT 2010. Tenth International IEEE Conference,* Allahabad, Uttar Pradesh, 2010, pp. 731-735.

[17] W. Chen, F. Kong, F. Mei, G. Yuan and B. Li, "A novel unsupervised anomaly detection approach for intrusion detection system," in *Big Data Security on cloud*, *2017. Third International IEEE Conference,* Beijing, 2017, pp. 69-73.

[18] A. J. Malik, *et al.*, "Binary PSO and random forests algorithm for probe attacks detection in a network," in *congress of evolutionary computation, 2011. CEC 2011. IEEE*, New Orleans, LA, 2011, pp. 662-668.

[19] N.F. Haq, A.R. Onik and F.M. Shah, "An ensemble framework of anomaly detection using hybridized feature selection approach (HFSA)," *SAI Intelligent System Conference*, London, 2015, pp. 989-995.

[20] O. Osanaiye, *et al.*, "Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing," *Eurasip Journal on Wireless Communications and Networking*, vol. 130, pp. 1-10, 2016.

[21] S. M. H. Bamakan, *et al.*, "An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization," *Neurocomputing*, vol. 199, pp. 90- 102, 2016.

[22] S. Laqtib, K. E. Yassini and M. L. Hasnaoui, "A technical review and comparative analysis of machine learning techniques for intrusion detection systems in MANET," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 3, pp. 2701-2709, 2020.

[23] K. Farnaha, M. Rahman and M.T. Ahmad, "An intrusion detection system for packet and flow-based networks using deep neural network approach," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 10, no. 5, pp.5514-5525, 2020.

[24] I. Abrar, *et al.*, "A machine learning approach for intrusion detection system on NSL-KDD dataset," in *smart electronics and communication, 2020. ICOSEC 2020. IEEE,* Trichy, India, 2020, pp. 919-924.

[25] H. Chae, B. Jo, S. Choi and T. Park, "Feature selection for intrusion detection using NSL-KDD," *Recent Advances in Comput Science*, pp. 184-187, 2013.

[26] N.K. Kanakarajan and K. Muniasamy, "Improving the accuracy of intrusion detection using GAR-forest with feature selection," in *Proceedings of 4th International Conference on Frontiers in Intelligent Computing: Theory Applications. Advances in Intelligent Systems and Computing*, vol. 404, 2016, pp. 539-547.

[27] S. Peddabachigari, A. Abraham, C. Grosan and J. Thomas, "Modeling intrusion detection system using hybrid intelligent systems," *Journal of Network and Computer Applications*, vol. 30, pp. 114-132, 2007.

[28] J. Gu, L. Wang, Y. Chunga and S. Wang, "A novel approach to intrusion detection using SVM ensemble with feature augmentation," *Computer and security*, vol. 86, pp. 53-62, 2019.

[29] F. Kuang, W. Xu and S. Zhang, "A novel hybrid KPCA and SVM with GA model for intrusion detection," *Applied Soft Computing*, vol. 18, pp. 178-184, 2014.

[30] I. Ahmad, M. Basheri, M.J. Iqbal and A. Rahim, "Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection," *IEEE Access*, vol. 6, pp. 33789-33795, 2018.
[31] Q. Zhoua, H. Zhoua and T. Lib, "Cost-sensitive feature selection using random forest: Selecting low-cost subsets of informative features," *Knowledge-based systems*, vol. 95, pp. 1-11, 2016.
[32] N. Farnaaz and M.A. Jabbar, "Random Forest Modeling for Network Intrusion Detection System," *Procedia Computer Science*, vol. 89, pp. 213- 217, 2016.
[33] W. Li, P. Yi, Y. Wu, L. Pan and J. Li, "A new intrusion detection system based on KNN classification algorithm in wireless sensor network," *Journal of Electronic and Computer Engineering*, pp. 1-8, 2014.
[34] M. Nikhitha and M.A Jabbar, "K Nearest neighbor-based model for intrusion detection system," *International Journal of Recent Technology and Engineering*, vol. 8, pp. 2258-2262, 2019.
[35] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Sixth International Symposium on Micro Machine and Human Science,* vol. 1, New York, 1995, pp. 39-43.
[36] S. Singh and A.K. Singh, "Web spam feature selection using CFS-PSO," *Procedia computer science,* vol. 125, pp. 568-575, 2018.
[37] Y. Yao, *et al.*, "Anomaly intrusion detection using hybrid MLP/CNN neural network," in *intelligent system design and applications, 2006. ISDA 2006. Sixth international conference IEEE,* Jinan, 2006, pp.1-6.
[38] M. Panda and M.R. Patra, "Ensemble voting system for anomaly-based network intrusion detection," *International Journal of recent trends in engineering and research*, vol. 2, pp.8–13, 2009.
[39] J. Lueckenga, *et al.*, "Weighted vote algorithm combination technique for anomaly based Smart Grid intrusion detection systems," in *International joint conference on Neural Networks, 2016. IJCNN,* Vancouver, 2016, pp. 2738-2742.

BIOGRAPHIES OF AUTHORS

Dr. Alwi M. Bamhdi, is an assistant Professor in the Department of computer sciences, Umm Al-Qura University, Saudi Arabia. He received his MSc and Ph.D. in Computer Science in 2014 from Heriot-Watt University, UK. His research interests include Mobile Ad Hoc networks, Wireless Sensor Networks, Information security, Internet of things, Cyber security, Computer Vision and simulation and performance evaluation.

Ms. Iram Abrar is currently persuing M.tech in the Department of Computer Science, University of Kashmir. Prior to this, she has completed her B.tech in computer science and engineering from Islamic university of science and technology. Her basic research interests include network security, IoT and machine learning.

Dr. Faheem Syeed Masoodi is currently working as an Assistant Professor in the Department of Computer Science, University of Kashmir. Earlier, he served College of Computer Science, University of Jizan, Saudi Arabia as an Assistant Professor. Prior to that, he performed his duties as a Research Scientist at NMEICT-Edrp project sponsored by Ministry of HRD, Govt. of India in 2015. He was awarded PhD in the domain of Network Security & cryptography by the Department of Computer Science, Aligarh Muslim University India in year 2014 and did his masters in computer sciences from University of Kashmir. His basic research interests include Cryptography & Network Security; and Internet of things (IOT).