IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

Hybrid Particle Swarm Optimization and Fuzzy *C*-Means Clustering for Network Intrusion Detection

Partha Sarathi Bhattacharjee[1*], Arif Iqbal Mozumder[2] , Shahin Ara Begum[3]

[1,2]Assam University, Silchar, Cachar, Assam, India
[3]Assam University, Silchar, Cachar, Assam, India

[*]*Corresponding Author: psbkls@gmail.com*

Available online at: www.ijcseonline.org

Accepted: 12/Sept/2018, Published: 30/Sept/2018

Abstract - Intrusion Detection systems (IDS) play an important role in network security and protection. Intrusion detection system uses either misuse or anomaly based techniques to identify malicious activities. To detect malicious activity, misuse detection systems is used to identify signatures or previously known malicious activities. On the other hand, anomaly based systems is used to identify unknown attacks. Intrusion detection system is now an essential tool to protect the networks by monitoring inbound and outbound activities and identifying suspicious patterns that may indicate a system attack. In recent years, some researchers have employed data mining techniques for developing IDS. In this paper, hybrid Particle Swarm Optimization (PSO) and Fuzzy *c*-means clustering for network Intrusion Detection is proposed to identify intrusion over NSL-KDD dataset. An attempt has been made to cluster the dataset into normal and the major attack categories i.e. DoS, R2L, U2R and Probe. The experimental results demonstrate the efficiency of the proposed approach.

Keywords - IDS, Fuzzy *c*-means Algorithm, PSO, Mutual Information, NSL-KDD Dataset

I. INTRODUCTION

Intrusion detection is a technique to monitor the events that occur in a computer system or network and analyze them for identifying the signature of intrusion. Intrusions are attempts to compromise the confidentiality, integrity or availability of computer or network. They are caused by attackers accessing a system from the internet by unauthorized users or by authorized user who misuse the privileges given to them. Anomaly detection and misuse detection are two general approaches to computer intrusion detection system. Misuse detection generates an alarm when a known attack signature is matched on the other hand anomaly detection identifies activities that deviate from the normal behaviour of the monitored system and thus has the potential to detect network attacks [1][2].

The network intrusion detection by using data mining techniques was proposed by [3]. It was proposed an unsupervised anomaly detection technique to assign a score to each network connection that reflects how anomalous the connection is and an association pattern analysis based module summarizes those network connections that are ranked highly anomalous by the anomaly detection module. Genetic algorithm is used to perform the detection of various types of network intrusions [4][5]. Network intrusion detection system based on data mining technology was proposed by [6] and [7]. Algorithm based on cascade Support Vector machine and graph based neural network for intrusion detection system were proposed by [8]. Intrusion Detection System based on Fuzzy clustering was proposed by [9]. Data mining methods for IDS consists of three steps *viz.* data pre-processing, feature selection and clustering. The duplicate samples are eliminated from the data set in pre-processing step followed by selection of most discriminated features by using principle component analysis in feature selection step. Finally, FCM is applied for clustering of selected features. An algorithm based on Artificial Immune System is proposed by [10] for anomaly detection in intrusion detection system. [11] proposed layered approach with certain rule learning classifier such as Genetic Algorithm, Ant search and Particle Swarm Optimization (PSO) respectively to detect network intrusion. Data mining techniques for identification of normal and attack data present over the network were proposed by [12]. The authors applied decision tree technique C4.5 to identify the attacks in the network.

The aim of the paper is to identify anomaly-based intrusion using Particle Swarm Optimization (PSO) and Fuzzy *c*-means algorithm. In the experiments, NSL-KDD 20 percent data set is used for detection of anomaly attack. This paper proposes PSO as a feature reduction tool and firstly it reduces the features of NSL-KDD 20 percent dataset and then Fuzzy *c*-means clustering is applied to identify the unknown attacks. The results obtained with the proposed method are compared with an existing method of feature selection by Mutual Information (MI) and then clustering by

the FCM over the same dataset. The rest of the paper is organized as follows: Section II gives a brief overview of PSO and MI. Section III presents the experimental setup along with the clustering algorithm description. Section IV presents the experimental results and discussions and finally Section V concludes the paper.

II. (A) PARTICLE SWARM OPTIMIZATION

PSO is an evolutionary computation technique inspired by behaviour of animals such as bird flocking or swarm behavior in fish. It is initiated with a random population called swarm, the PSO algorithm tries to find feasible solutions called particles. Each particle of the swarm represents a candidate solution in the *d*-dimensional search space where particle moves [13]. The PSO is a meta-heuristic technique, generally used in optimization problem. Intrusion Detection System (IDS) is one of the domains, where PSO can be used as an optimizer in order to improve the efficiency of the IDS.

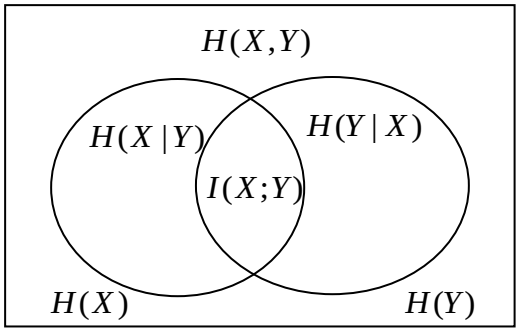
II. (B) MUTUAL INFORMATION

The Mutual Information of two random variables is a technique to determine the mutual dependence between the two variables. It assesses the amount of information obtained about one random variable, through the other random variable. The concept of mutual information is intractably linked to that of entropy of a random variable which defines the amount of information apprehended in a random variable [14].

The Mutual Information is defined as follows [15]:

$$I(X;Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (1)$$

Pictorially it can be defined by the following diagram:



where, the area contained by both circles is the joint entropy H(X,Y). The circle on the left is the individual entropy H(X), with the conditional entropy H(X|Y). The circle on the right is H(Y), with the conditional entropy H (Y|X). The intersection point is the mutual information I(X;Y).

III. EXPERIMENTAL SETUP

The experiments are carried out on MATLAB software to identify network intrusions. For identifying intrusions, firstly the network attack is identified by Fuzzy *c*-means clustering for NSL-KDD 20 percent dataset with 42 features. Then the features of NSL-KDD 20 percent dataset are reduced by MI and clustered using Fuzzy *c*-means clustering algorithm. Finally PSO has been used as a feature reduction tool to reduce the features of NSL-KDD 20 percent dataset and then Fuzzy *c*-means clustering is applied to find the optimal results of attack detection.

Dataset

NSL-KDD is a dataset with a large amount of quality data which follows the real time data and is used for training and testing of intrusion detection. The NSL-KDD dataset is the superior version of the KDD CUP99 dataset [16]. A detailed analysis on the NSL-KDD data set using various machine learning techniques is presented by Revathi and Malathi [17].

A thorough analysis of NSL KDD dataset is made using data mining based clustering algorithms like *k*-means and Fuzzy *c*-means clustering algorithms [18].

The different types of attack in NSL-KDD dataset are mentioned in Table 1. The normal traffic in the network is mentioned in the dataset as 'normal'.

Table 1: Different types of attack in NSL-KDD data set

Category	Attack Type
DOS	Smurf, Neptune, Back, Teardrop, Pod, Land
U2R	Buffer_overflow, Rootkit, loadmodule, perl
R2L	Warezclient, Guess_passwd, Warezmaster, Imap,ftp_write, Multihop, Phf, Spy
Probe	Satan, Ipsweep, Portsweep, Nmap

Feature Selection using Mutual Information (MI)

Selection of most discriminated features of NSL-KDD 20 percent dataset using Mutual Information is performed by Algorithm 1 [19].

Algorithm 1 MI based feature selection

Input: Initial set of NSL-KDD 20 percent dataset say Z.

Output: Discriminated features of NSL-KDD dataset say *S*.

Step 1: For each feature $f \in Z$ compute $I(C; f)$.

Step 2: Find the feature f that minimized $I(C; f)$: set $Z \leftarrow Z \setminus \{f\}$; set $S \leftarrow \{f\}$.

Step 3: Repeat until repeat until the desired number of features are selected.

a) For all couple of variable (f, s) with $f \in Z, s \in S$, compute $I(f, s)$, if it is not already available.

b) Choose feature *t* as the one that maximize

$$I(C; f) - \beta \sum_{s=S} I(f; S); \text{ set } Z \leftarrow Z \setminus \{f\} \quad (2)$$

Step 4: Return set *S* as output.

Feature Selection using Particle Swarm Optimization (PSO)
PSO based selection of most discriminated features of NSL-KDD 20 percent dataset is performed by Algorithm 2 [13].

Algorithm 2 PSO based feature selection

Input: The swarm size (*m*), acceleration constant (c_1 and c_2), weight inertia (*w*), maximum generation (*mg*), and fitness threshold (*ft*).

Output: Discriminated features of NSL-KDD 20 percent dataset

Step 1: Initialize: $pbest_i$=0, *gbest*=0, *Itr*=0 and initialize population with random positions
and velocities on *d*-dimensions (*d*=1,2,…, 42)

Step 2: **while** *itr*<*mg* or *gbest* <*ft* **do**
For *i*=1 to *m* do
fitness (*i*)=Evaluate (*i*)
if fitness (*i*)> fitness ($pbest_i$) **then**
fitness($pbest_i$)=fitness(*i*)
update $p_{id}=x_{id}$
end if
if fitness(*i*)> *gbest* **then**
gbest=fitness (*i*)
update *gbest=i*
end if
for each dimension *d* ***do***
update the velocity vector and particle position
end for
end for
itr=itr+1
end while

Step 3: Return most discriminated features

Fuzzy *c*-means Algorithm
Fuzzy *c*-means algorithm on discriminated features of NSL-KDD 20 percent dataset is Algorithm 3 [20].

Algorithm 3 Fuzzy *c*-means clustering for Intrusion detection

Input: The features of NSL-KDD dataset obtained by using Algorithm 1 and Algorithm 2 respectively.

Output: Detection of attacks in the network.

Step 1: Fuzzy *c*-means algorithm:

The objective function *Z* is minimized as follows:

$$Z = \sum_{i=0}^{n} \sum_{j=1}^{c} u_{ij}^{m} ||x_i - c_j||^2 \quad (3)$$

where *m* is any real number whose value is greater than 1 and less than ∞, u_{ij} is the degree of membership of x_i in the cluster *j*, x_i is the i^{th} value of n dimensional measured data, c_j is the center of dimension of the cluster.

Step 2: The cluster center matrix *c* is randomly initialized and fuzzy partition matrix *U* is created through an iterative optimization of the objective function shown above by updating of membership u_{ij} and the cluster centers c_j as follows:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (4)$$

$$c_j = \sum_{i=1}^{n} (u_{ij})^m x_i \Big/ \sum_{i=1}^{n} (u_{ij})^m \quad (5)$$

This iteration will stop when, $||u^{(k+1)} - u^{(k)}|| < €$ where value of € lies between 0 and 1 and closed to 0; and *k* are the iteration number.

Step 3: Initialize $U=[u_{ij}]$ *matrix,* $U^{(0)}$

Step 4: At *k*-step, calculate the centers vectors

$$c_j = \sum_{i=1}^{n} (u_{ij})^m x_i \Big/ \sum_{i=1}^{n} (u_{ij})^m \quad (6)$$

Step 5: Update $u^{(k)}$, $u^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (7)$$

Step 6: *If* $||u^{(k+1)} - u^{(k)}|| < €$ then STOP; otherwise return to step 3 where the value of € lies between 0 and 1 and closed to 0.

Step 7: The output from the Hybrid FCM clustering gives the result of attack detection for NSL-KDD 20 percent dataset.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Experimental results obtained with FCM for NSL-KDD 20 Percent dataset with 42 features is presented in Table 2 [18].

Table 2: Results of Fuzzy *c*-means for NSL-KDD Dataset with 42 features

Clusters	Iteration	Normal	DOS	Probe	R2L	U2R	Number of attacks in each cluster	Percentage of attack detection
Cluster-1	1	479	46	1031	0	0	1556	6.17
	2	479	46	1031	0	0	1556	6.17
	3	1629	339	42	35	0	2045	8.12
	4	10060	497	870	141	2	11570	**45.93**
Cluster-2	1	1629	339	42	35	0	2045	8.12
	2	1629	339	42	35	0	2045	8.12
	3	1369	8390	352	34	3	10148	**40.28**
	4	479	46	1031	0	0	1556	6.17
Cluster-3	1	1369	8390	352	34	3	10148	40.28
	2	10060	497	870	141	2	11570	**45.93**
	3	10060	497	870	141	2	11570	**45.93**
	4	1369	8390	352	34	3	10148	40.28
Cluster-4	1	10060	497	870	141	2	11570	**45.93**
	2	1369	8390	352	34	3	10148	40.28
	3	479	46	1031	0	0	1556	6.17
	4	1629	339	42	35	0	2045	8.12

Table 2 shows the maximum percentage of attack detection by Fuzzy *c*-means clustering algorithm for NSL-KDD 20 percent dataset with 42 features is 45.93%.

The experimental results and analysis of intrusion detection for NSL-KDD 20 Percent dataset with 10, 15, 20, 25, 30, 35 and 42 features using Hybrid MI-FCM is presented in Table 3, Table 4, Table 5, Table 6, Table 7, Table 8 and Table 9, respectively.

Table 3: The detection of attack by Hybrid Fuzzy *c*-means – MI algorithm for NSL-KDD 20 Percent dataset with 10

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	316	8168	1	0	262	8747	34.72
2	12813	1055	10	206	1553	15637	62.07
3	148	7487	0	0	103	7738	30.72
4	365	28	0	3	513	909	3.60

Table 4: The detection of attack by Hybrid Fuzzy *c*-means – MI algorithm for NSL-KDD 20 Percent dataset with 15 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	566	65	0	3	1098	1732	6.87
2	2617	422	3	12	283	3337	13.25
3	11119	8603	3	71	203	19999	79.38
4	579	319	5	128	1003	2034	8.07

Table 5: The detection of attack by Hybrid Fuzzy *c*-means – MI algorithm for NSL-KDD 20 Percent dataset with 20 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	11	7	0	0	499	517	2.05
2	10352	9146	11	201	1600	21310	84.59
3	2408	69	0	4	0	2481	9.85
4	1119	22	0	4	190	1335	5.30

Table 6: The detection of attack by Hybrid Fuzzy *c*-means – MI algorithm for NSL-KDD 20 Percent dataset with 25 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	12862	7823	11	198	1259	22153	87.94
2	23	26	0	1	13	63	0.25
3	12807	7764	11	198	1227	22007	87.36
4	554	1390	0	10	1017	2971	11.79

Table 7: The detection of attack by Hybrid Fuzzy *c*-means – MI algorithm for NSL-KDD 20 Percent dataset with 30 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	12813	1055	10	206	1553	15637	62.07
2	316	8168	1	0	262	8747	34.72
3	365	28	0	3	513	909	3.61
4	148	7487	0	0	103	7738	30.72

Table 8: The detection of attack by Hybrid Fuzzy *c*-means – MI algorithmfor NSL-KDD 20 Percent dataset with 35 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	479	46	0	0	1031	1556	6.18
2	10061	497	8	141	870	11577	45.95
3	1629	339	0	35	42	2045	8.12
4	1369	8390	3	34	352	10148	40.28

Table 9: The detection of attack by Hybrid Fuzzy *c*-means – MI algorithmfor NSL-KDD 20 Percent dataset with 42 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	316	8168	1	0	567	9052	35.93
2	12813	1055	10	206	1553	15637	62.07
3	365	28	0	3	513	909	3.6
4	148	7487	0	0	103	7738	30.72

The experimental results and analysis of intrusion detection for NSL-KDD 20 Percent dataset with 10, 15, 20, 25, 30, 35 and 42 features using Hybrid PSO-FCM is presented in Table 10, Table 11, Table 12, Table 13, Table 14, Table 15 and Table 16 respectively.

Table 10: The detection of attack by Hybrid Fuzzy *c*-means – PSO algorithm for NSL-KDD 20 Percent dataset with 10 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	120	19	0	0	690	829	3.29
2	12844	9160	11	206	1132	23353	92.7
3	1678	8126	0	15	286	10105	40.11
4	459	47	0	3	431	940	3.73

Table 11: The detection of attack by Hybrid Fuzzy *c*-means – PSO algorithmfor NSL-KDD 20 Percent dataset with 15 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	1127	26	0	4	689	1846	7.32
2	1890	61	0	4	0	1955	7.76
3	10068	9098	11	201	561	19939	79.15
4	486	53	0	0	1039	1578	6.26

Table 12: The detection of attack by Hybrid Fuzzy *c*-means – PSO algorithm for NSL-KDD 20 Percent dataset with 20 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	11082	8597	3	72	186	19940	79.15
2	923	312	5	127	1376	2743	10.88
3	1302	292	3	8	231	1836	7.28
4	448	46	1	3	565	1063	4.22

Table 13: The detection of attack by Hybrid Fuzzy *c*-means – PSO algorithm for NSL-KDD 20 Percent dataset with 25 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	566	45	0	3	246	860	3.41
2	32	37	0	0	907	976	3.87
3	12329	7759	11	196	1145	21440	85.11
4	550	1412	0	10	56	2028	8.05

Table 14: The detection of attack by Hybrid Fuzzy *c*-means – PSO algorithm for NSL-KDD 20 Percent dataset with 30 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	11830	6643	10	197	929	19609	77.84
2	12753	9172	11	207	1196	23339	92.64
3	448	46	1	3	565	1063	4.22
4	424	57	0	1	592	1072	4.25

Table 15: The detection of attack by Hybrid Fuzzy *c*-means – PSO algorithm for NSL-KDD 20 Percent dataset with 35 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	10691	573	8	142	881	12295	48.80
2	208	7751	0	0	600	8559	33.97
3	2260	484	3	65	667	3479	13.80
4	617	7260	0	3	312	8192	32.52

Table 16: The detection of attack by Hybrid Fuzzy *c*-means – PSO algorithm for NSL-KDD 20 Percent dataset with 42 features

Cluster	Attack Category					No. of attacks detected	% of attack detection
	Normal	DOS	U2R	R2L	Probe		
1	11830	6643	10	197	929	19609	77.84
2	424	57	0	1	592	1074	4.26
3	12753	9172	11	207	1196	23339	92.65
4	336	45	0	3	563	947	3.76

The comparative analysis of Hybrid MI-FCM with Hybrid PSO-FCM is presented in Table 17 and in Figure 1.

Table 17: Comparison of maximum percentage of attack detection by Hybrid MI-FCM and Hybrid PSO-FCM over the NSL-KDD 20 percent dataset

Number of features	Hybrid MI-FCM (maximum % of attack detection)	Hybrid PSO-FCM (maximum % of attack detection)
10 features	62.07	92.7
15 features	79.38	79.15
20 features	84.59	79.15
25 features	87.36	85.11
30 features	62.07	92.64
35 features	45.95	48.8
42 features	62.07	92.65

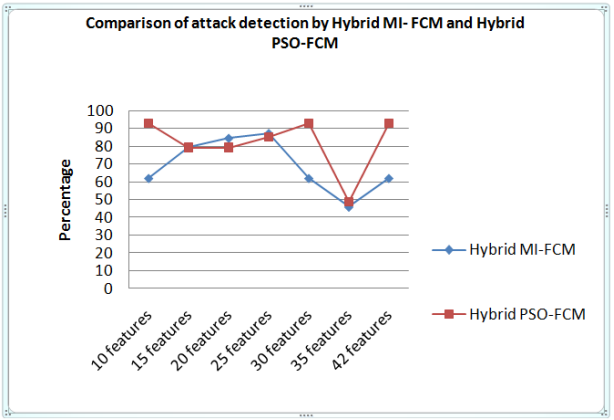


Figure 1: Comparison of Hybrid MI-FCM and Hybrid PSO-FCM Clustering for anomaly attack detection over the NSL-KDD 20 percent dataset

The experimental results (Table 17 and Figure 1) demonstrate that the proposed hybrid intrusion detection approach by PSO and FCM outperforms the hybrid approach based on MI and FCM for anomaly attack detection.

V. CONCLUSION

This paper presents an data mining approach based on Particle Swarm Optimization and Fuzzy *c*-means clustering algorithm to detect the major network attack categories i.e. Normal, DoS, R2L, U2R and Probe over the NSL-KDD dataset. With the hybrid approach implemented here, called Hybrid Particle Swarm Optimization and Fuzzy *c*-means clustering (Hybrid PSO-FCM), for network intrusion detection it is found that 92.70% attack is detected for the NSL-KDD 20 percent dataset as compared to 87.36% attack detected by Hybrid MI-FCM and maximum of 45.93% of attack as detected by FCM clustering algorithm. As observed from the experimental results the Hybrid PSO-FCM algorithm outperforms the MI-FCM clustering algorithm for attack detection over the considered dataset and maybe deployed in efficient detection of anomaly based network attacks.

REFERENCES

[1] Roger Storlokken (2007), "Labelling clusters in an anomaly based IDS by means of clustering quality indexes", Department of Computer Science and Media Technology,Gjovik University College

[2] M.Shivakumar, R.Subalakshmi, S. Shanthakumari and S.John Joseph (2013), "Architecture for Network-Intrusion Detection and Response in open Networks using Analyzer Mobile Agents", IJSRNSC, Vol.1, Issue 4, pp.3-7

[3] Raghunath ,B. R. and Mahadeo, S. N. (2008), "Network Intrusion Detection System (NIDS)", International Conference on Emerging Trends in Engineering and Technology", IEEE, 2008

[4] Benaicha, S. E., Saoudi, L., Guermeche, B., Eddine, S. and Lounis, O. (2014), "Intrusion detection system using genetic algorithm", Science and Information Conference (SAI), IEEE-2014, pp. 564–568

[5] Manmohan Dagar and Rashmi Popli (2018), "Honeypots: Virtual Network Intrusion Monitoring System", IJSRNSC, Vol.6, Issue 2, pp.45-49
[6] Zhao, Y. (2016), "Network intrusion detection system model based on data mining",17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), IEEE, Shanghai, China, pp. 155–160
[7] D Gupta, S Singhai, S Malik and A Singh (2016), "Network intrusion detection system using various data mining techniques", IEEE International Conference on Research Advances in Integrated Navigation Systems (RAINS)
[8] A.K. Siddique and T Farooqui,, (2017), "Improved Ensemble Technique based on Support Vector Machine and Neural Network for Intrusion Detection System", International Journal Online of Science, 3(11)
[9] Harish, B.S. and Kumar, S.A., (2017), "Anomaly based intrusion detection using modified fuzzy clustering", International Journal of Interactive Multimedia and Artificial Intelligence, 4(6), pp.54-59
[10] R.K. Das, M Panda, S Dash and S.S Dash (2018) "Application of Artificial Immune System Algorithms in Anomaly Detection", Progress in Computing, Analytics and Networking, Springer, Singapore, pp. 687-694
[11] A. Panigrahi and M.R. Patra (2018), "A Layered Approach to Network Intrusion Detection Using Rule Learning Classifiers with Nature-Inspired Feature Selection", In Progress in Computing, Analytics and Networking, Springer, Singapore, pp. 215-223
[12] R Sahani, C Rout, J.C. Badajena, A.K. Jena and H. Das (2018), "Classification of Intrusion Detection Using Data Mining Techniques", Progress in Computing, Analytics and Networking, Springer, Singapore, pp. 753-764
[13] A Ahmed, Dowlat Elngar, A. El, Mohamed, A. and Fayed, F. M. Ghaleb (2013), "A Real-Time Anomaly Network Intrusion Detection System with High Accuracy", Inf. Sci. Lett. 2, No. 2, pp.49-56
[14] Lan, Yuan-Dong (2017), "A Hybrid Feature Selection based on Mutual Information and Genetic Algorithm.", Indonesian Journal of Electrical Engineering and Computer Science 7, No. 1, pp. 214-225.
[15] Nojun Kwak, , Choi, Chong-Ho (2002), "Input Feature Selection by Mutual Information Based on Parzen Window", IEEE transactions on pattern analysis and machine intelligence, Vol. 24, no. 12
[16] Tavallaee, Mahbod, Bagheri, Ebrahim, Lu, Wei and Ali A. Gorbani (2009), "A Detailed Analysis of the KDD CUP 99 Dataset", In Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009, IEEE, pp. 1-6
[17] Revathi, S. and Malathi, A. (2013), "A Detailed Analysis on NSL-KDD Dataset Using Various Machine Learning Techniques for Intrusion Detection", IJERT, 2013, Vol. 2 Issue 12.
[18] P.S. Bhattacharjee, S. A. Begum, and Md, Fujail Abul Kashim (2017), "A Comparison of Intrusion Detection by K-Means and Fuzzy C-Means Clustering Algorithm over the NSL-KDD Dataset", IEEE-ICCIC 2017
[19] Cang, Shuang (2011), "A Mutual Information based Feature Selection Algorithm", 4th International Conference on Biomedical Engineering and Informatics (BMEI), IEEE, pp. 2241-2245
[20] Ren Wuling, Cao, Jinzhu and Wu, Xianjie (2009), "Application of Network Intrusion Detection Based on Fuzzy C-Means Clustering Algorithm", In Intelligent Information Technology Application IITA 2009, Third International Symposium , IEEE,Vol. 3, pp. 19-22

Authors Profile

Mr. Partha Sarathi Bhattacharjee received his Master Degree in Computer Application from IGNOU, New Delhi, India and M. Phil in Computer Science from Vinayaka Missions University, Salem, Tamilnadu, India. He is currently Research Scholar in the Department of Computer Science, Assam University, Silchar, Assam, India. .His research interests are network security, data mining, genetic algorithm and fuzzy logic.

Mr. Arif Iqbal Mozumder received Master in Computer Science from Assam University, Silchar. He is currently Research Scholar in the Department of Computer Science, Assam University, Silchar, Assam India .His research interests are in image processing, artificial neural networks, genetic algorithms and fuzzy logic.

Dr. (Mrs.) Shahin Ara Begum received her M.Sc. in Computer Science from Jamia Millia Islamia, New Delhi, India and Ph. D in Computer Science from Assam University, Silchar, Assam, India. She is working as an Associate Professor in the Department of Computer Science, Assam University, Silchar. Her research interests are in Machine Learning, Soft Computing Techniques, Pattern Recognition and Data Mining.

