INCREMENTAL DEEP NEURAL NETWORK INTRUSION DETECTION IN FOG BASED IOT ENVIRONMENT: AN OPTIMIZATION ASSISTED FRAMEWORK

Aftab Alam Abdussami
Research Scholar, Dept of Computer Application, Integral University, Dasouli,
Lucknow, Uttar Pradesh 226026, India
abdussami@iul.ac.in

Dr. Mohammed Faizan Farooqui
Asst. Professor, Dept of Computer Application, Integral University, Dasouli,
Lucknow, Uttar Pradesh 226026, India
ffarooqui@iul.ac.in

Abstract-

IoT has gained more attention on sharing data and directing respective tasks. IoT also has the potential to enhance the lifestyle of people. Fog computing has been evolved to process information to make its service more efficient. However, the evolution of various technologies also raises difficulties in network security. This paper intends to introduce a new intrusion detection model in fog-based IoT considering the phases like Preprocessing, Feature Extraction, and Intrusion Detection. Initially, data normalization is carried out in the raw data (considered as pre-processing). Then, the pre-processed data is subjected to the feature extraction phases, where the proposed entropy-based features, gain information, and gain ratio features are extracted. Subsequently, the extracted features are subjected to the attack detection phase, where the Incremental Deep Neural Network (DNN) will detect the presence of an intruder in the network based on the data attributed. The training of DNN is made optimal via tuning the weights. This optimization is carried out by a Modified Electric Fish Optimization Algorithm (M-EFO). At last, the supremacy of the proposed developed model is examined via evaluation over extant techniques.

Keywords: Attack detection; Fog; Optimized DNN; Incremental learning; M-EFO Optimization; Fog Computing

Nomenclature

Abbreviation	Description
CEP	Complex event processing
DDoS	Distributed Denial-Of-Service
DL	Deep Learning
DNN	Deep Neural Network
EFO	Electric Fish Optimization Algorithm
FC	Fog Computing
EHO	Elephant Herding Optimization
GI	Gain information
IL	incremental learning
IMT	Internet of Medical things
IDS	Intrusion Detection System
KNN	K-Nearest Neighbour
FNR	False Negative Rate
FDR	False Discovery rate
FPR	False-positive rate
GWO	Grey Wolf Optimization

IoT	Internet of Things
IMT	Internet of Medical things
M-EFO	Modified EFO
MCC	Matthews Correlation Coefficient
MFO	Moth Flame Optimization
NPV	Negative Predictive Value
OLDA	Online Learning Detection Algorithm
RF	Random Forest
SVM	Support Vector Machine
WOA	Whale Optimization Algorithm

1. INTRODUCTION

Fog Computing moves the storage and computing resources closer to the Internet of Things devices. FC arises when the need for immediate responsive tasks emerges and increases in IoT applications. Fog Computing is a decentralized system that is subjective to the context-aware of the information related to data sources such as response time, location, resources disbursed by the service, Etc. It is provided at the edge node in the network, either in physical or virtual form. [Manimurugan (2021)]

The propagation of IoT devices and their appliances in various fields, namely, smart homes, smart cities, smart health Etc., has offered several advantages in the recent progressions. IoT networks are practicing remarkable development, and these devices are bound to reach about 50 billion by the end of 2021. However, this development comes with various challenges [Cristiano *et al.* (2020)] [Rabia Latif *et al.* (2019)]. Accordingly, the foremost confront is the security of such linked devices that are increasingly under attack. Conversely, there is a shortage of sufficient resources (computational and storage), which characterize IoT devices and are necessary for employing security solutions like network abnormality alleviation, which is generally done by IDS on IoT networks [Marcos V.O *et al.* (2020)a] [Quoc-Dung Ngo *et al.* (2020)b] [K.Mandal *et al.* (2020)c].

The security demands in IoT network generally comes as network abnormalities, primarily when deviation occurs from the normal network traffic flow. Such anomalous network flows include Probing attacks and DDoS attacks. A botnet consists of a larger count of hijack nodes or network systems regulated by malevolent users. These systems or nodes were deployed for executing numerous kinds of attacks [Bohan Li *et al.* (2020)]. Three features generally categorize a botnet attack: "similarity of attack sources, the divergence between normal and attack network traffic flow, and automation of attack execution."

The fog computing concept is deployed to improve the shortage of needed resources for functioning abnormality mitigation models in IoT networks and protect and ensure the proficient function of IoT devices in-network. Fog computing is envisaged to alleviate storage, computational, energy utilization, and latency requirements by bringing these sources to the network edge [8]. In this manner, IoT devices and appliances could obtain a faster and better response and relief from doing operations that stretch their resources and decrease their efficacy.

The contribution is described here:

- Proposes an intrusion detection model, where the proposed entropy-based features, gain information, and gain ratio features are considered to relate to the behaviour of attackers or intruders.

- Deploys Optimized DNN with incremental learning concept for detecting the presence of intruders, where a new M-EFO algorithm fine-tunes its weights.

In this work, section II briefs the reviews on IDS models. The architecture of fog-based IoT is given in Section III. Proposed IDS system: pre-processing and feature extraction are elaborated in section IV. Section V briefs optimized DNN with incremental learning via M-EFO algorithm. Finally, Section VI specifies the result and the work is concluded by section VII.

2. LITERATURE REVIEW

2.1 *Related works*

In 2021, [Manimurugan *et al.* (2021)] implemented an IDS model with the combination of Fog with cloud in IoT to support IoT-based smart city applications. Fog computing has offered different services, and cloud computing has given data storage at top-level management to support various IoT-based applications. This model

has provided Network-based IDS (NIDS) to detect the anomalies in the network by applying Principle Component Analysis (PCA) technique with the improved Naïve Bayes classifier (PCA) technique. The experiment was conducted on the "UNSW-NB15 dataset" to demonstrate the efficiency of the suggested IDS model concerning the detection rate, recall, precision, and accuracy.

In 2021, [Kumar *et al.* (2019)] have suggested IDS on the Internet of Medical Things (IoMT) based on fog-cloud computing with an ensemble learning framework. First, it was proposed by considering the different classifiers in the first and second levels, random forest, naive Bayes, decision tree in level 1, and the XGBoost was used to get the classification results with identification of attack or standard instances. Then, deployment architecture was presented by considering both heterogeneous and dynamic networks, where Infrastructure as a Service (IaaS) was assessed on the cloud side and the Software as a Service (SaaS) on the fog side. Finally, this model has been experimented on the "ToN-IoT" dataset. The experimental analysis showed that the proposed IDS attained superior performance in a false alarm, detection, and accuracy.

In 2020, [Ahmed *et al.* (2020)] analyzed a comprehensive attack detection model for cyber-attacks in IoT utilizing DL. In addition, the attack detector was implemented on fog nodes. Six DL models were evaluated over other DL models to discover the most excellent one to perform superior. DL techniques require more extensive storage, widespread computation, and higher power, owing to their higher ability to extract unknown attacks and known attacks. However, every DL technique was computed for detecting different attacks, and they offered superior detection accuracy than existing models.

In 2020, [Sibi *et al.* (2020)] had developed the solution for ransomware. A most special attack besieged the individuals and industries and IoT. The majority of the systems do not succeed in ransomware recognition, such as Anti-Virus and IDPS. IDH was developed to detect ransomware attacks with three major parts like "audit watch, Honey folder, Complex event processing." Honey folder acts as an earlier caution system when unlawful activities occur modeled by deploying the SLA model. CEP made data aggregation. Further, this technique has detected diverse ransomware attacks.

In 2020, [Geethapriya *et al.* (2020)] proposed that in IMT, every medicinal device gets embedded with sensors and transmits data over wireless communication. It was very supportive in health care appliances for making everything as smart. However, illegal activities could take place. Therefore, novel mobile agent-oriented IDS were constructed for preventing attacks in linked medical devices. The adopted work has exploited "machine learning and regression algorithms" for detecting attacks.

In 2020, [Bohan *et al.* (2020)] had introduced the online learning recognition model for identifying malicious nodes in the network. The malicious nodes in IoT networks launch a variety of attacks. However, they could be identified. These make the nodes achievable to receive and securely send the messages. This procedure consumed more time for collecting every message from each node. The OLDA computed the reliability of paths in-network, and the clustering model noticed every node trust and malevolent nodes. The improved online learning detection scheme has proven that high stability and accuracy were good.

In 2020, [Hamid *et al.* (2020)] have offered the attack defense schemes, where intrusions were detected and analyzed in distributed IoT systems. Every node has to be concerned about detecting intrusions in a neighbor node. Therefore, every suitable node has protection mechanisms for protecting the systems when malevolent nodes subsist in the network. Further, a systematic model was introduced depending upon the "Stochastic Petri Net technique." The defense mechanisms improved the life span of the system in a functioning locality.

In 2019, [Sudqi *et al.* (2019)] had offered a lightweight IDS using Multilayer Perceptron (MLP) on a vector space representation of fog computing. The suggested model was executed on "Australian Defense Force Academy Linux Dataset (ADFA-LD) and Australian Defense Force Academy Windows Dataset (ADFA-WD)," which has exploited different attacks on diverse applications. The proposed model has employed the feature extraction method based on n-gram transformation using modified vector space representation. In addition, the matrix formatting was compressed using the sparse matrix, where the zero values were compensated through the linear correlation coefficient. Furthermore, the number of features was reduced using mutual information for feature selection. The experimental results have revealed that the suggested model has improved the accuracy of detection against attacks.

3. THE ARCHITECTURE OF FOG BASED IOT ENVIRONMENT

The present architecture is designed for operating at a foggy computing layer. Based on the appliance, it is essential to include numerous fog devices with implanted IDS. Every fog device encompasses detection modules, which carry out the classification without communication with the cloud, therefore eliminating latency. Every fog device is accountable for observing its connected IoT networks. The traffic on every IoT network is monitored by its relevant fog node that functions in immoral mode. Substantial alterations in traffic on an IoT network may take place. If these alterations are malevolent to DoS attacks, the recognition technique will recognize and instigate countermeasure activities [6].

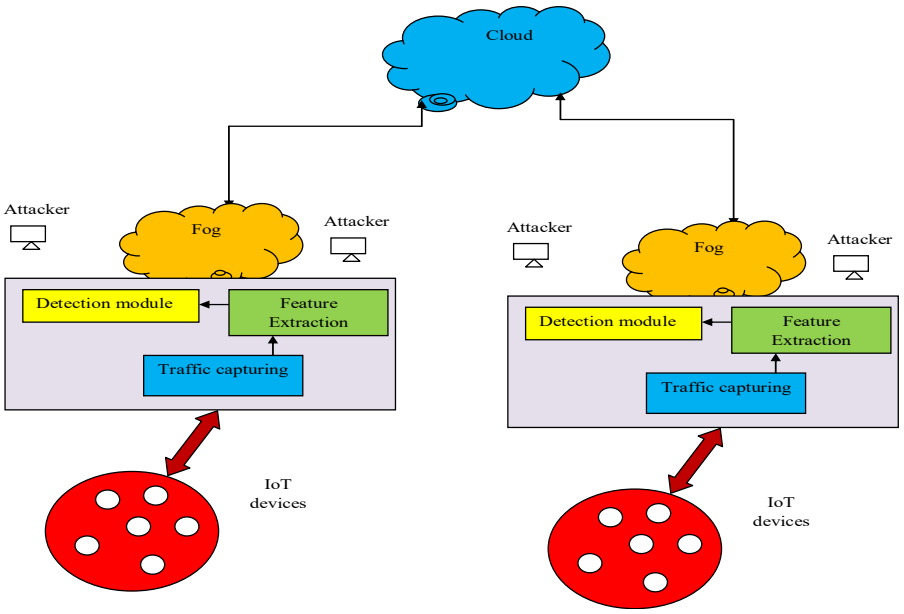


Fig. 1. Architecture of Fog based IoT environment

4. PROPOSED IDS SYSTEM USING INCREMENTAL DEEP NEURAL NETWORK MODEL

As stated, the proposed intrusion detection module involves certain phases to process the data traffic. The corresponding data is pre-processed first, then extracting the relevant information or features trained to the deep model. Moreover, the incremental learning concept is also involved in the proposed work. Also, the optimization logic is assisted with the model to make the precise detection results about the presence of intruders.

The process flow of the detection system is given in Figure 2.

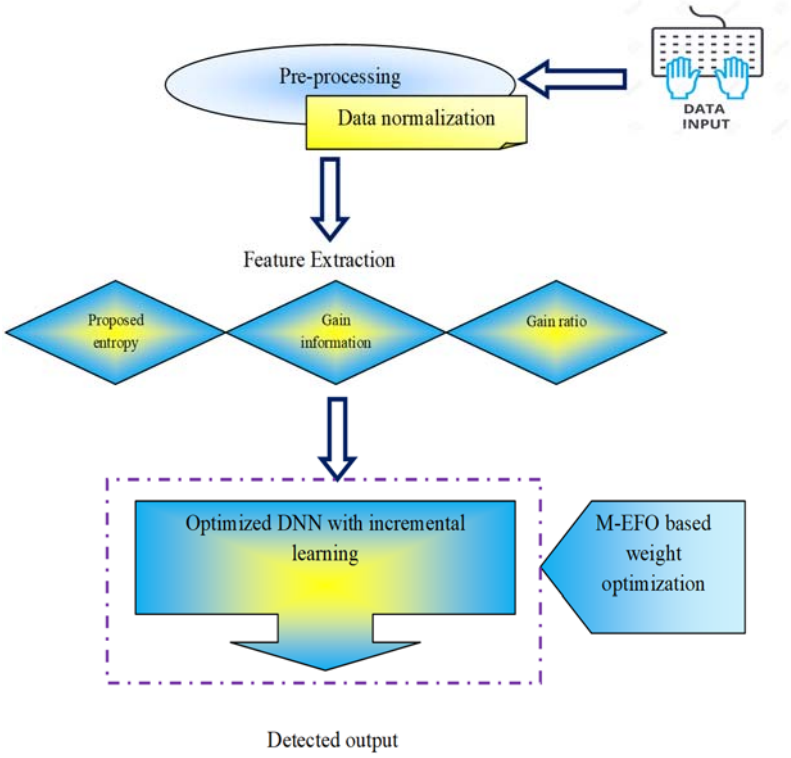


Fig. 2. Intrusion detection framework

4.1 *Pre-processing*

Initially, pre-processing of data is carried out by the data normalization process [Stephen Watts *et al.* (2020)]. Data normalization is defined as the organization of data to appear similar across all records and fields, and it increases the cohesion of entry types leading to cleansing and higher quality data. By carrying out data normalization, here, the entire data are normalized between 0 and 1. Thus, the attained pre-processed data is denoted by (Da_{pre}).

4.2 *Extracting proposed entropy, gain information, and gain ratio features*

From Da_{pre} the proposed entropy, gain knowledge and gain ratio features are extracted.

Proposed entropy: The entropy describes the impurities of data, and it computes the heterogeneity of input data regarding its classification [6]. Conventionally, entropy is evaluated based on the proportion of records; however, as per the proposed entropy, the evaluation takes place based on a weight factor denoted by w. Eq. (1) shows the entropy of the dataset d with C classes, where, $\rho\left(\frac{i}{d}\right)$ refer to the proportion of records that belongs to rank i in d. Here, it w is computed as shown in Eq. (2), $(m - bh_i) \geq 0$ $m = b = 1$ h referring to the entropy value.

$$En(d) = -\sum_{i=1}^{C} \rho\left(\frac{i}{d}\right) * \log_2\left[\rho\left(\frac{i}{d}\right)\right] * w \tag{1}$$

$$w = \frac{m - bh_j}{\sum_{j=1}^{n}(m - bh_i)} \tag{2}$$

GI: "It is attained by subtracting the entropy of the set d by the entropy that is obtained when partitioning d using the attribute a" [6]. It is computed as in Eq. (3), wherein, $p(a)$ refers to the set of values that a could presume, x refers to d its element, and dx refers to the subset d created by the data $a = x$.

$$GI(d,a) = En(d) - \sum_{x \in p(a)}\left[\frac{dx}{d} * En|dx|\right] \tag{3}$$

Gain Ratio (GR): It is deployed to discover the most delicate features of traffic packets [6]. It is computed as in Eq. (4).

$$GR(d,a) = \frac{GI(d,a)}{En(d)} \tag{4}$$

These features $En(d)$ $GI(d,a)$ $GR(d,a)$ are considered the final feature set Fe, which is subjected to train the DNN with incremental learning.

5. OPTIMIZED DNN WITH INCREMENTAL LEARNING WITH M-EFO ALGORITHM

5.1 *Optimized DNN model with incremental learning*

Generally, DNN [Ghadaa.A *et al.* (2021)] is regarded as the most noteworthy computational network comprising ways TO interconnect nodes and numerous hidden layers with nodes. DNN algorithm constructs the model via the three most important steps. At first, the model topology determines the neurons for every layer and counts layers with their relationships. Subsequently, the forward propagation with its activation function and perceptron classifier is used employing the artificial neurons. At last, the backpropagation is deployed with optimizer and loss function.

(i) The model topology:

- **Input layer:** It initializes the data for NN purposes.
- **Hidden layers:** It acts as an intermediary layer amongst output and input layers in which all the computation is carried out.
- **Output layer:** It generates the outcomes (i.e.), attack, or normal with attack types.

(ii) Forwarding propagation:

It tends to predict the outcomes (regular or attack) utilizing perceptron classifier. The perceptron is indicated as in Eq. (5).

$$Y = \sum_{j=1}^{p} J_j \; W_j + B \tag{5}$$

In Eq. (5) p refers to the numbers of nodes in a layer, J refers to the values of these nodes, W refers to the weights, which are optimally tuned by the M-EFO algorithm, and B refers to the bias of these nodes.

(iii) Backpropagation:

It is a well-known method for training a DNN model via weight and bias modifications. It consists of optimizers and loss functions. In addition, the cost function, also known as the loss function, diminishes the value for reaching the optimal values for modelling parameters.

*(iv)Incremental Training***:**

Here, a total training method is proposed. Initially, the set of classes is separated into two sets. First, the bigger or "core" set is deployed for training a primary network. Subsequently, the small set or "demo" set is deployed for training cloned branch networks with diverse sharing configurations. From training resultants, sharing curve vs. accuracy is produced, from which the optimal sharing structures and architecture are elected.

Objective:

The objective Obj of the developed scheme is to minimize the error as given in Eq. (6), wherein Er signifies the error between actual and predicted value.

$$Obj = Min(Er) \tag{6}$$

5.2 *Proposed M-EFO algorithm*

Though the conventional EFO [Yilmaz *et al.* (2020)] model encompasses a variety of enhancements, it suffers from specific limitations in maintaining convergence rate and speed. Hence, certain modifications are needed in the conventional process. Generally, self-improvement is considered to be capable in conventional optimization schemes [B.R. Rajkumar (2013)] [Geeta S. Navale *et al.* (2020)] [Ali Mortazavi (2019)] [Yong Zhu (2017)] [Venkata Rao *et al.* (2020)] [Mukund Wagh *et al.* (2019)] [Kodad S.R *et al.* (2019)] [Amolkumar *et al.* (2019)] for better convergence. The steps of the proposed M-EFO are as follows.

Initially, the population of electric fish are spread randomly via the search space by considering the space bounds, wherein, X_{uv} refers to u^{th} an individual position in N a D -dimensional space, $X_{\max v}$ and $X_{\min v}$ refers to upper and lower bounds for dimension v, correspondingly, $\phi \in (0,1)$ refers to arbitrary values drawn from a uniform distribution [Yilmaz *et al.* (2020)].

$$X_{uv} = X_{\min v} + \phi\left(X_{\max v} - X_{\min v}\right) \tag{7}$$

The frequency holds a significant role in the M-EFO model for balancing exploitation and exploration and is deployed for determining if an individual performs passive or active electrolocation.

5.2.1 *Active electrolocation***:**

The dynamic ranges of u^{th} an individual (r_i) are portrayed by its value of amplitude (B_i). The functional range computation in EFO is specified in Eq. (8). Conventionally, cartesian distance is computed based on X_{uv} and X_{kv} , however, in the developed M-EFO model, the Minkowski-based distance evaluation is done as shown in Eq. (9).

$$r_i = \left(X_{\max v} - X_{\min v}\right)B_i \tag{8}$$

$$D_{uk} = \left(\sum_{i=1}^{n}\left(X_u - X_k\right)^P\right)^{\frac{1}{P}} \tag{9}$$

If at least a single neighbor is there in the active sensing region, EFO deploys Eq. (10); else, Eq. (11) is deployed, wherein, k refers to randomly selected individual from neighbor set of u^{th} the individual.

$$X_{uv}^{can} = X_{uv} + \phi\left(X_{kv} - X_{uv}\right) \tag{10}$$

$$X_{uv}^{can} = X_{uv} + \phi r_i \tag{11}$$

Here, $\phi \in (-1,1)$ Eq. (10) refers to arbitrary count produced from an even distribution and X_{uv}^{can} refers to the candidate location of the u^{th} individual.

5.2.2 Passive electrolocation:

Conventionally, the new location is updated based on X_{rv} and X_{uv}. However, as per the proposed M-EFO, the location is updated based on the best position (X_{best}) and levy $(Levy(\beta))$, as shown in Eq. (12).

$$X_{uv}^{new} = X_{uv} + \phi(X_{rv} - X_{best}) + Levy(\beta) \tag{12}$$

The last phase of passive location is to adjust a constraint of u^{th} the individual as per Eq. (13) to increase the probability of a trait to be varied, wherein $ran(0,1)$ refers to the arbitrary values drawn from the uniform distribution.

$$\begin{aligned} X_{uv}^{can} &= X_{\min v} + \phi(X_{\max v} - X_{\min v}) \\ ran(0,1) &\le ran(0,1) \end{aligned} \tag{13}$$

If v^{th} the constraint value of an u^{th} individual goes beyond the bounds of search space, it is reallocated to bounds of space, which it goes beyond.

$$X_{uv}^{can} = \begin{cases} X_{\min v} & X_{uv}^{can} < X_{\min v} \\ X_{uv}^{can} & X_{\max v} > X_{uv}^{can} > X_{\min v} \\ X_{\max v} & X_{uv}^{can} > X_{\max v} \end{cases} \tag{14}$$

6. RESULTS AND DISCUSSIONS

6.1. *Simulation Set up*

The developed IDS scheme employing DNN with IL + M-EFO model was executed in MATLAB**.** The dataset for analysis was downloaded from https://research.unsw.edu.au/projects/toniot-datasets. Accordingly, the performance of adopted approach was measured over extant models such as DNN-kNN [Cristiano *et al.* (2020)], DNN + M-EFO, DNN [Ghadaa *et al.* (2021)], RF [Rezai *et al.* (2009)], SVM [Rezai *et al.* (2009)], DNN with IL + WOA [Andrew Lewis *et al.* (2016)], DNN with IL + GWO [Andrew Lewis *et al.* (2014)], DNN with IL + MFO [Syedali *et al.* (2015)], DNN with IL + EHO [Wang *et al.* (2015)], and DNN + EFO [Yilmaz *et al.* (2020)] regarding certain positive, negative and neutral measures. The performance analysis was performed for varied learning rates: 60, 65, 70, 75, 80, 85, and 90.

6.2. *Evaluation metrics*

The performance of the developed model is illustrated via the metrics such as accuracy, sensitivity, precision, specificity, FPR, FNR, FDR, F1-score, MCC, and NPV.

Accuracy: To predict the accurate prediction of the method.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

Here, TP the true positive, TN the actual negative, FP the false positive, and FN the false negative.

Sensitivity: Ability of the measure for classifying the data to detect the actual positive value and is expressed as,

$$Sen = \frac{TP}{TP + FN} \tag{16}$$

Specificity: Ability of the measure for classifying the data to detect the actual negative value and is expressed as,

$$Spe = \frac{TN}{TN + FP} \tag{17}$$

Precision: Quantifies the number of positive predicted result that belongs to the positive class.

$$\Pr e = \frac{TP}{TP + FP} \tag{18}$$

FPR: Probability of rejecting the null values of data and is represented as,

$$FPR = \frac{FP}{FP + TN} \tag{19}$$

FNR: Probability of false alarm and it is also known as miss rate, and it is expressed as,

$$FNR = \frac{FN}{FN + TP} \tag{20}$$

F1-score: Offers a single score that balances the concern of recall and precision, and it is expressed as,

$$F1 - score = \frac{2TP}{2TP + FP + FN} \tag{21}$$

NPV: Ratio of true negative to the sum of true negative and false negative and it is expressed as,

$$NPV = \frac{TN}{TN + FN} \tag{22}$$

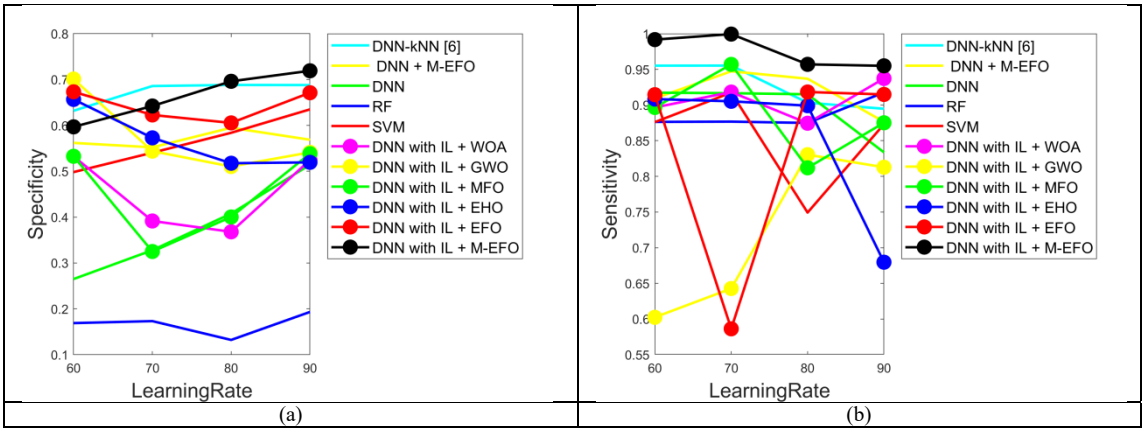
MCC: It is used in machine learning as a measure of the quality of binary classification, and it is expressed as,

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{23}$$

6.3. *Performance Analysis*

The proposed intrusion detection model (DNN with IL + M-EFO model) is computed over extant models concerning "positive measures like accuracy, sensitivity, specificity, precision and negative measures like FDR, FNR, FPR and neutral measures like F1-score, MCC and NPV". On examining the whole graphs, the presented DNN with IL + M-EFO model has obtained better outcomes than compared schemes. In particular, nominal negative values assure a more good detection rate of the model. For example, from Fig. 3(c), the suggested DNN with IL + M-EFO model has accomplished improved accuracy value than DNN-kNN, DNN + M-EFO, DNN, RF, SVM, DNN with IL + WOA, DNN with IL + GWO, DNN with IL + MFO, DNN with IL + EHO, DNN with IL + M-EFO schemes when the learning rate is 60. Primarily, it can be noticed that specific measures such as accuracy, precision, NPV, and sensitivity decrease with an increase in learning rates. However, the adopted DNN with IL + M-EFO model has exposed superior values that ensure the better performance of the adopted optimization concept.

On the other hand, specific measures such as F1-score, MCC, and specificity increase with learning rates. In addition, negative metrics such as FDR and FPR increase with an increase in learning rates, while FPR decreases with an increase in learning rates. Hence, the superior speeds of positive measures and lower rates of aggressive actions have assured the enhancement of the adopted scheme in detecting the presence of intruders. This analysis has proven the impact of proposed optimization logic in the profound learning concept with optimal system training. Specifically, the convergence of error minimization ensures accurate detection of an attack.



(a) (b)

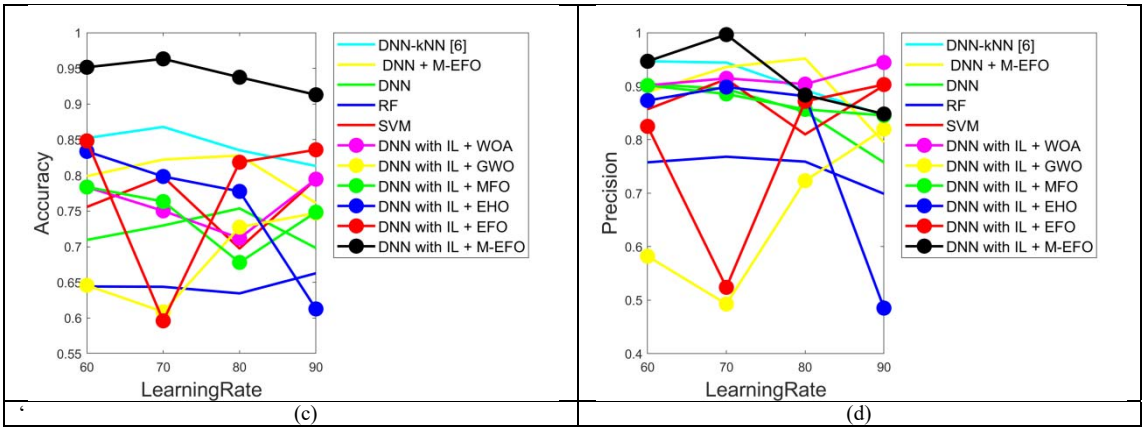


Fig. 3. Performances of developed method over extant models for (a) specificity (b) sensitivity (c) accuracy (d) precision

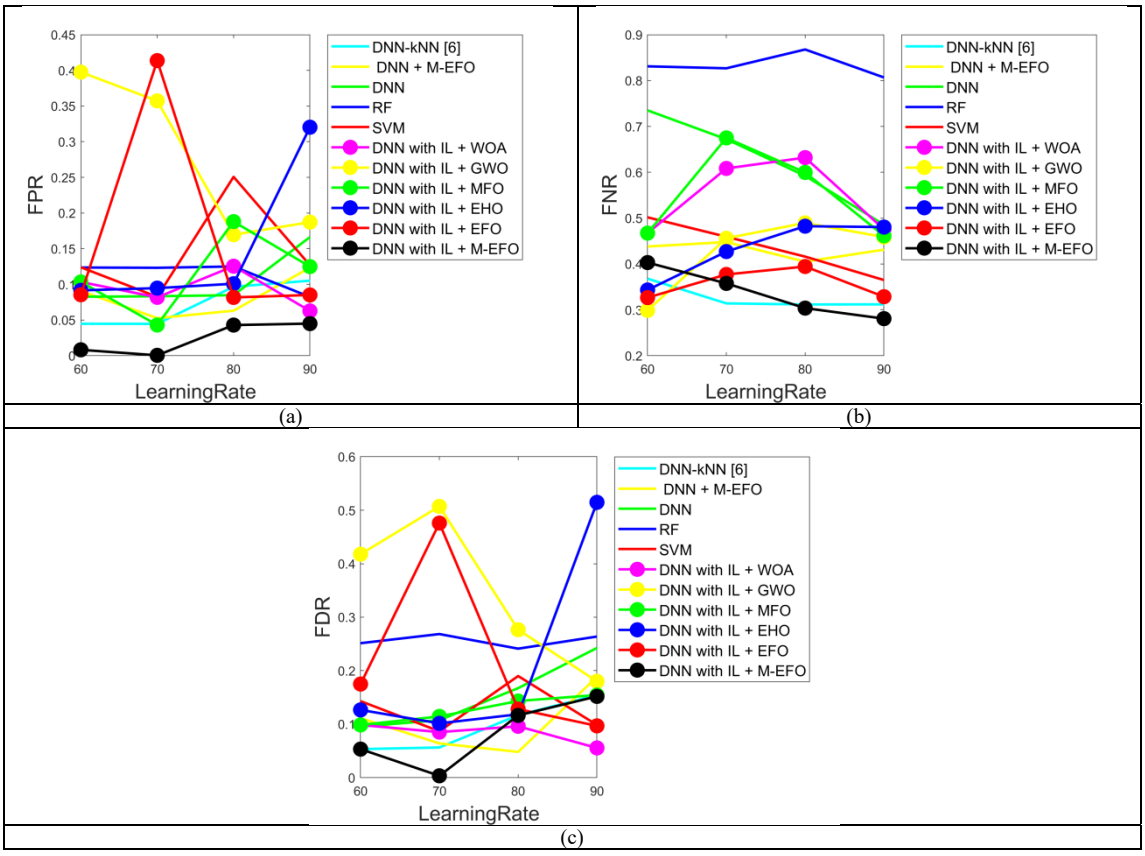


Fig. 4. Performances of developed method over extant models for (a) FPR (b) FNR (c) FDR

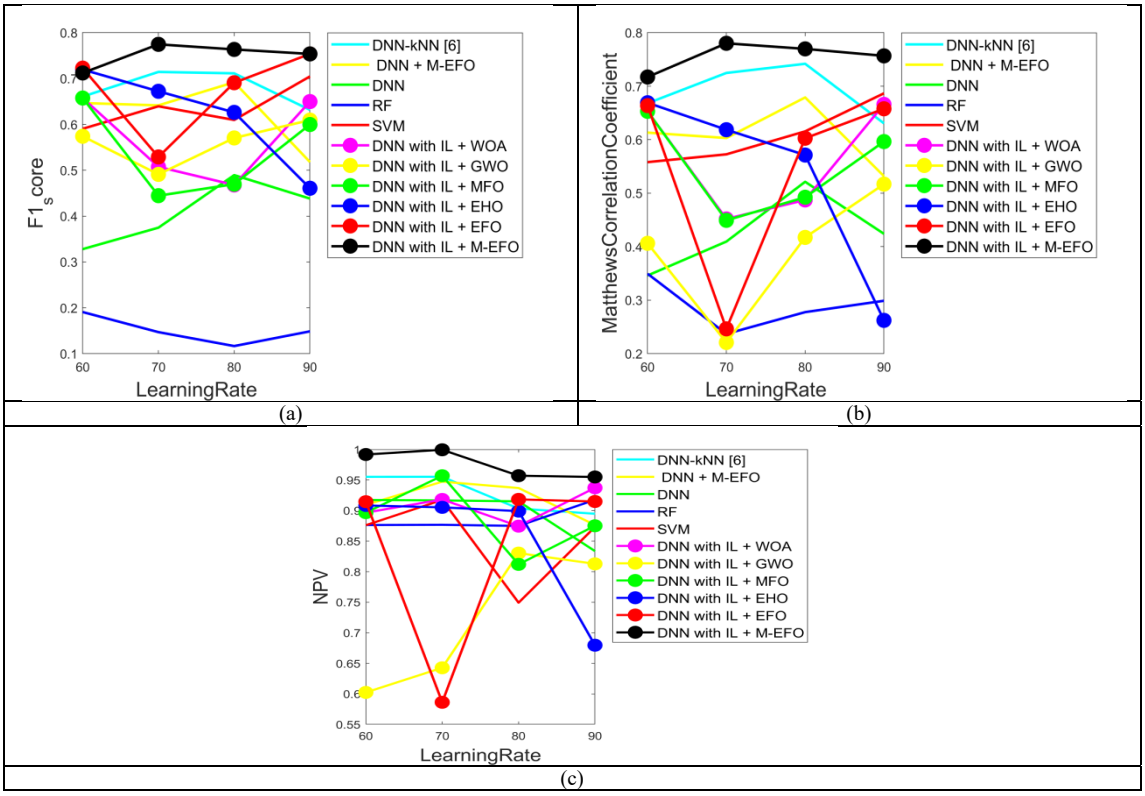


Fig. 5. Performances of developed method over extant models for (a) F1-score(b) MCC (c) NPV

6.4. *An overall analysis of proposed and conventional models*

Table I describes the overall analysis of the adopted model over conventional approaches. Here, research is performed for varied metrics such as "accuracy, sensitivity, specificity, precision and negative measures like FDR, FNR, FPR and neutral measures like F1-score, MCC and NPV". On noticing the analysis resultants, the developed DNN with IL + M-EFO model has obtained optimal values when evaluated over existing schemes. Predominantly, on detecting accuracy values from Table I, the adopted DNN with IL + M-EFO scheme has attained better value, while the compared DNN with IL + EFO model has accomplished minimal accuracy value 0.59609. Thus, the accuracy of the adopted model is 61.6% superior to traditional DNN with IL + EFO scheme. In addition, on examining the precision, the developed DNN with IL + M-EFO model has attained higher values than the extant schemes. Thus, the overall assessment shows the impact of the developed DNN with IL + M-EFO model on better overall results.

TABLE I. AN OVERALL ANALYSIS OF DEVELOPED SCHEME OVER TRADITIONAL MODELS REGARDING VARIED METRICS

Metrics	**Developed DNN with IL+ M-EFO**	Traditional EFO	MFO	EHO	GWO	WOA	SVM	RF	DNN	DNN + M-EFO	DNN-kNN
Accuracy	0.96332	0.59609	0.76337	0.79843	0.60868	0.75021	0.79793	0.64379	0.72974	0.82217	0.86792
Specificity	0.64232	0.62282	0.32506	0.57289	0.54389	0.39162	0.54057	0.17326	0.32788	0.5522	0.68589
Sensitivity	0.99947	0.58633	0.95698	0.90527	0.64278	0.91818	0.91743	0.87675	0.91664	0.94732	0.95532
Precision	0.99643	0.52408	0.88544	0.89844	0.49281	0.91473	0.91317	0.76822	0.89475	0.93585	0.94415
FPR	0.000534	0.41367	0.04302	0.094732	0.35722	0.081821	0.082568	0.12325	0.083365	0.052677	0.044682

F1-score	0.77463	0.52906	0.44462	0.67223	0.49106	0.50748	0.63957	0.14707	0.37485	0.64155	0.71449
MCC	0.78018	0.24644	0.44927	0.61867	0.22098	0.45197	0.57268	0.23706	0.40962	0.60313	0.72478
FNR	0.35768	0.37718	0.67494	0.42711	0.45611	0.60838	0.45943	0.82674	0.67212	0.4478	0.31411
NPV	0.99947	0.58633	0.95698	0.90527	0.64278	0.91818	0.91743	0.87675	0.91664	0.94732	0.95532
FDR	0.003572	0.47592	0.11456	0.10156	0.50719	0.085274	0.086831	0.26851	0.10763	0.064147	0.056502

6.5. *Convergence Analysis*

Fig. 6 describes the convergence (cost) analysis of adopted DNN with IL + M-EFO scheme over traditional schemes such as GWO, EHO, MFO, WOA, and EFO for varied iterations. Here, analysis is done by varying the iterations from 0, 10, 20, 30, 40, and 50. On observing the analysis outcomes, the proposed DNN with IL + M-EFO model has attained minimal values for all iterations compared to the existing schemes. Initially, from iteration 0 to iteration 20, the cost values are higher for proposed and evaluated models. However, as the iteration count increases, better (minimal cost) outputs are attained. Finally, from iteration 20 to 50, the cost values reduce for both the proposed and compared models. However, the adopted DNN with IL + M-EFO scheme exhibits the least values compared to the existing ones. Predominantly, the presented approach has accomplished the least-cost deal (almost 0.0055) with its successive enhancement in the process, particularly the levy flight evaluation. Thus, the overall assessment shows the enhancement of the presented model.

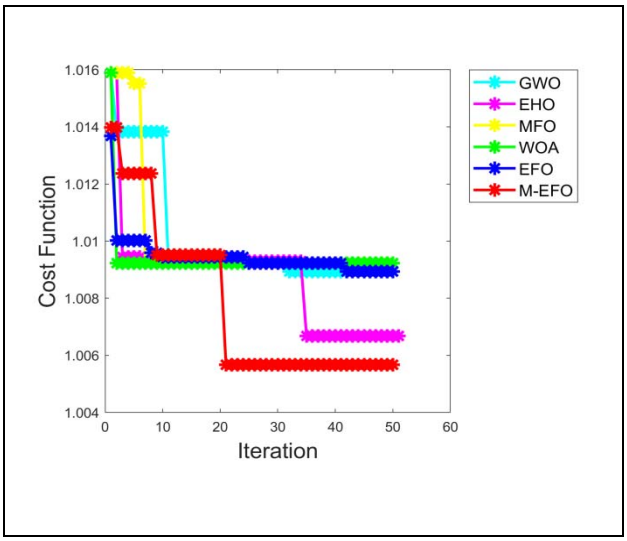


Fig. 6. Convergence analysis of developed approach over compared approaches

7. Conclusion

This paper introduced an intrusion detection model that involved three stages, namely "(i) Pre-processing, (ii) Feature Extraction, and (iii) Intrusion Detection." Primarily, the input data was pre-processed via data normalization. From the normalized data, proposed entropy, gain information, and gain ratio features were extracted, which were then subjected to optimized DNN with incremental learning, where the weights of DNN were optimally tuned using M-EFO. From the resultants, the adopted DNN with IL + M-EFO scheme has attained a better value (0.96332), while the compared DNN with IL + EFO model has accomplished a minimal accuracy value 0.59609. Thus, the accuracy of the adopted model was 61.6% superior to traditional DNN with IL + EFO scheme. Thereby, the development of the suggested DNN with IL + M-EFO scheme was validated effectively.

8. ACKNOWLEDGEMENT

This paper is acknowledged under Manuscript Number IU/R&D/2021-MCN0001276

9. REFRENCES

[1] Abdussami,A.A.; Farooqui,F.M. (2020): A Systematic Literature Review on Fog Computing. International Journal of Advanced Science and Technology, 29 (7), 12755-12769.

[2] Almiani,M.; AbuGhazleh, A.; Al-Rahayfeh,A.; Atiewi,S.; Abdul Razaque.(2019): Deep Recurrent Neural Network For IoT Intrusion Detection System. JournalPreProof.

[3] Al-Hamadi, H.; Chen, I.; Wang, D.; Almashan, M. (2020): Attack and Defense Strategies for Intrusion Detection in Autonomous Distributed IoT Systems. IEEE ACCESS.

[4] Antonio de Souzaa, C.; Westphalla, B.C.; Machadob, B.R.; Sobrala, M.B.J.; Vieira, S.G. (2020): Hybrid approach to intrusion detection in fog-based IoT environments. Journal Pre Proof.

[5] B. R. Rajakumar. (2013): Impact of Static and Adaptive Mutation Techniques on Genetic Algorithm. International Journal of Hybrid Intelligent Systems, 10, No. 1, pages: 11-22,DOI: 10.3233/HIS-120161.

[6] Bagaa, M.;Taleb, T.; Bernabe,B.J.; AntonioSkarmeta. (2020): Machine Learning Security Framework for IoT Systems. IEEE ACCESS.

[7] Bhayo,J.; Hameed,S.; Shah,A.S. (2020): An Efficient Counter-Based DDoS Attack Detection Framework Leveraging Software Defined IoT (SD-IoT).IEEE ACCESS.

[8] Chakkaravarthy, S.S.; Sangeetha, D.; Cruz, M.V.; Vaidehi,V.; Raman, B. (2020): Design of Intrusion Detection Honeypot Using Social Leopard Algorithm to Detect IoT Ransomware Attacks. IEEE ACCESS, VOLUME 8.

[9] de Assis, M.O.V;. Carvalho, F.L.; Rodrigues, J.J.P.C.; aime Lloret, Proença Jr,L.M. (2020): Near real-time security system applied to SDN environments in IoT networks using convolutional neural network. IEEE ACCESS.

[10] Entezari-Maleki, R.; Rezaei, A.; Minaei-Bidgoli,B. (2009): Comparison of Classification Methods Based on the Type of Attributes and Sample Size. Department of Computer Engineering.

[11] Farivar, F.; Haghighi, S.M.; Jolfaei,A.; Alazab, M. (2019): Artificial Intelligence for Detection, Estimation, and Compensation of Malicious Attacks in Nonlinear Cyber Physical Systems and Industrial IoT. IEEETransactions on Industrial Informatics.

[12] Halbhavi,S.; Kodad,S.F.; Ambekar, S. K.; Manjunath,D. (2019).Enhanced Invasive Weed Optimization Algorithm with Chaos Theory for Weightage based Combined Economic Emission Dispatch. Journal of Computational Mechanics, Power System and Control, 2,No.3, pp.19-27.

[13] Hassan,Z.; Mehmood,A.; Maple,C.; Khan,A.M.; Aldegheishem,A. (2020): Intelligent Detection of Black Hole Attacks for Secure Communication in Autonomous and Connected Vehicles.IEEE ACCESS, Volume 8.

[14] Hossain,M.; Xie,J. (2020): Third Eye: Context-aware Detection for Hidden Terminal Emulation Attacks in Cognitive Radio-enabled IoT Networks. IEEE Transactions on Cognitive Communications and Networking.

[15] https://www.bmc.com/blogs/data-normalization/

[16] Jadhav,N.A.; Gomathi,N. (2019): DIGWO: Hybridization of Dragonfly Algorithm with Improved Grey Wolf Optimization Algorithm for Data Clustering. Multimedia Research, 2,No.3, pp.1-11.

[17] Khan,Y.A.; Latif,R.; Latif, S.; Tahir,S.; Batool,G.; Saba,T. (2019): Malicious Insider Attack Detection in IoTs Using Data Analytics. IEEE Access.

[18] Kim,A.; Oh,U.; Ryu,J.; Lee,K. (2020): A Review of Insider Threat Detection Approaches With IoT Perspective. IEEE ACCESS, Volume 8.

[19] Kumar,P.; Gupta, P.G.; Tripathi, R. (2021): An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for IoMT networks. Computer Communications, 166, pp. 110-124.

[20] Li, B.;Ye, R.;Gu, G.; Liang, R.; Liu,W.; Cai, K. (2020): A detection mechanism on malicious nodes in IoT. Computer Communications

[21] Maithem, M.; Ghadaa,A. (2021): Network intrusion detection system using deep neural networks.J.Phys:Conf.Ser.1804 012138,2021.

[22] Mandal,K.; Rajkumar,M.;Ezhumalai, P.;Jayakumar,D.;Yuvarani,R.(2020): Improved security using machine learning for IoT intrusion detection system. ScienceDirect.

[23] Manimurugan, S. (2021): IoT-Fog-Cloud model for anomaly detection using improved Naïve Bayes and principal component analysis. Journal of Ambient Intelligence and Humanized Computing.

[24] Mirjalili, S.; Andrew Lewis,A. (2016): The Whale Optimization Algorithm. Advances in Engineering Software, 95, pp. 51-67.

[25] Mirjalili,M.S.; Lewis,A. (2014): Grey Wolf Optimizer.Advances in Engineering Software, 69, pp.46–61.

[26] Mirjalili,S. (2015):Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowledge-Based Systems, 89, pp. 228-249

[27] Mohammed,H.; Hasan,R.S.; Awwa,F. (2020): Fusion-On-Field Security and Privacy Preservation for IoT Edge Devices: Concurrent Defense Against Multiple Types of Hardware Trojan Attacks. IEEE ACCESS, Volume 8.

[28] Mortazavi,A. (2019): Interactive fuzzy search algorithm: A new self-adaptive hybrid optimization algorithm. 81, pp. 270-282.

[29] Navale,G.S.; Mali, N.S. (2020): Self-Adaptive Optimization for Improved Data Sanitization and Restoration. International Journal of Uncertainty, 28, No.03, pp.391-420.

[30] Ngo,D.Q.; Nguyen, T.H.; Le, H.V.; Doan,Hieu Nguyen,H. (2020): A survey of IoT malware and detection methods based on static features. ICT Express.

[31] Rahman,A.M.; Asyharia, L.S.; Leong,G.B.; Satrya, M.; Tao,H.; Zolkipli,F.M. (2020): Scalable Machine Learning-Based Intrusion Detection System for IoT-Enabled Smart Cities. Journal Pre-proof.

[32] Rao,V.R.; Keesari,K.H. (2020): A Self Adaptive population Rao algorithm for optimization of selected bio-energy systems. Journal of Computational Design and Engineering, 8, pp. 69-96.

[33] Samy, A.; Yu, H.; Zhang ,H. (2020): Fog-based Attack Detection Framework for Internet of Things Using Deep Learning. IEEE TRANSACTIONS and JOURNALS.

[34] Thamilarasu, G.; Odesile, A.; Hoang, A.(2020): An Intrusion Detection System for Internet of Medical Things. IEEE ACCESS, VOLUME 8.

[35] Wagh, B.M.; Gomathi, N. (2019): Improved GWO-CS Algorithm-Based Optimal Routing Strategy in VANET.Journal of Networking and Communication Systems,2,No.1, pp.34-42.

[36] Wang, Gai-Ge, Deb, Suash, Coelho, Leandro. (2015): Elephant Herding Optimization. DOI:10.1109/ISCBI.2015.8,

[37] Yılmaz, Selim, Sen, Sevil. (2020): Electric fish optimization: a new heuristic algorithm inspired by electrolocation. Neural Computing and Applications. 32. 10.1007/s00521-019-04641-8.

[38] Zhu,Y.; Jiang,W.; Kong,X.; Quan,L.; Yongshun Zhang,Y. (2017): A chaos wold optimization algorithm with self-adaptive variable step-size. AIP Advances,7.

Authors Profile:

Aftab Alam Abdussami is a Ph.D. student in Computer Application at the Integral University. He holds a Master's degree in computer from Sikkim Manipal University (2006). His area of expertise is Network Security, Cyber Security, Intrusion Detection System, Intrusion Prevention System, Internet of Things, Fog Computing, Ethical Hacking, Artificial Intelligence, Deep Neural Network. He holds many professional certifications like, C|CISO, CEHv9, CCNA R&S, CCNA Security, MCP and Certified Novell Engineerv5. He holds instructor certification from Cisco for CCNA (R&S) and CCNA Security.

Dr. Mohammed Faizan Farooqui currently working as Associate Professor in the Department of Computer Application. His area of expertise is requirement engineering, web technology, operating system, DBMS, computer organization, graph theory, computer graphics and animation and data warehousing & data mining