Modeling of Deep Learning based Intrusion Detection System in Internet of Things Environment

Mohammad Hammoudeh [1], Saeed M. Aljaberi [2]

[1] Manchester Metropolitan University, UK

[2] Artificial Intelligence department, Dubai Police, Dubai, UAE

Emails: m.hammoudeh@mmu.ac.uk ; eljabri@live.com

Abstract

The Internet of Things (IoT) has become a hot popular topic for building a smart environment. At the same time, security and privacy are treated as significant problems in the real-time IoT platform. Therefore, it is highly needed to design intrusion detection techniques for accomplishing security in IoT. With this motivation, this study designs a novel flower pollination algorithm (FPA) based feature selection with a gated recurrent unit (GRU) model, named FPAFS-GRU technique for intrusion detection in the IoT platform. The proposed FPAFS-GRU technique is mainly designed to determine the presence of intrusions in the network. The FPAFS-GRU technique involves the design of the FPAFS technique to choose an optimal subset of features from the networking data. Besides, a deep learning based GRU model is applied as a classification tool to identify the network intrusions. An extensive experimental analysis takes place on KDDCup 1999 dataset, and the results are investigated under different dimensions. The resultant simulation values demonstrated the betterment of the FPAFS-GRU technique with a higher detection rate of 0.9976.

Keywords: IoT, Security, intrusion detection, Feature selection, Deep learning, KDDCup 1999 dataset

1. Introduction

The Internet of Things (IoT) has been presented initially by British researcher Kevin Ashton in 1999, whereas they defined a scheme which can enable physical objects to be linked to the Internet through various sensor nodes [1, 2]. The IoT could be determined as a network of connected devices which could receive and send information when it is in a dynamic/static state [3]. IoT device collects information through devices, like sensor nodes and RFID tags, to a certain environment/event for providing smart solutions for distinct problems. It became promising due to the exponential increase of techniques like innovative data analyses algorithm, wireless transmission, and cloud computing [4]. Hence, the IoT was employed in many features like smart homes, industries, smart vehicles, and healthcare, for putting them on networks and digitalize them. Each gathered information could be interchanged among each party in the IoT networks; e.g., the information could be interchanged among a device and human, and another reasonable framework [5].

Over the past decades, the world has observed a remarkable transformation of AI and its application in many fields [6-8]. It has permitted the formation of actual functioning system for several workloads which is deliberated to be a kind of fiction in the previous years. Now, it is promising to implement and develop an AI scheme which identifies persons in public and airport places and links their activity to their whole life recording information. Further, AI assisted security system provides the capability for identifying

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021

social activity and triggers promising alarm if suspected behaviour takes place. Intrusion prevention and detection system monitor the device, network, and system for malicious activity and policy violation is the important counter measure against cyberattacks.

Using an extensive spectrum, the prevention and detection schemes differ from anti-virus software to classified systems observing the traffics of a whole support network. Generally, IDSs are classified into 2 sets [9-11], i.e., signature based detections (malicious pattern is previously recognized) and anomaly based detections (no pattern is provided). Conventional systems and methods can be failed directly appropriate to advanced computing infrastructure and paradigm as abovementioned. New intrusion prevention and detection systems and algorithms are required for catering the novel computing framework and recently developing cyber threats and attacks, considering the aspects like computing environment heterogeneity, algorithmic scalability, complexity, and data diversity [12-14].

Wen [15] proposed an approach of CC-IDS technique based BPNN, involving BPNN algorithms, NN-CC-IDS, and ABC method, that is employed for conducting CC-IDS experiments according to BPNN method; Proposes an ABC-NN method; design a CC-IDS according to BPNN model. In [16], the interrupted information has been classified and detected using ML through MapReduce method. The main aim of Hadoop MapReduce method is to decrease the range of database's ideal weights which is determined by reducer's method. Next, employing DT classifiers in detecting information. These DT classifiers utilize suitable classifiers for deciding the class label for nonhomogeneous leaves node. The DT fragments provide rough sections although the leaves level classifiers produced information regarding the quality which effect the labels inside part.

Singh and Ranga [17] discussed an efficient network based IDS using an ensemble based ML method with 4 classifications viz., bagged tree, Boosted tree, RUSBooted, voting scheme, and subspace discriminant. The voting algorithms are integrated into the architecture of obtaining combined last predictions. Krishnaveni et al. [18] proposed an efficient IDS for CC platform with classification and ensemble FS methods. This technique is based on the ensemble FS techniques, i.e., employed to the election of valued reduce featured set from the provided intrusions dataset. Whereas the ensemble classifier could capably merge an individual classifier for producing strong classifiers with the voting techniques. An ensemble based technique efficiently classify the network traffics behaviour as attack/normal. In [19], an effective IDS is intended for solving the problems which aren't favorably influence the economic advancement of the CC and improve the security of the CC from malicious attacks. The IDSs are modelled by the presented FDHO based DRN for detecting networks intrusion. But, the presented FDHO algorithms are proposed with incorporating FAT using DHOA, correspondingly. Furthermore, the recognition of malicious attacks is performed by a DRN considerably increase the training speed, reduce the computation difficulty, and generate efficient detecting result.

This study designs a novel flower pollination algorithm (FPA) based feature selection with gated recurrent unit (GRU) model, named FPAFS-GRU technique for intrusion detection in the IoT platform. The proposed FPAFS-GRU technique is mainly designed to determine the presence of intrusions in the network. The FPAFS-GRU technique involves the design of FPAFS technique to choose an optimal subset of features from the networking data. Besides, a deep learning based GRU model is applied as a classification tool to identify the network intrusions. An extensive experimental analysis takes place on KDDCup 1999 dataset and the results are investigated under different dimensions.

2. Design of FPAFS-GRU Technique

In this paper, a novel FPAFS-GRU technique is presented for intrusion detection in the IoT platform. The FPAFS-GRU technique involves two major processes. At the first stage, the FPAFS technique is designed to choose an optimum subset of features. Then, in the second stage, the chosen features are fed into the GRU model to perform classification process.

2.1 Process involved in FPAFS Technique

Primarily, the FPAFS technique is designed to choose an optimum subset of features. The biotic pollination, cross-pollination, abiotic pollination, and self-pollination components are explained in a field optimized and induced in flower pollination approach [20]. The pollination tasks enclose an order of

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021

tedious processes from plant generation rules. The flower and equivalent pollen gametes are inclined to give a consistent solution to the optimized problem. The flower constancy is obvious as accurate solution that could be perceptible. Against global pollinations, the pollinator transfers pollen in longer distances to maximum fitting. At others, local pollination has been approved to inside a lesser area of limited flower is performed in shading water. Global pollination occurs with possibilities that are called switch probabilities. If the phase is unconcerned, the local pollination is interchanged. In FPA technique, there are 4 principles as provided in the following:

- The live and cross pollinations are named global pollination and the carry of pollen pollinators implement the levy fight.
- Abiotic and self-pollination were mentioned as local pollinations.
- Pollinators are insects that are able of evolving flower constancy. It can be determined as production probabilities to two implemented flowers.
- The communication of global and local pollinations is controlled to switch possibilities.

Therefore, the 1st and 3rd rules are demonstrated as:

$$x_q^{t+1} = x_p^t + \gamma \times L(\lambda) \times \left(g_* - x_p^t\right) \quad (1)$$

where $x_p^t =$ pollen vector at iteration t; g_* implies the present optimum solution in another present creating outcome; $\gamma =$ a refers the scale factor for managing step size, and L represents strength of pollination which is connected with step size of levy distributions. The levy flight has been determined as the group of arbitrary calculations that is the length of all jumps which execute levy probability distributions function through infinite change. Next, L signifies the levy distributions as given:

$$L \sim \frac{\lambda \times \Gamma(\lambda) \times \sin\frac{\pi\lambda}{2}}{\pi} \times \frac{1}{S^1 + \lambda} S \gg S_0 0, \quad (2)$$

where $\Gamma(\lambda) =$ typical gamma function. In the case of local pollination, the 2nd and 3rd rules are demonstrated as:

$$x_p^{t+1} = x_p^t + \varepsilon\left(x_q^t - x_k^t\right) \quad (3)$$

where x_q^t and $x_k^t = 2$ pollens in diverse flowers to same plants. During arithmetic format, if x_q^t and x_k^t come in the same species were elected in homogeneous populations that are demonstrated as local arbitrary walk and ε have composed to a uniform distribution from 0 and 1.

At this point, FS has been demonstrated as binary optimized problem in which the searching agents are limited with binary {0, 1} values. At this time, every solution endures characterized as 1D vector is the length of vector that is dependent upon value attribute from dataset. All the cells of vector have 2 values, (1 or 0), where value 1 illustrates the neighboring attribute as elected, but zero demonstrated that attributes as non-selected. The FS problem is considered as multi-objective optimized problem in which two contrast aims that existed reached; selecting lesser amount of features and superior classifier accuracy. For resolving the multi-objective problem, two binary optimized techniques are projected. During FS problem, the solution has been termed as optimum one that is collected to minimal amount of features together with maximal classifier accuracy. The purpose is for finding the balance amongst amount of attributes and classifier accuracy; the Fitness Function (FF) in Eq. (4) has been utilized from optimized techniques for estimating these solutions.

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021

$$Fitness = \alpha \Upsilon_R(D) + \beta \frac{|R|}{|N|} \quad (4)$$

where $\Upsilon_R(D)$ represents the classifier error rate. So, |R| represents the cardinality of elected feature subsets but $|N|$ illustrates the entire features in actual datasets, α and β are 2 parameters to the importance of ordering quality as well as subset length, $\alpha \in [0,1]$ and $\beta = (1 - \alpha)$ implemented during the study.

2.2 Process involved in GRU Technique

During classification process, the chosen features are fed into the GRU model to allot proper class labels [21]. The LSTM has been determined as time recurrent neural network (RNN). It can be comprised of cell, input, output, and forget gates. The unit is valued at various time intervals, and three gates were utilized from controlling the data flow. The forget gate is accomplished utilizing a simple 1-layer NN. The activation of gate has been evaluated as implementing in Eq. (5).

$$f_t = \sigma\big(W[x_t, h_{t-1}, C_{t-1}] + b_f\big) \quad (5)$$

At this time, x_t indicates the input order; h_{t-1} illustrates the present block output; C_{t-1} refers the standard LSTM block memory; b_f showcases the bias vector; W signifies the separate weight and logistic sigmoid functions. An input gate has new memory established by simple NN together with tanh activation functions and presenting memory block effects. The task has been estimated in Eqs. (6) and (7):

$$i_t = \sigma(W[x_t, h_{t-1}, C_{t-1}] + b_i) \quad (6)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh([x_t, h_{t-1}, C_{t-1}]) + b_c \quad (7)$$

The output gate has been included with outcome of existing LSTM block. It can be defined as implementing in Eqs. (8) and (9):

$$\sigma_t = \sigma(W[x_t, h_{t-1}, C_{t-1}] + b_o) \quad (8)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (9)$$

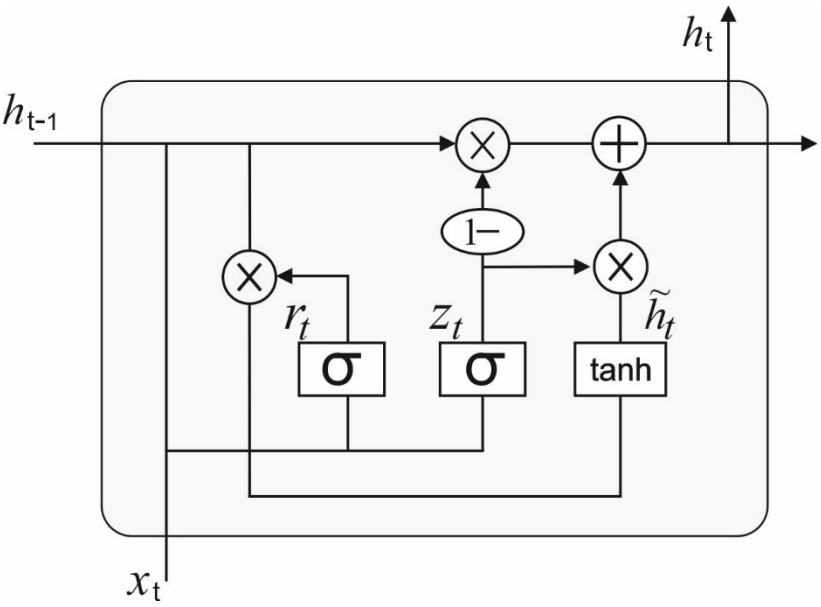

Fig. 1. Structure of GRU

The gated recurrent unit (GRU) is an alternative technique of gate depends on recurrent unit which comprises minimal framework and same function of LSTM unit, as shown in Fig. 1. The GRU has 2 gates: reset and upgrade. $r^{(t)}$ and $z^{(t)}$ represents the measure of reset as well as upgrade gates at time step t individually. $x_i \in R^n$ depicts the 1D input vectors for GRU block at time step t. $\tilde{h}^{(t)}$ demonstrates

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021

the outcome candidate of GRU blocks. $h^{(t-1)}$ indicates the recurrent GRU block outcome of time step $t-1$ and the present outcome at time t is $h^{(t)}$. This function is evaluated as:

$$z^{(t)} = \sigma\left(W_z x^{(t)} + U_z h^{(t-1)} + b_z\right) \quad (10)$$

$$r^{(t)} = \sigma\left(W_r x^{(t)} + U_r h^{(t-1)} + b_r\right) \quad (11)$$

$$\tilde{h}^{(t)} = tanh\left(W x^{(t)} + U(r^{(t)} \odot h^{(t-1)}) + b\right) \quad (12)$$

$$h^{(t)} = (1 - z^{(t)}) \odot h^{(t-1)} + z^{(t)} \odot \tilde{h}^{(t)} \quad (13)$$

where W_z, W_r, and W are feedforward weights and U_z, U_r, and U are RNN weight of upgrade, reset, and resultant candidate activations. b_z, b_r and b are biases of these gate and resultant candidate activations $\tilde{h}^{(t)}$ respectively.

3. Experimental Validation

This section offers a detailed result analysis of the FPAFS-GRU technique on the applied KDDCup 1999 dataset. It includes a set of 125973 instances with 41 features under two classes.

Fig. 2 investigates the FS results of the FPA-FS technique with other FS models interms of best cost (BC). The results portrayed that the PSO-FS technique has accomplished worse outcomes with the maximum BC of 0.009139. In line with, the GA-FS technique has resulted to certainly improved results with the BC of 0.008150. Followed by, the BGOA-FS technique has offered moderate outcome with the BC of 0.006530. Moreover, the CBOA-FS technique has accomplished near optimal BC of 0.002345 and 0.006530. However, the FPA-FS technique has reached superior results with the least BC of 0.002289.

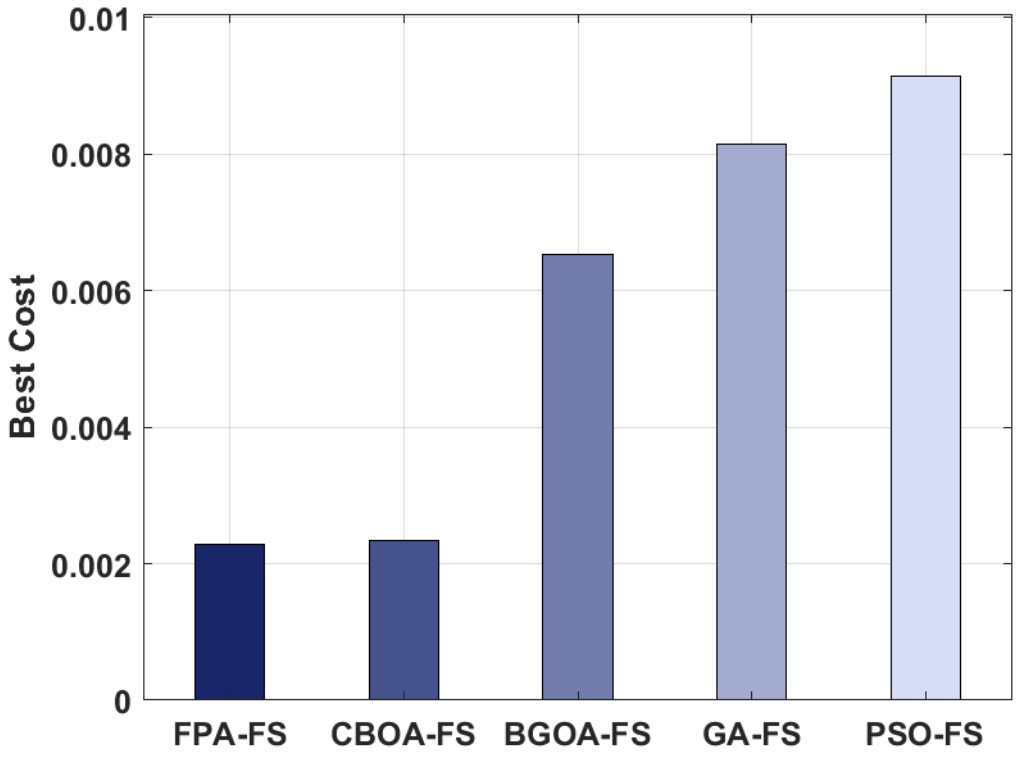

Fig. 2. BC analysis of FPA-FS technique

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021

An extensive comparative analysis of FPAFS-GRU technique with and without FS process is made in Table 1.

Fig. 3 investigates the results analysis of the FPAFS-GRU technique with and without FS interms of accuracy. The figure exhibited that the FPAFS-GRU technique has resulted in supreme performance with maximum accuracy. On examining the results without FS, it is evident that the DT and ensemble models have portrayed a lower accuracy value of 0.8868 and 0.8995. At the same time, the RF and LR techniques have attained a slightly improved accuracy of 0.9069 and 0.9037. However, the FPAFS-GRU technique has achieved a higher accuracy of 0.9645. On inspecting the results with FS, it is apparent that the RF and LR models have portrayed lower accuracy values of 0.8939 and 0.9013. Moreover, the DT and ensemble techniques have reached a certainly increased accuracy of 0.9104 and 0.9203. But the FPAFS-GRU technique has resulted in a maximum accuracy of 0.9959.

Fig. 4 examines the results analysis of the FPAFS-GRU technique with and without FS interms of precision. The figure showed that the FPAFS-GRU technique has resulted in utmost performance with maximum precision. On analyzing the performance of the results without FS, it is shown that the DT and RF models have depicted a lower precision value of 0.9000 and 0.9090. In line with, the LR and ensemble techniques have attained a slightly improved precision of 0.9150 and 0.9120. But the FPAFS-GRU technique has achieved a higher precision of 0.9690. On reviewing the results with FS, it is deceptive that the RF and LR models have portrayed lower precision values of 0.9070 and 0.9130. Furthermore, the DT and ensemble techniques have reached a certainly increased precision of 0.9190 and 0.9280. But the FPAFS-GRU technique has resulted in a maximum precision of 0.9898.

Fig. 5 investigates the results analysis of the FPAFS-GRU technique with and without FS interms of detection rate. The figure exhibited that the FPAFS-GRU technique has resulted in supreme performance with the maximum detection rate. On inspecting the results with FS, it is apparent that the DT and LR models have portrayed lower detection rate values of 0.9903 and 0.9929. Additionally, the RF and ensemble techniques have reached a certainly increased detection rate of 0.9942 and 0.9938. But the FPAFS-GRU technique has resulted in a maximum detection rate of 0.9976.

Table 1 Result Analysis of Proposed with Existing Methods in terms of Accuracy

Methods	Accuracy (%)		Precision (%)		Detection Rate (%)	
	None	With-FS	None	With-FS	None	With-FS
Proposed Model	0.9645	0.9959	0.9690	0.9898	0.9513	0.9976
Decision Tree	0.8868	0.9104	0.9000	0.9190	0.9885	0.9903
Random Forest	0.9069	0.8939	0.9090	0.9070	0.9933	0.9942
Logistic Regression	0.9037	0.9013	0.9150	0.9130	0.9578	0.9929
SVM+GBT+DT+RF+LR (ensemble)	0.8995	0.9203	0.9120	0.9280	0.9934	0.9938

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021

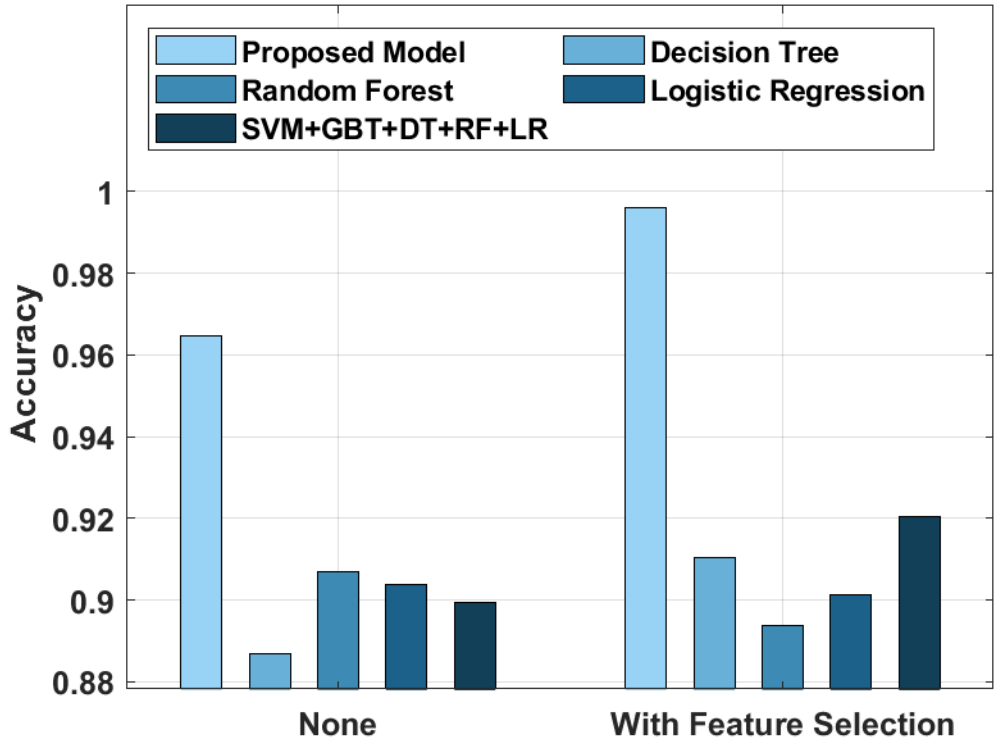

Fig. 3. Accuracy analysis of FPAFS-GRU technique with and without FS

Fig. 4. Precision analysis of FPAFS-GRU technique with and without FS

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021

Fig. 5. Detection rate analysis of FPAFS-GRU technique with and without FS

4. Conclusion

In this paper, a novel FPAFS-GRU technique is presented for intrusion detection in the IoT platform. The proposed FPAFS-GRU technique is mainly designed to determine the presence of intrusions in the network. The FPAFS-GRU technique involves the design of FPAFS technique to choose an optimal subset of features from the networking data. Besides, a deep learning based GRU model is applied as a classification tool to identify the network intrusions. An extensive experimental analysis takes place on KDDCup 1999 dataset and the results are investigated under different dimensions. The resultant simulation values demonstrated the betterment of the FPAFS-GRU technique with a higher detection rate of 0.9976. As a part of future extension, feature reduction approaches can be involved prior to the classification process.

References

[1] Almomani, A., Alauthman, M., Albalas, F., Dorgham, O. and Obeidat, A., 2020. An online intrusion detection system to cloud computing based on NeuCube algorithms. In Cognitive Analytics: Concepts, Methodologies, Tools, and Applications (pp. 1042-1059). IGI global.

[2] Besharati, E., Naderan, M. and Namjoo, E., 2019. LR-HIDS: logistic regression host-based intrusion detection system for cloud environments. Journal of Ambient Intelligence and Humanized Computing, 10(9), pp.3669-3692.

[3] Mehibs, S.M. and Hashim, S.H., 2018. Proposed network intrusion detection system in cloud environment based on back .propagation neural network Journal of University of Babylon for Pure and Applied Sciences, 26(1), pp.29-40.

[4] Singh, D.A.A.G., Priyadharshini, R. and Leavline, E.J., 2018. Cuckoo optimisation based intrusion detection system for cloud computing. International Journal of Computer Network and Information Security, 11(11), p.42.

[5] Ghosh, P., Karmakar, A., Sharma, J. and Phadikar, S., 2019. CS-PSO based intrusion detection system in cloud environment. In Emerging Technologies in Data Mining and Information Security (pp. 261-269). Springer, Singapore.

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021

[6] Deshpande, P., Sharma, S.C., Peddoju, S.K. and Junaid, S., 2018. HIDS: A host based intrusion detection system for cloud computing environment. International Journal of System Assurance Engineering and Management, 9(3), pp.567-576.

[7] Chiba, Z., Abghour, N., Moussaid, K., El Omri, A. and Rida, M., 2019. New anomaly network intrusion detection system in cloud environment based on optimized back propagation neural network using improved genetic algorithm. International Journal of Communication Networks and Information Security, 11(1), pp.61-84.

[8] Manickam, M. and Rajagopalan, S.P., 2019. A hybrid multi-layer intrusion detection system in cloud. Cluster Computing, 22(2), pp.3961-3969.

[9] Umamaheswari, K. and Sujatha, S., 2017. Impregnable Defence Architecture using Dynamic Correlation-based Graded Intrusion Detection System for Cloud. Defence Science Journal, 67(6).

[10] Mahajan, V. and Peddoju, S.K., 2017, August. Deployment of intrusion detection system in cloud: a performance-based study. In 2017 IEEE Trustcom/BigDataSE/ICESS (pp. 1103-1108). IEEE.

[11] Jelidi, M., Ghourabi, A. and Gasmi, K., 2019, April. A hybrid intrusion detection system for cloud computing environments. In 2019 International Conference on Computer and Information Sciences (ICCIS) (pp. 1-6). IEEE.

[12] Balamurugan, V. and Saravanan, R., 2019. Enhanced intrusion detection and prevention system on cloud environment using hybrid classification and OTS generation. Cluster Computing, 22(6), pp.13027-13039.

[13] Li, Z., Xu, H. and Liu, Y., 2017. A differential game model of intrusion detection system in cloud computing. International Journal of Distributed Sensor Networks, 13(1), p.1550147716687995.

[14] Tummalapalli, S.R.K. and Chakravarthy, A.S.N., 2021. Intrusion detection system for cloud forensics using bayesian fuzzy clustering and optimization based SVNN. Evolutionary Intelligence, 14(2), pp.699-709.

[15] Wen, L., 2021. Cloud Computing Intrusion Detection Technology Based on BP-NN. Wireless Personal Communications, pp.1-18.

[16] Murugan, I., 2021. Supervised classifier approach for intrusion detection on KDD with optimal mapreduce framework model in cloud computing. Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 14(4), pp.1115-1125.

[17] Singh, P. and Ranga, V., 2021. Attack and intrusion detection in cloud computing using an ensemble learning approach. International Journal of Information Technology, 13(2), pp.565-571.

[18] Krishnaveni, S., Sivamohan, S., Sridhar, S.S. and Prabakaran, S., 2021. Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing. Cluster Computing, pp.1-19.

[19] Sobin Soniya, S. and Maria Celestin Vigila, S., 2021. Feedback deer hunting optimization algorithm for intrusion detection in cloud based deep residual network. International Journal of Modeling, Simulation, and Scientific Computing, p.2150047.

[20] Porkodi, V., Singh, A.R., Sait, A.R.W., Shankar, K., Yang, E., Seo, C. and Joshi, G.P., 2020. Resource provisioning for cyber–physical–social system in cloud-fog-edge computing using optimal flower pollination algorithm. IEEE Access, 8, pp.105311-105319.

[21] Yan, B. and Han, G., 2018. LA-GRU: Building combined intrusion detection model based on imbalanced learning and gated recurrent unit neural network. security and communication networks, 2018.

[22] https://www.unb.ca/cic/datasets/nsl.html

DOI: 10.5281/zenodo.5501286
Received: February 19, 2021 Accepted: July 29, 2021