A Hybrid Approach to Mitigate False Positive Alarms in Intrusion Detection System

Sachin and C. Rama Krishna

Abstract The aim of intrusion detection systems (IDSs) is to detect the malicious traffic and dynamic traffic which changes according to network characteristics, so intrusion detection system should be adaptive in nature. Many of IDS have been developed based on machine learning approaches. In proposed approach, experiment have been carried out on KDD-99 dataset with three classes DoS attack, other attacks and normal (without any attack). Paper checks the potential capability of optimization-based features with artificial neural network (ANN) classifier for the different types of intrusion attacks. A comparative analysis with ANN and other optimizer with ANN has been carried out. The experimental results show that the accuracy of intrusion detection using particle swarm optimization with genetic algorithm (PSO_GA) improves the results significantly by reducing false positive alarms and also improve individual class detection.

Keywords PSO · IDS · ANN · GA · KDD99 · IDPS

1 Introduction

Intrusion is an action executed to break the security of one's system and to misuse. Mainly there are two threats in any system first is malware and other is intruder. Intruder may be defined as a threat which always used to break the system and to mislead the system [1]. For the solution of intruder many researchers used to introduce a detection system termed as intrusion detection system. This paper gives a spotlight on the intrusion detection system and its working and how one can enhance its working by reducing false positive alarms. Researcher provides many techniques

Sachin (✉) · C. Rama Krishna
Department of Computer Science and Engineering, National
Institute of Technical Teachers Training and Research, Chandigarh, India
e-mail: sachin.cse@nitttrchd.ac.in

C. Rama Krishna
e-mail: rkc_97@yahoo.com

S. Smys et al. (eds.), *International Conference on Computer Networks and Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 15, https://doi.org/10.1007/978-981-10-8681-6_77

like firewall, encryption [2] to protect the interior system from intruder or intrusion but because of its some drop out intruder make their way to affect the system without any harm to itself [2].

Now a day it is very difficult to make any system free from intrusion. The main function of IDS is to detect the unknown or abnormal activity in a particular system and to resolve this activity in very less interval of time [3]. IDS are used to protect or prevent various penetration and inner structure of computers. This system consist of several hardware and software which is used to determine unexpected events which is going to give an indication like attack is going to happen, attack is happening in your system or attack is happened in your system. These are such indication given by the IDS. It can be classified as its working type of the system like it warns before attack or it can warn while attack is in process or it warns after attack [3]. There are three components of IDS (a) sensor which is used to generate events and to sense the traffic of network or activity of the system (b) console which is used to control sensor, events and alerts and (c) detection engine which is used for the generation of alerts after receiving variation from security events. It is also used to maintain the data of sensors' events in any database.

This paper gives its contribution towards the enhancement of IDS system and its working on intrusions. KDD-99 is used in this paper as a data set having 41 features for analysis. The main goal of this paper is to attain accuracy in the working of IDS by mitigating the false positive alarms. This work represents the various processes processed for integrating IDS. Many machine learning approaches whether it is supervised or unsupervised techniques are utilized to enhance the efficiency of intrusion detection system which is explained clearly in Sect. 2. Organization of rest paper is as follows; Sect. 2 with related work, Sect. 3 gives the description about the proposed work, analysis and evaluation is explained in Sect. 4 and at the end summary of whole work is given in Sect. 5.

2 Related Work

Machine learning is a study that permits the computer or any system to learn without being programmed. There are two types of approaches which are used to enhance the working of IDS system. In supervised approach, they are capable to create the function from the given training data. There are several techniques given by several researchers. To apply any technique or approach it is important to have a study on the IDS system, its function appropriately so, [4] gives a brief study on the function of IDS which mainly focuses on four major classes of detection methods (a) classification (b) statistical (c) information theory and (d) clustering. Ahmed et al. [4] Also used to spotlight on the problems ascertain during its function. To develop the efficiency in the function of IDS [5] gives a fuzziness based semi-supervised learning approach which is termed as SSL. SSL is used to improve the performance of classifier for IDS where unlabeled samples are assisted with some other supervised learning approach. To improve the security of IDS [6] provides a deep neural network which is used to alert the system about attack and then sensor used to recognize the malicious attack. There is some other supervised approach to develop IDS like k-NN, fuzzy K-NN etc.

Ambusaidi et al. [7] Proposes a combination of IDS with LSSVM with dataset KD-99 which is also used in this paper for enhancing the intrusion detection system. Now if we talk about the un-supervised learning approach, it is a method of machine learning where a model is perfect for observations. There are number of approach like hybrid approach, domain approaches etc. Erin Liong et al. [8] Provides an un-supervised approach for learning purpose termed as DH approach which stands for deep hashing approach. Patel and Jhaveri [9] gives a study on the several machine learning techniques like ANN, SVM, Q-learning and Bayesian network to recognize the malware nodes and also recognize the misbehaving of nodes in the system. A xeromorphic cognitive approach is given by Alom and Taha [10] to detect intrusion in cyber security with the help of deep learning. These are some techniques given by researcher to enhance the IDS and make it efficiently.

3 Proposed Work

As discussed in Sect. 1 this paper works on KDD-99 data set. In this section, KDD-99 dataset with 41 features is processed for optimization, and learning to attain accuracy, precision and recall. Whole methodology is proposed into three phase, in first phase implementation of KDD-99 data set, in second phase optimization is processed on the feature and at last they allow for learning in third phase. KDD-99 dataset is a bench mark dataset and recognize by many users. For testing we use to select 10% of KDD-99 which contains 41 features from KDD-91. Where 10% KDD-dataset contains 494,021 connections and we set 311,030 connections for our work. These connections are labeled as normal or attack which classified into four categories: Denial of service which is termed as (DOS), Probe (port scanning) Unauthorized access to root person termed as U2R and last unauthorized remote login to machine [11]. There are three kind of features (1) basic feature (2) content based feature and (3) time based feature [11]. These features are labeled like $X_{1,}$ X_2 … X_n. After labeling these features are used for optimization in our work.

Optimization is done by two algorithm PSO and learning approach. Initially feature is applied on PSO for optimization to obtain fitness value. If any feature is not optimized by PSO then these un-optimized feature is allowed on learning algorithm for further optimization process. If all the features are optimized by these two algorithms, then it further moves to check convergence. If convergence is not accepted, then these feature again used for optimization by PSO and Learning algorithm. If convergence check is ok, then we move to next phase which is learning.

After labeling the dataset feature optimization is done by PSO which stands for particle swarm optimization where swarm denotes to collection of particles. In the process, PSO Particles floats through the hyper-dimensional search space. PSO is a population based search algorithm which is based on simulation on the social behavior of birds within a flock. Variation in the position of particle in a search space is depend upon the psychological tendency of each particle to imitate the development of other. In PSO swarm consists of a set of particles where each particle demonstrates a potential solution [12]. The position of particle is varying with respect to its own experience and of the neighbor particles. PSO is used to optimize the value of objective function. Every particle in space used to mobile to find the point where optimized function is obtained where 'z' is the position of particles in time 'h' having velocity 'u'. Every particle has its local and global best position in the space. Global best position is the position of a particle which is close to optimal value and all the particles move towards the global best position. The global position of particle will vary with the motion of particles. The changed position is obtained using equation [12].

$$z_r(h) = z_r(h-1) + u_r(\mathrm{h}) \tag{1}$$

$$u_r(h) = Iu_r(h-1) + L_1V_1\left(z_{p\mathrm{best}_r} - z_r(h)\right) + L_2V_2\left(z_{g\mathrm{best}} - z_r(h)\right) \tag{2}$$

where, 'I' is the weight of inertia and L_1 and L_2 are the learning factors and V_1 and V_2 are the random values. Using Eqs. 1 and 2 we can obtain fitness value of the feature after optimization. If all the features are optimized, then we proceed to further step that is convergence check otherwise we use genetic algorithm to optimize the feature which are left by PSO algorithm. In this work genetic algorithm is utilizes to obtain the optimal and near-optimal threshold for feature selection. Genetic algorithm is a technique which is used to evaluate true and approximate solution to optimization and search problem. In this paper GA is used for find the solution of optimization.

In GA, the main and initial step is to demonstrate the problem in such a manner that GA able to resolve the problem as it works on binary coding. In GA, chromosomes are used to gradually evolve with the help of biological operations. It might be possible to obtain greater feature from big data but it takes extra time and computational steps hence, we used GA for selected feature for random generation [13]. After initializing, every individual chromosome are computed by fitness function and according to this value of chromosome which is associated with fitness gives better result as compared to un-fit value. Crossover permits the search to determine the efficient way to obtain solution and optimization, and also allow chromosomal material from different parent to be combined in a single child. Crossover gives a way to introduce new information into the population [13, 14]. Here, GA uses Roulette selection for selecting best features from the feature sets. Roulette selection is similar to the game of roulette, every features get a slice of wheel but the features which is more fit gets larger slice as compare to other. In short, they are chosen in terms of their fitness value:

$$p_s = \frac{S_i}{\sum_k S_k} \tag{3}$$

After optimization with Genetic algorithm if, all the features are optimized they proceed to further step otherwise the optimization with GA is repeated. This procedure is repeated unless the entire features are optimized. After optimization every

Fig. 1 Labeling of KDD-99 data set

feature convergence is tally if it is ok then our feature are allow for learning by neural networks [13, 14].

Genetic algorithm

```
Simple_Algorithm
{
Initialize the population;
Evaluate Fitness Function;
While ( The number of generation > maximum number)
        {
Selection ; // Natural selection, Survival of fittest
Crossover ; // Reproduction, Propagate favorable characteristics
Mutation; // Mutation
            Calculate fittest function;
             }
  }
```

Now, optimized feature are permitted to learn from neural networks which is our phase3. From the reference of bio-logical neurons, there is an introduction of artificial neural network which is a set of neurons similar to the neuron in human nervous systemin computer system. These neurons are used to learn patterns and relationship in the given data. In this work neuron network are used to learn the feature which is optimized from PSO and GA. NN do not require any explicit codes of any particular problems this is because of the learning rules in NN. These rules helps network to gain knowledge from the given or available dataset and implement this knowledge in problems according to requirement. NN is a network which contains number of nodes having node function related to it that is used to evaluate the output from the local parameters. This node function depends upon the variation in local parameters. Hence it is termed as information processing system where neurons are used to process the information and signals are transmitted through connection links. These links have some weight which is multiplied with respect to the incoming signals. Neural net may be of single layer or multiple layers [15] (Fig. 1).

Optimized features are converted into neurons and applied as input like N_1 and N_2 having weight M_1 and M_2. It may be single layer or multi-layer of neurons. In the input layer raw information i.e. optimized features are input in the networks. In hidden layer evaluation is done between input data and hidden layer in terms of their connecting weights. On the basis of activity of neurons in hidden layer output varies (Fig. 2).

Net weight is calculated as:

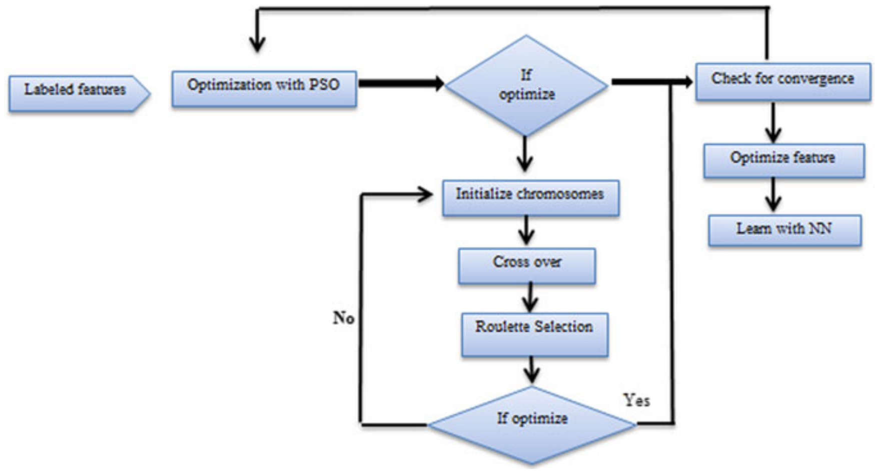

Fig. 2 Optimization process

$$\mathrm{Net} = \mathrm{N}_1\mathrm{M}_1 + \mathrm{N}_2\mathrm{M}_2 \tag{4}$$

and can be written as

$$\mathrm{Net\,input} = \sum\nolimits_m N_m M_m \tag{5}$$

```
Algorithm
Step 1: Input the KDD-99 data set having 41 features and label all the features.

Step 2: Optimize these features with PSO to obtain fitness value using equation1 and 2.

Step 3: If we obtain optimized value by PSO then proceed further for convergence check otherwise apply genetic algorithm for optimization and uses unless we obtain optimized features.
After obtaining optimized feature use to check convergence of whole feature set.
Rolette selection is used for feature selection in GA having equation 4.
Step 4: Check convergence of entire feature set.
If ( convergence = yes)
{
Learn the feature with neural networks and evaluate accuracy, precision, F-measure and recall.
Else
{
Repeat step 3
}
```

Table 1 Evaluation of the parameters among ANN, ANN with GA, ANN with PSO and ANN with GA_PSO

	ANN	ANN with GA	ANN with PSO	ANN with GA_PSO
Accuracy	89	92.23	92.34	94.23
Precision	88	89.23	90.32	92.33
Recall	87	88.56	91.26	96.33
F-measure	85	88.25	86.23	91.13

Table 2 The attack type from KDD CUP 99 dataset

Normal	Dos	R2L	U2R	Probe
Normal	Smurf Processtable Pod Land	PHF Xlock Send-mail Guess_password	Root-kit Eject Perl Buffer overflow	Portsweep Satan Saint M-scan

4 Description of Dataset

As discussed above experiments are executed using KDD-99 having 41 feature sets. These features are used for optimization and then learning and now they are use to analyze in terms of attack. In this work, we use to evaluate the accuracy rate in an intrusion detection system. In the analysis we take data on the basis of number of intrusions. Attacks generally fall into four categories (1) Dos, (2) Probe, (3) R2L (4) U2R. In our analysis we uses three categories (1) other attack which consist of probe, R2L and U2R (2) DoS-attack, and (3) Normal attacks (non-attacks). In this work we evaluate the accuracy, precision, recall, and F-measure in various cases:

Case 1: Evaluation of accuracy, precision, f-measure and recall is given by ANN individually, ANN with GA, ANN with PSO and ANN with combined GA_PSO which is represented in Table 1. In this case we evaluate the efficiency of IDS by applying ANN individually or with GA and PSO or by hybrid of both algorithms with ANN.

Case 2: In this case, we evaluate the efficiency for single ANN on three attack condition (a) other attack consist of probe, R2L and U2R, (b) Dos attack, and (c) Normal or non-attacks condition. Similarly we evaluate the efficiency for ANN with PSO, ANN with GA and at last ANN with both GA_PSO. Here we evaluation efficiency of algorithms in terms of accuracy, precision, recall and F-measure. If we spot some light on attacks we are considered. Dos attack which stands for denial of service attack in Dos attack hidden attacks is done by user which is shown in the system. This type of attack may be done by single intruder or a group of intruders. It makes the system unavailable to its real user. Probe attack is a kind of attack where intruder used to break the security by trial methods. R2L attack stands for remote to user attack. And at last U2R attacks it is the type of attack where intruder starts on the system as a normal user and spoil all the activities of the systems [16] (Fig. 3).

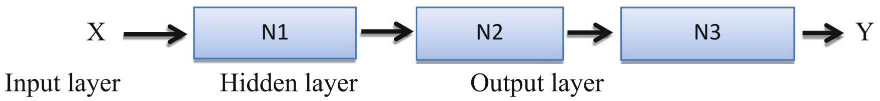

Fig. 3 Layer of MNN

Table 3 The static data to analyze the efficiency of the approaches for above discussed attack

Algorithm type	Types of attack	Accuracy	Precision	Recall	F-measure
(a) ANN	Other attack	86.23	87.23	86	83
	Dos attack	87.23	88.33	88	84
	Normal attack	91.23	90.13	87	86
(b) ANN with GA	Other attack	91.13	87.23	86.75	87.23
	Dos attack	90.23	90.13	84.23	86.23
	Normal attack	94.11	89.99	94.23	89.13
(c) ANN with PSO	Other attack	93.13	89.23	92.23	84.23
	Dos attack	92.13	88.13	89.13	83.23
	Normal attack	90.13	84.23	92.13	87.34
(d) ANN with GA and PSO	Other attack	95.62	90.23	92.33	89.23
	Dos attack	89.23	89.23	93.33	90.13
	Normal attack	96.23	93.13	99.23	90.23

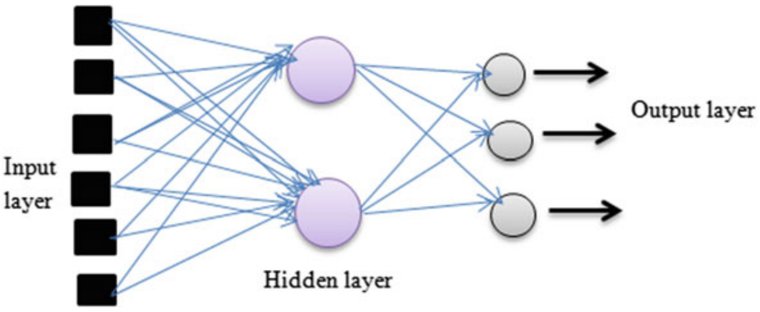

Fig. 4 Various layers in neural networks

4.1 Experimental Results

In this section we analyze the statistical data by simulation. Graphical result of Tables 1 and 3 is given by simulation process (Fig. 4; Table 2).

Figure 5 shows the simulated analysis of Table 1 in terms of accuracy, precision, recall, and F-measure. In this figure analysis on efficiency is demonstrated from all the

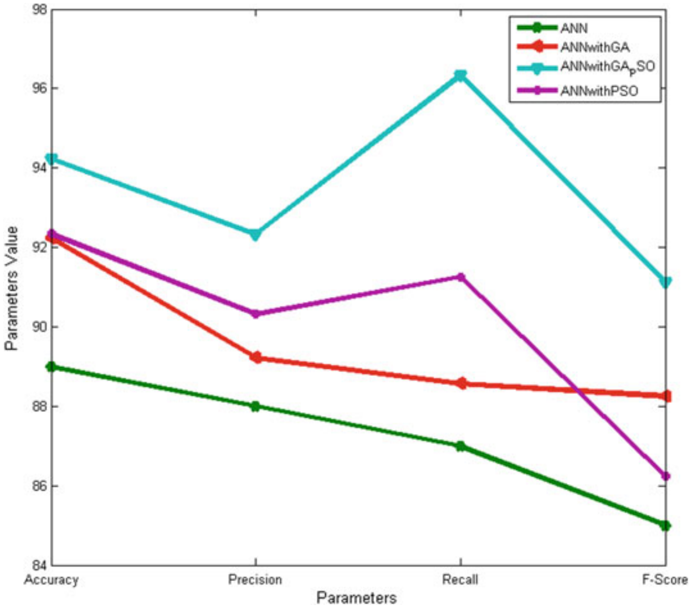

Fig. 5 Simulated graph of Table 1

four algorithm that are ANN represented by green line, ANN with PSO represented by purple line, ANN with GA represented by red line and ANN with both PSO and GA represented by blue line. Analysis demonstrates that ANN with both SPO and GA gives better result in terms of all the four parameters (accuracy, precision, recall, and F-measure).

Observation 1: In Fig. 5 Parameters analysis of different classifier and proposed approach has been shown. In analysis, parameters like precision, recall, accuracy and F-measure vary according to classifier but one analysis about proposed approach (PSO with GA in neural network) is very clear that it shows significant improvement in all parameters. If analysis have to be made only over proposed approach then recall parameter have shown significant improvement then other parameters. So it gives clear indication of reducing false negative rate and attacks identification is effective in proposed approach because of optimize weight given by PSO_GA approach.

Figures 6 and 7 gives the analytic result in terms of accuracy (represented by black line), recall (represented by red line), precision (represented by blue line), and F-measure (represented by green line) of two approaches that are ANN and ANN with GA for the parameters (Dos attack, other attack, and Non or normal attack).The entire four graphs demonstrate the better efficiency of the algorithm ANN with PSO_GA.

Observation 2: In Fig. 6, depth analysis of all three classes in ANN and ANN_GA has been shown. In this analysis, we have tried to show what is the significance of our approach. We continue this discussion in observation 3 also. So, at first point if analysis is done over normal class and on other class which not any attack

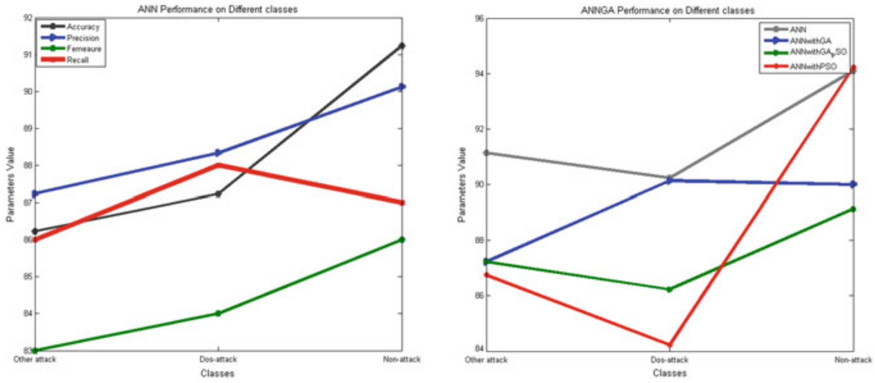

Fig. 6 Analysis with ANN and ANN _GA

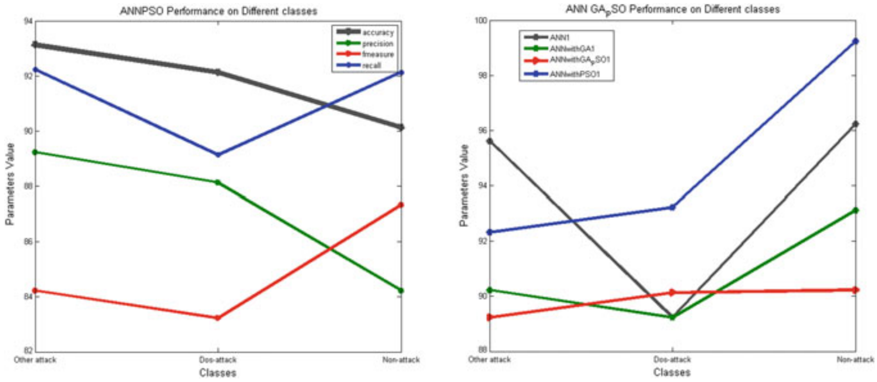

Fig. 7 Analysis with ANN _PSO and ANN with PSO_GA

in both cases ANN and ANN with GA performed well compared to other parameters like precision, recall, and F-measure but ANN_GA still has better accuracy than ANN, so feature weighted by optimization somehow performs well because of reducing overlapping information learning. If analysis is done through DOS attack only it also shows higher accuracy in ANN with GA, so we can conclude Feature optimize weight is better approach. So how it is going to improve optimization that have been discussed in next part.

At last from the whole analysis it can be concluded that algorithm ANN with PSO_GA gives better result for all the attacks we examined in our work.

Observation 3: In Fig. 7, analysis continues from observation 2 and tried to find out the significance of optimization and improvement in detection of different classes by classification. If analysis have to be made then both graph shows the effective recall but only for normal class, which reduces the false positive rate. This improvement can be seen with all classes like DOS attack and other attacks. But effective results have been seen in the in proposed approach where the improved detection have been done for other attacks too. So PSO optimization is effective but PSO with GA i.e. the proposed approach gives more improved results in other attack and normal class.

5 Conclusion

This paper investigates the optimization weight of feature and how to reduce the statistically overlapping between features and improving the optimization by hybridization is worthy or not. In the proposed approach optimization base features with artificial neural network has been used and experiment shows that this hybrid approach not only improves cumulative accuracy, precision, recall, and F-score but also improves all individual parameters among three classes (i.e. DOS attack, Normal and other attacks).

References

1. Ashfaq, R.A.R., et al.: Fuzziness based semi-supervised learning approach for intrusion detection system. Inf. Sci. **378**, 484–497 (2017)
2. Lin, W.C., Ke, S.W., Tsai, C.F.: CANN: An intrusion detection system based on combining cluster centers and nearest neighbors. Knowl. Based Syst. **78**, 13–21 (2015)
3. Harendra, V., Mirza, S., Mali, N.: Intrusion detection system. Int. J. Adv. Res. Eng. Sci. Technol. **3** (2016)
4. Ahmed, M., Mahmood, A.N., Jiankun, H.: A survey of network anomaly detection techniques. J. Netw. Comput. Appl. **60**, 19–31 (2016)
5. Ashfaq, R.A.R., et al.: Fuzziness based semi-supervised learning approach for intrusion detection system. Inf. Sci. **378**: 484–497 (2017)
6. Kang, M.J., Kang, J.W.: Intrusion detection system using deep neural network for in-vehicle network security. PLoS ONE **11**(6), e0155781 (2016)
7. Ambusaidi, M.A., et al.: Building an intrusion detection system using a filter-based feature selection algorithm. IEEE Trans. Comput. **65**(10): 2986–2998 (2016)
8. Erin Liong, V., et al.: Deep hashing for compact binary codes learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2475–2483 (2015)
9. Patel, N.J., Jhaveri, R.H.: Detecting packet dropping nodes using machine learning techniques in Mobile ad-hoc network: A survey. In: International Conference on 2015 Signal Processing and Communication Engineering Systems (SPACES), pp. 468–472. IEEE (2015)
10. Alom, MZ., Taha, T.M.: Network intrusion detection for cyber security on neuromorphic computing system. In: International Joint Conference on 2017 Neural Networks (IJCNN), IEEE (2017)
11. Tavallaee, M., et al.: A detailed analysis of the KDD CUP 99 data set. In: IEEE Symposium on 2009 Computational Intelligence for Security and Defense Applications, CISDA 2009. IEEE (2009)
12. Sierra, M.R., Coello, C.A.C.: Improving PSO-based multi-objective optimization using crowding, mutation and e-dominance. In: Evolutionary Multi-criterion Optimization, vol. 3410, Springer, Berlin, Germany (2005)
13. Kim, Kyoung-jae, Han, Ingoo: Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. Expert Syst. Appl. **19**(2), 125–132 (2000)

14. Lipowski, A., Lipowska, D.: Roulette-wheel selection via stochastic acceptance. Physica A: Stat. Mech. Appl. **391**(6), 2193–2196 (2012)
15. Vassiliadis, S., et al.: Artificial neural networks and their applications in the engineering of fabrics. In: Woven Fabric Engineering, InTech (2010)
16. Paliwal, S., Gupta, R.: Denial-of-service, probing and remote to user (R2L) attack detection using genetic algorithm. Int. J. Comput. Appl. **60**(19), 57–62 (2012)

Sachin studying in Department of Computer Science and Engineering, in National Institute of Technical Teachers Training and Research, Chandigarh, India. His research area includes denial of service attacks and network security.

C. Rama Krishna working as a professor in Department of Computer Science and Engineering, in National Institute of Technical Teachers Training and Research, Chandigarh, India. His research area includes denial of service attacks and network security.