ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 9, September 2015

Development of Hybrid Intrusion Detection System and Its Application to Medical Sensor Network

Karthik G[1], Geetha T[2], Nagappan A[3]

Research Scholar, Dept. of Computer Science Engineering, V.M.K.V Engineering College, Vinayaka Missions Research Foundation Deemed University, Salem, Tamil Nadu, India.

Assistant Professor, Dept. of Computer Science Engineering, V.M.K.V Engineering College, Vinayaka Missions Research Foundation Deemed University, Salem, Tamil Nadu, India.

Principal, V.M.K.V Engineering College, Vinayaka Missions Research Foundation Deemed University, Salem, Tamil Nadu, India.

ABSTRACT: Intrusion detection is the challenge to monitor and probably prevent the attempts to intrude into or otherwise compromise your system and network resources. One of the recent methods for identifying any abnormal activities staging in a computer system is carried out by Intrusion Detection Systems (IDS) and it forms a major portion of system defence against attacks.In the literature, various techniques for Intrusion Detection have been proposed in recent years. One of the methods proposed is Intrusion Detection System (IDS) based on Fuzzy Bisector- Kernel Fuzzy C-means clustering technique and Bayesian Neural Network. In the previous work, the dimensionality of the data played a major role in obtaining the better detection rate. In order to overcome the dimensionality issue, feature selection would be right choice to improve the detection rate without compromising the computation time. In the extended work, LDA+CS (Linear Discriminant Analysis + Cuckoo search) will be developed by combining LDA and CS. LDA is a commonly used technique for dimensionality reduction. Here, CS will be incorporated with the intention of assisting the ill-conditioning issue by selecting an “optimal” subset of features that result in an intermediate lower dimensional subspace. Then, feature reduced dataset is grouped into clusters with the use of Fuzzy Bisector- Kernel Fuzzy C-means clustering (FB-KFCM). In the classification step, the centroids from the clusters are taken for training the Bayesian Neural Network. For the online identification of intrusion detection node, test data is given to the trained network, which outputs if the data is intruded or not. The entire system will be applied to medical sensor network to find the intrusion behaviour by simulating the networks in JAVA. Finally, the performance of the system will be analysedusing KDD CUP 99 dataset in terms of accuracy.

KEYWORDS: Intrusion Detection System, Classification, Clustering, LDA-CS, Fuzzy Bisector- Kernel Fuzzy C-means clustering (FB-KFCM), Bayesian Neural Network, KDD CUP 99.

I. Introduction

An Intrusion Detection System (IDS) is software and/or hardware, which is designed for identifying the undesirable efforts for enhancing the computer security systems [9]. Especially, the wireless sensor devices has given rise to a wider range of amazing applications in various walks of our lifethat involve environment and habitat monitoring, healthcare applications and many more. But, simultaneously, the sensor nodes have produced the same number of threats caused by attackers, whose intention is to achieve access to the network and the data transferred inside it. Till now, numerous classical security methodologies exist for the purpose of avoiding these intrusions [10]. The intrusion detection systems fall into two important categories. One category is for analyzing the network traffic and the other is to analyze the operating system audit trails. These systems use either the rule-based misuse detection or anomaly detection naturally [13] and their power relies on the ability of the security personnel developing them to a larger extent. The first category is capable of identifying the known attack types alone. On the contrary, the second category is

DOI: 10.15680/IJIRCCE.2015. 0309041

subjected to the generation of false positive alarms. Therefore, several machine learning techniques have been applied for designing IDS.These machine learning techniques include neural networks, linear genetic programming, Support vector machines, Bayesian Networks, Multivariate adaptive regression splines and Fuzzy inference systems. [14]. Likewise, several data mining techniques have been developed as well to detect the key features or parameters that help in defining intrusions [12].

Hence, IDS should lower the quantity of data to be processed and this is more vital in case of real-time detection. Data filtering, data clustering and feature selection can achieve reduction of data. Clustering can be done to obtain the hidden patterns in the data and the essential features used for detection purposes. Better classification is possible with feature selection, which searches for the subset of features that excellently classifies the training data [15]. The classical cluster analysis works by assigning each datum to exactly one cluster. But, the fuzzy cluster analysis improves this requirement by using gradual memberships. This helps in dealing with the data that simultaneously belong to more than one cluster. The intrusion detection systems (IDS) extensively use the Clustering methodologies and in particular, the fuzzy approaches seem to be more efficient than the other clustering algorithms in use. Fuzzy C-Means clustering model (FCM)was initially introduced by Dunn in 1974 and it was extended and generalized by Bezdek in 1983 [11]. Generally, the techniques for dimensionality reduction concentrate either on choosing a suitable subset from the original set of I attributes or on mapping the initial I-dimensional data onto the K-dimensional space, where K<I[16].

Most of the recent feature extraction techniques involve linear transformations of the original pattern vectors to new vectors of lower dimensionality [5].The renowned technique for reducing the dimensionality is the Principal Component Analysis. But, problems arise in this method with the selection of number of directions. It does not perform the computation of principal component in high dimensional feature spaces that have relation to input space by some nonlinear map [18]. Linear Discriminant analysis feature reduction technique is the new scheme employed in the field of cyber attack detection. This method reduces the number of input features in addition to improving the classification accuracy. Moreover, the training and the testing time of the classifiers can be decreased by this method through the selection of most discriminating features [5].

The way the optimal set of features is selected forms the major problem encountered by most of the researchers. This isbecause; all the features are not related to the learning algorithm. In some situations, irrelevant and redundant features can produce noisy data that can distract the learning algorithm and degrade the detector accuracy, leading to time consuming training and testing processes. Feature selection was proved to have a considerableeffecton the performance of the classifiers [17]. A feed-forward neural network classically trained using back-propagation can be regarded as an effective classifier of theactionsproduced bythehead of severely disabled people [24], [25], [26]. Yet, there are demerits with the standard neural networks because it offers poor generalisation ability when provided with limited training data. Bayesian techniques have been applied to neural networks in the recent yearsfor enhancing the accuracy and robustness of neural network classifiers [19]. In our previous research [23], it has been shown that the Bayesian neural network is capable of classifying the head movement commands consistently even with limited training data.

However, various researchers have found the dimensional reduction of huge data input and optimal feature selection as the major problems in intrusion detection systems. Similarly in the existing work, it plays a major role that reduces the accuracy of the system. So, wehaveproposed a new method called LDA+CS (Linear Discriminant Analysis + Cuckoo search) to overcome this. The proposed technique is implemented using JAVA PROGRAMMING, employing KDD CUP 99 dataset. The evaluation metric utilized is the accuracy and the comparative analysis is made against other techniques such as FCM+ Bayesian network and FB-KFCM+ Bayesian network.

The rest of the paper is organized as follows: A brief review of researches related to the proposed technique is presented in section 2. The proposed intrusion detection technique is presented in Section 3. The detailed experimental results and discussions are given in Section 4. The conclusions are summed up in Section 5.

II. LITERATURE SURVEY

Literature review presents several techniques based on intrusion detection system. It has received a lot of interest among the researchers because it is widely applied for preserving the security within a network. Here we present some of the technique for intrusion detection system. Anomaly Intrusion Detection System (IDS) have various drawbacks like complex computation and inefficiency in real time detection. So, in order to reduce the computational complexity, Zhiyuan Tan *et al* [1]have designed a method called Linear Discriminant Analysis (LDA). They have used the difference distance map for selecting the significant features. Here, the high-dimensional feature vectors were transformed into a low-dimensional domain by the designed method initially. Then, based on the Euclidean distance on the simple, low dimensional feature domain, they have identified the similarity between the new incoming packets and a normal profile. The experimental results were based on the pre-calculated threshold, which differentiates normal and abnormal network packets. DARPA 1999 IDS dataset was used here to evaluate their proposed method.

But the conventional Linear Discriminant Analysis (LDA) feature reduction technique has drawbacks that were not suitable for non-linear dataset. In general, the huge sized network traffic data used in intrusion detection system have ineffective information that affects the system accuracy. So in order to overcome this drawback, Shailendra Singh and Sanjay Silakari [2] have designed an efficient feature reduction method called Generalized Discriminant Analysis (GDA). The number of input features was reduced by this method. Also, the classification accuracy was increased and the time required for classifier in training and testing was reduced by selecting the most discriminating features. The performance of their designed method was evaluated by Artificial Neural Network (ANN) and C4.5 classifiers. The experimental results have shown that the accuracy of their designed method was improved.

The previously used k-means clustering algorithm in intrusion detection system have various drawbacks such as computation complexity and the selection of initial central point affects the algorithmic results. So, Li Tian and Wang Jianwen [3] have designed an improved k-means clustering algorithm, which introduced the optimized dynamic central point cyclic method. The improved clustering method applied in the intrusion detection system has enhanced the fault detection rate of abnormal detection and has reduced the false drop rate effectively as well. Finally, the algorithm was evaluated by KDD cup 99 dataset to show that the accuracy of data classification and the detection efficiency has increased significantly. Also, the experimental results have revealed that the designed algorithm has achieved the desired objectives with a higher detection rate and higher efficiency.

The different issues found in the intrusion detection system were regular updating, lower detection, capability to unknown attacks, non-adapting high false alarms rate, high resources consumption and many others. However, because of the importance of soft computing in the intrusion detection system, Hafiz Muhammad Imran *et al* [4] have introduced an efficient soft computing method to select the optimum subset of features. Here, to get better results, they have provided a hybrid method called LDA + GA for feature transformation and selection. LDA was chosen here as a feature reduction method because it outperformed PCA. Also, the dataset used here for training and testing was standard NSL-KDD dataset. Further, to classify the network traffic into normal or intrusive activities, they have used an outstanding classification method called RBF. The experimental results of our designed method have shown that the selection of optimal subset of features has reduced the time consumption rate and increased the accuracy ratio as well.

The existing intrusion detection system makes use of the entire irrelevant features. Hence, to produce an effective and efficient classification process, a well-defined feature extraction algorithm was essential. RupaliDatti and Bhupendraverma [5] have suggested an efficient feature extraction method called as Linear Discriminant Analysis (LDA) for intrusion detection system. The back propagation algorithm was employed to perform the classification process. This method aims to identify the significant input features that are computationally efficient and effective in constructing IDS. It is apparent from their experimental results that the proposed model has offered improved and robust representation of data. This is because, it has achieved 97% of data reduction and about 94% of training time reduction. In addition, the accuracy achieved in identifying the new attacks is found to be more or less thesame.Thenumber of computer resources as well as both the memory and the CPU timespent on detecting an attack was also decreased. The experimental results have shown that their method was reliable for detecting intrusion.

To deal with the multiclass problem in intrusion detection system, Snehal A. Mulay*et al* [6] have designed a decision-tree-based support vector machinethatuses support vector machines and decision tree in a combined fashion. The non-time consuming training and testing processes may be viewed as the benefits of this method, which in turn increases the system efficiency. At first, the dataset was split into two subsets from root to the leaf until every subset contains only one class. This had a larger impact on the classification performance of their system. Though the final results for the designed method was not presented, it can be known that the multiclass pattern recognition problems can be solved using the tree structured binary SVMs and the resultant intrusion detection system could be of more speed than the other methods.

Shingo Mabu*et al* [7] have developed a GNP-based fuzzy class-association-rule mining with sub attribute utilization and classifiers that rely on the extracted rules. It is capable of consistently utilizing and combining the discrete and continuous attributes in a rule and can efficiently extract severalsuperior rules for classification. As an application, intrusion-detection classifiers for both misuse detection and anomaly detection have been developed and their effectiveness was proved using KDD99Cup and DARPA98 data. The experimental results of misuse detection depict that the designed method offers high DR and low PFR that serve as the two important criteria for security systems.

Gang Wang et al, [8] have proposed an intrusion detection method called as FC-ANN that depends on ANN and fuzzy clustering. The fuzzy clustering technique was employed to partition the heterogeneous training set into numerous homogenous subsets. In such a way, the complexity of each of the sub training set was decreased and as a result, the detection performance was increased. The experimental results using the KDD CUP 1999dataset demonstrates the effectiveness of their method, in particular, for low-frequent attacks like R2L and U2R attacks in terms of detection precision and detection stability.

III.PROPOSED INTRUSION DETECTION SYSTEM USING LOCALITY PRESERVING CUCKOO SEARCH ALGORITHMS AND BAYESIAN NEURAL NETWORK

Intrusion detection system is a device used to identify whether the input data is intruded or not. The process is done by classifying the huge amount of input data in to different groups or classes by clustering. In the previous work, we have designed an Intrusion Detection System (IDS) based on Fuzzy Bisector- Kernel Fuzzy C-means clustering technique and Bayesian Neural Network. The method was implemented using JAVA PROGRAMMING with KDD CUP 99 dataset. In this existing work, the dimensionality of the data in clustering stage was a major drawback that reduces the detection rate of the IDS. So, in order to overcome the problem based on the dimensionality of the data, we design a hybrid intrusion detection system called LDA-CS (Linear Discriminant Analysis-Cuckoo Search) here. In our proposed hybrid intrusion detection system, the input dataset consists of large number of data with various attacks. So, classifying this huge dataset is difficult and time consuming and there is also a possibility of increasing the error rate. The different attacks found in our datasets are, DOS (Denial of Service attack), R2L (Remote to Local (User) attack), U2R (User to Root attack) and Probing Surveillance. To overcome the drawbacks of the previously done works, we have introduced a new method called LDA-CS here, which will improve the detection rate of our intrusion detection system. Our proposed method consists of two phases, namely, the training phase and the testing phase. For training and testing, we have used the KDD cup 99 dataset in our method. The general architecture of our proposed method is shown in Fig.1.

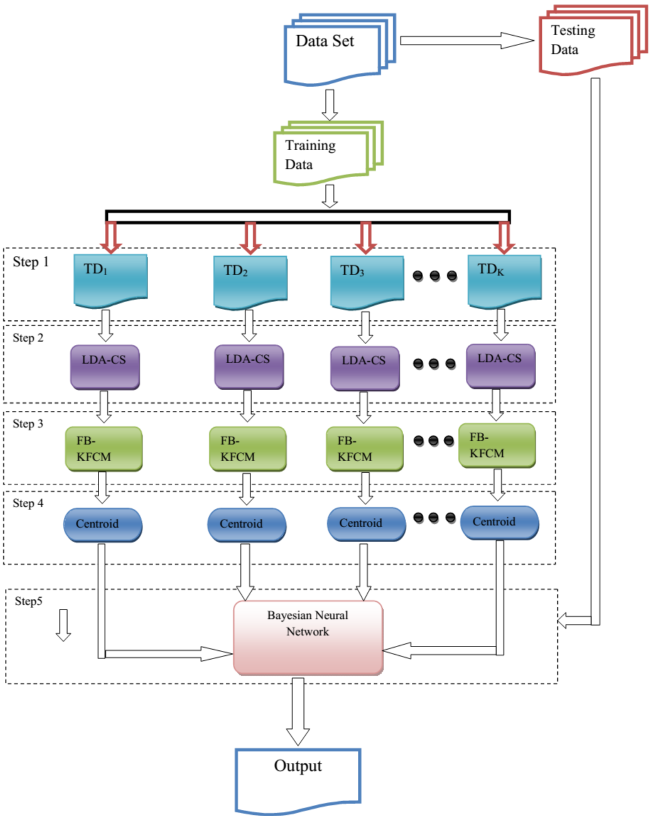


Fig.1. Proposed Intrusion Detection System

A. *Training Phase:*

The training phase consists of various processing stages such as the input data set is clustered and classified using various techniques like LDA-CS, FB-KFCM and Bayesian Neural Network. Here, we have used the KDD cup 99 dataset that is huge in size. It consists of approximately 4,900,000 single connection vectors, each of which contains 41 features. In general, the classifier delivers more accurate results only while using complete linear feature space. But, the direct application of this dataset to the classifier has various drawbacks such like the classifier becomes biased due to architecture complexity and training as well as testing efficiency decreases. It also results in increasing memory consumption rate and computational cost. In order to overcome these problems, it is best to adopt some approachesfor selecting the optimal subset of features from a linear space of features. Hence, Cuckoo Search algorithm that is commonly called as LDA-CSisapplied in this work to select the optimal subset of linear feature space. The LDA-CS consists of four stages of processing as follows:

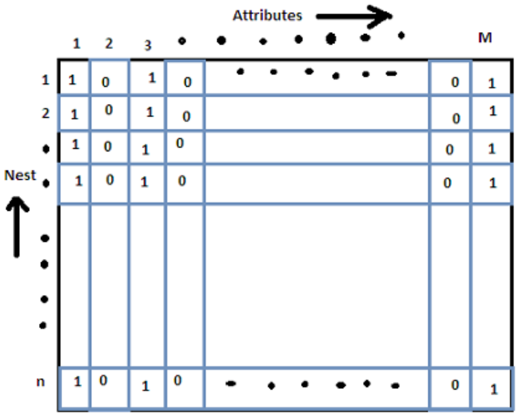
(i) Initialization

(ii) Fitness calculation and Nest update

(i) Initialization

In the cuckoo search algorithm, a fixed host nest is built at a size of $n \times M$. Here n is the number of nest and M is the number of attributes. The fixed host nest is an index to select the relevant features from the original dataset. Here, the class for each nest is not defined in the fixed host nest. So, in order to determine the class for each host nest based feature, we have used a classifier called LDA here. It is used to identify whether the host nest based data is intruded or not. The fixed nest built is shown in Fig.2.

DOI: 10.15680/IJIRCCE.2015. 0309041



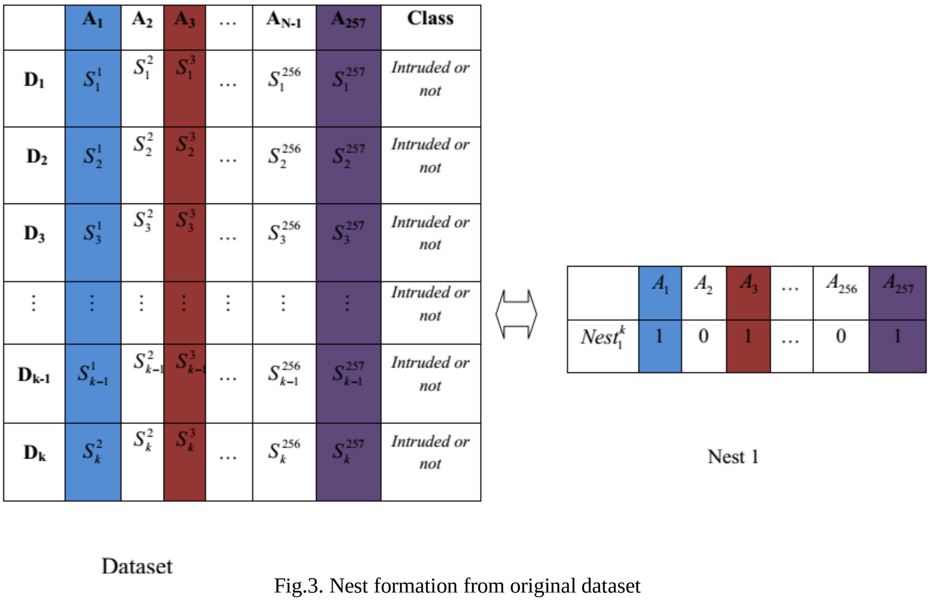
	1	2	3	•	• • • • • •		M
1	1	0	1	0	• • • • • •	0	1
2	1	0	1	0		0	1
•	1	0	1	0		0	1
Nest •	1	0	1	0		0	1
• • • •							
n	1	0	1	0	• • • • • •	0	1

Fig.2 Fixed Nest

Further initialization process is done based on the fixed host nest.

(ii) Fitness Calculation and Nest update

In this stage, the fixed host nest built is randomly assigned with a probability of 1s and 0s. After this selection, both the 1s and 0s based relevant features from the original dataset will increase the computational complexity. So, based on the cuckoo search algorithm the nest with 1s is selected and the random nest with 0 is neglected. This results in dimensionless feature subset based on the neglected nest with 0s. The selected subset contains features that are relevant to the selected host nest with a size of $n_1 \times m$ where, m is the dimension of the reduced subset. Finally, a dimension reduced N number of training feature subset is obtained from an original set (N) where, the size is $N < M$. The general data set and the host nest obtained is shown in Fig.3.


Dataset
Nest 1

	A$_1$	A$_2$	A$_3$	…	A$_{N-1}$	A$_{257}$	Class
D$_1$	S_1^1	S_1^2	S_1^3	…	S_1^{256}	S_1^{257}	*Intruded or not*
D$_2$	S_2^1	S_2^2	S_2^3	…	S_2^{256}	S_2^{257}	*Intruded or not*
D$_3$	S_3^1	S_3^2	S_3^3	…	S_3^{256}	S_3^{257}	*Intruded or not*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	*Intruded or not*
D$_{k-1}$	S_{k-1}^1	S_{k-1}^2	S_{k-1}^3	…	S_{k-1}^{256}	S_{k-1}^{257}	*Intruded or not*
D$_k$	S_k^2	S_k^2	S_k^3	…	S_k^{256}	S_k^{257}	*Intruded or not*

	A_1	A_2	A_3	…	A_{256}	A_{257}
$Nest_1^k$	1	0	1	…	0	1

Fig.3. Nest formation from original dataset

Here, the dimension reduced subset contain only valuable information and has some data about some of the other features. Furthermore, the subset with relevant feature is given to LDA for classification. The LDA has various stages of processing, which doesn't change the location but only tries to provide more class separately and draws a decision region between the given classes. The input to LDA is N dimensional training subset that belongs to different class v with N_i samples in the i^{th} class. The first stage of the LDA is to group the subset of data into two different classes, which are attack or not. For each subset, the within-class distance and the between-class distance is computed for two different classes. For N number of training datasets, the mean vector and the covariance matrix is calculated for each class of the complete data set. It is given as in the Eq (1) below.

$$N = \sum_{i=1}^{n} N_i \qquad \text{Eq(1)}$$

Where, N represent the total number of training subset were, N_i represents the number of training datasets in class i. Naturally, the number of classes is i. The scatter matrix is calculated by eigen decomposition that is applicable to high dimensional data. The within class and between class scatter matrix calculated is represented as W_C and B_C. The scattering matrix are represented by the Eq (2) below.

The Between class scatter matrix B_C is represented as:

$$B_C = \sum_{i=1}^{n} N_i / N(v_i - v)(v_i - v)^T \qquad \text{Eq (2)}$$

The within class matrix W_C is represented by the Eq (3):

$$W_C = \sum_{i=1}^{n} \sum_{j=1}^{N_i} 1/N(z_j^{(i)} - v_i)(z_j^{(i)} - v)^T \qquad \text{Eq (3)}$$

Here, the mean for the i^{th} class,V_i is represented by the Eq (4):

$$v_i = 1/N_i \sum_{j=1}^{N_i} z_j^{(i)} \qquad \text{Eq (4)}$$

Similarly, the total mean of the class for the whole dataset is represented by the Eq (5) given below.

$$v = 1/N \sum_{i=1}^{n} \sum_{j=1}^{N_i} z_j^{(i)} \qquad \text{Eq (5)}$$

Finally, a discriminant function is determined based on the following by the Eq (6).

$$Y_{LDA} = tr[(w_C)^{-1} B_C] \qquad \text{Eq (6)}$$

Fitness for each training subset is obtained based on the LDA classifier. A dimension reduced subset of feature is obtained and applied to LDA in single iteration. Likewise, the process is repeated until the global best solution is obtained. Here, N number of training subset is given as input to the LDA classifier and n number of fitness is obtained for each subset. The N number of fitness functions determined for the fixed host nest is, $f = f_1, f_2, f_3, \ldots . f_N$ .Among this, the best fitness is found and replaced as X $_{best}$. Finally, the accuracy of our system is determined based on the ratio of the total number of correct predictions to the actual data set size. The fitness function f is calculated by the Eq (7) given below.

$$fitness = 1 - Accuracy \qquad \text{Eq (7)}$$

In order to generate a new solution, Levy Flight is performed that provides a random walk. The new solution $y_{(t+1)}$ is determined based on the Eq (8) given below, but maintain the current best.

$$y_{(t+1)} = y_{(t)} + \alpha \oplus Levy(\lambda) \qquad \text{Eq (8)}$$

Where, $\alpha > 0$ is the step size, But in most cases, we use $\alpha = 1$ .This has an infinite variance with an infinite mean. Here, the consecutive steps of a cuckooessentially form a random walk process that obeys a power-law step-length distribution with a heavy tail. In addition, a fraction of the worst nests can be abandoned, so that the new nests can be built at new locations by random walks and mixing. The mixing of the solutions can be performed by random permutation according to the similarity/difference to the host eggs. The flow diagram of the designed LDA-CS is shown in Fig.4.

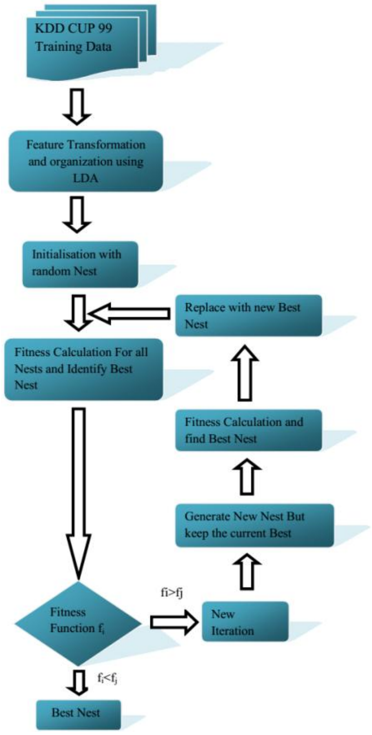

KDD CUP 99 Training Data
Feature Transformation and organization using LDA
Initialisation with random Nest
Replace with new Best Nest
Fitness Calculation For all Nests and Identify Best Nest
Fitness Calculation and find Best Nest
Generate New Nest But keep the current Best
fi>fj
Fitness Function f_i
New Iteration
f_i<f_j
Best Nest

Fig.4. LDA-CS Flow Diagram

The optimal dimensionality reduced features are further clustered using a technique called FB-KFCM (Fuzzy Bisector-Kernel Fuzzy C-Mean).

B. *Clustering using FB-KFCM*

Clustering is one of the common methods used to group the optimal features obtained from LDA-CS. KFCM is one of the clustering methods used previously. But, it is not suitable for large datasets. So, we have proposed a new method for effective clustering by incorporating Fuzzy Bisector with fuzzy C-means clustering called as FB-KFCM here. In order to obtain better results, here we have used a modified technique by incorporating Fuzzy Bisector called as FB-KFCM. The general operation of the newly incorporated fuzzy bisector is based on the optimal features and Minimum Squared Error (MSE) parameters. The initial stage of the fuzzy bisector is that it initially selects a cluster based on the above parameters and is divided in to two using fuzzy c-means technique. The process has several stages and each contains single bisection, which increases the number of clusters by one. The input dataset to the FB-KFCM algorithm is represented as: $X=\{x_1,x_2,...,x_d\}$ where, d is the size or dimension of the dataset. Further, the input dataset is clustered and grouped into n number of clusters as represented by Eq (9) below.

$$Q=\{C_1,C_2,...,C_N\} \qquad \text{Eq (9)}$$

Here, each grouped cluster has data x_i belonging to Q_i . Also, the data inside the i^th cluster C_i is represented as: $C_i=\{D_1,D_2,...,D_k\}$ where, k is the number of data in the i^th cluster. Each cluster has a group of n number of data. The proposed FB-KFCM is shown in Fig.5.

DOI: 10.15680/IJIRCCE.2015. 0309041

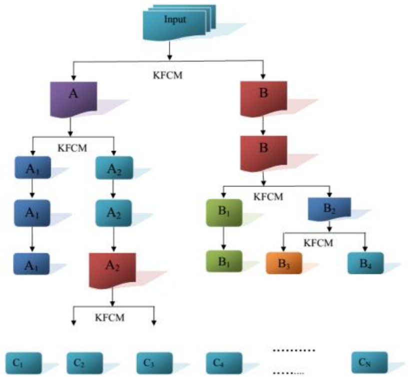


Fig.5 FB-KFCM

The FB-KFCM clustering includes N+ 1 stages and in each stage, the input data is divided into two clusters by KFCM algorithm. For each input data X, two clusters are formed and are further divided in to two clusters such as A and B. In the next stage, one among the two clusters is taken anditis divided into two based on the KFCM. The total number of clusters in this stage is three and likewise, there are m number of clustering stages and are grouped in to n clusters as denoted by the followingEq (10).

$$Q=\{C_1,C_2,...,C_N\} \qquad \text{Eq (10)}$$

Then, for each grouped cluster, the Mean Square Error is detected based on the Euclidian distance between the data points and centroid. The MSE of the ith cluster is represented by the Eq (11) given below.

$$MSE_i=\frac{1}{Ni}\sum_{k=1}^{Ni}\| C_k - ci \|^2 \qquad \text{Eq (11)}$$

Finally, for N number of clustering stages, the data points in the clusters is represented as $D_1,D_2,...,D_K$ and MSE of the clusters are represented as $E_1,E_2,...,E_K$. Each stage of the process is carried out by the KFCM, which has totally N+ 1 stages. Hereafter, the centroid for each cluster is to be calculated for further process. The centroid based classification has various advantages such as less time consumption and reduced complexity. The centroid of the ithcluster is calculated by the Eq (12) given below.

$$W_i = \sum_j D_j \Big/ K \qquad \text{Eq (12)}$$

Based on the above equation, the centroid for each cluster is calculated and given to the classification process.

C. Classification using Bayesian Neural Network:

Classification in intrusion detection is to train the centroid based grouped data obtained from FB-KFCM. The centroid of each cluster is trained by classifier to identify whether the input data is intruded or not. In our proposed system, we have used Bayesian Neural Network for better classification. Bayesian neural network is the improved version of artificial neural network to obtain robust classification result. In Bayesian Neural Network Classifier (BNNC), the weight decay parameter can be adjusted automatically to obtain the optimal solution during training. The whole data can be used for training without any need of separate validation. The centroid value obtained from each cluster of the input data is given to the BNNC for training. Let the centroid input to the Bayesian classifier be represented by the Eq (13) below.

$$W_i;0<i\le(N+1) \qquad \text{Eq (13)}$$

The general neural network contains three layers, namely, the input layer, the hidden layer and the output layer. Initially, the centroid obtained from each cluster is given as input to the Bayesian neural network to select the prior probability distribution for model parameters. Second is the fact that the prediction are made with respect to the posterior parameter distribution obtained by updation of the prior function. The Bayesian neural network is formed based on the above two properties. Let the input be the vector of real centroid value W_i. The output of each input centroid is trained by varying the weight at each node to obtain the best classification result. The architecture of the Bayesian neural network is shown in Fig. 6.

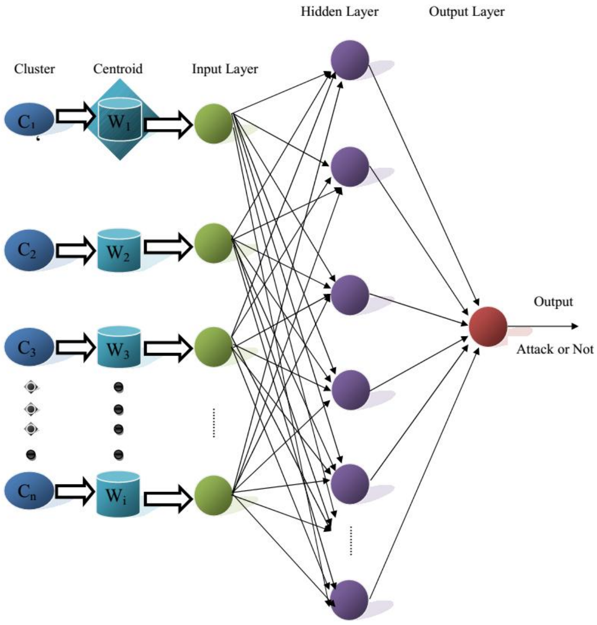


Fig.6 Bayesian Neural Network Classifier(BNNC)

The output for the single hidden layer based Bayesian neural network is computed based on the Eq (14) given below.

$$y_k(x) = V_0((b_k + \sum_{i=1}^{M} W_{ki} P(x))) \qquad \text{Eq (14)}$$

$$Where,\ P(x) = \tan b_j + \sum_{j=1}^{d} W_{ij} x_j \qquad \text{Eq (15)}$$

Here, W_{ij} is the weight on the connection from the input unit j to the hidden unit i. Similarly, W_{ki} is the weight on the connection from hidden unit i to output unit k. Also, the biases of the hidden and output unit are b_k and b_j. The activation function of the output layer is V_0. Further, to avoid larger weights, a weight decay function is added to the data error function e_D. Particularly, for classification problem, In Eq (16):

$$T_e = e_D + \sum_{h=1}^{H} J_h e_{Wh} \qquad \text{Eq (16)}$$

Where, T_e is the total error function, J_h is a non-negative parameter for the distribution of other parameters such as weights and biases. Here, e_{Wh} is the weight error for the hth group of weights and biases and H is the number of groups of weights and biases in the neural network. Hereafter, the parameters such as weights and biases are grouped into a single W dimensional weight vector w. According to the given weight w, the posterior distribution of the given data D is represented by the Eq (17) below.

$$P(w/D,\mu)=\frac{P(D/w,\mu)P(w/\mu)}{P(D/\mu)} \qquad \text{Eq (17)}$$

Where, $\mu=\{J_1,J_2,.....J_H\}$.

Also, the prior distribution of the weight is represented by the Eq (18) :

$$P(w/\mu)=\frac{1}{Z_W(\mu)}\exp\left(-\sum_{h=1}^{H}J_h e_{Wh}\right) \qquad \text{Eq (18)}$$

Where, $Z_W(\mu)=\prod_{h=1}^{H}\left(\frac{2\pi}{J_H}\right)^{W_{h/2}}$

The posterior density for the parameters is proportional to the product of prior and hence, the training process is carried out for all clustering centroid $W_i;0<i\le(N+1)$. After training, the test data is given to our Bayesians trained neural network to determine whether the output data is attacked or not.

IV. Experimental Result

The performance of our proposed method is analyzed using KDD CUP 99 dataset. It is one of the most widely used dataset in intrusion detection systems. In section 4.1, data set description and experimental setup are given. In section 4.2, details about the evaluation metric employed are given. In section 4.3, comparative analysis is provided. Finally, in section 4.4, the application in medical sensor network is given.

A. *Experimental Setup and Data set Description:*

The proposed technique is implemented using JAVA PROGRAMMING on a system having 8GB RAM and 3.2 MHz processor. To evaluate the performance of the proposed technique, we have used KDD CUP 99 DATASET for testing and evaluation. The KDD CUP 99 dataset used here is a version of the original 1998 DARPA intrusion detection evaluation program. Also, it is one of the publicly available data set that has actual attacks [20]. So, we have used the dataset here to design and evaluate our intrusion detection system.

The KDD CUP 1999 dataset used here was obtained from raw TCP dump data for a length of nine weeks. The dataset is made of large number of network traffic activities that include both normal and malicious connections, which has five million connection records as training data and two million as test data. Each instance has 41 features which are marked as normal or an attack. Totally 38 different attacks are found in both training and testing data, which falls into four main categories such as PROBE, denial of service (DOS), remote to local(R2L) and user to root(U2R). [21].

The KDD Cup99 dataset are available in three different files such as KDD Full Dataset that contains 4898431 instances, KDD Cup 10% dataset that contains 494021 instances and KDD Corrected dataset that contains 311029 instances. In table 1, the details about the KDD full and KDD 10% dataset are given. Table 1 explores the number of samples present in each category before and after the reduction of duplicate samples with percentage of reduction. Similarly, Table 2 contains detail information on KDD Corrected and GureKDD dataset along with before and after the reduction of redundancy samples with percentage of reduction. The reduction of duplicate samples is based upon algorithm 1. The Table 2 elaborates the forth mentioned four attack category on KDD Cup 3 different datasets with number of samples in each category and percentage of reduction after applying algorithm 1. Each sample of the dataset represents a connection between two network hosts according to network protocols. It is described by 41 attributes, out of which 38 are continuous or discrete numerical attributes and 3 are categorical attributes. Each sample is labelled as either normal or one specific attack. The dataset contains 23 class labels, out of which 1 is normal and remaining 22 are different attacks. The total 22 attacks fall into four categories as forth-mentioned attacks [22].

Table.1.The details categories of attack categories in "KDD FULL & 10%" Dataset.

Category	KDD Cup 99FULL Dataset	After Removing Duplicate Samples	% rate of Reduction	KDDCup99 10% Dataset	After Removing Duplicate Samples	%rate of Reduction	Dataset Class
Normal	972781	812814	16.44	97278	87832	9.71	NORMAL
Back	2203	968	56.06	2203	968	54.88	DOS
Pod	264	206	21.97	264	206	21.97	DOS
Land	21	19	9.52	21	19	9.52	DOS
Smurf	2807886	3007	99.89	280790	641	99.77	DOS
Teardrop	979	918	6.23	979	918	6.23	DOS
Neptune	1072017	242149	77.41	107201	51820	51.66	DOS
Nmap	2316	1554	32.90	231	158	31.60	PROBE
Satan	15892	5019	68.42	1589	906	42.86	PROBE
Ipsweep	12481	3723	70.17	1247	651	47.79	PROBE
Portsweep	10413	3564	65.77	1040	416	60.00	PROBE
Phf	4	4	0.00	4	4	0.00	R2L
Guess_pwd	53	53	0.00	53	53	0.00	R2L
Ftp_write	8	8	0.00	8	8	0.00	R2L
Imap	12	12	0.00	12	12	0.00	R2L
Spy	2	2	0.00	2	2	0.00	R2L
Multihop	7	7	0.00	7	7	0.00	R2L
Warezclient	1020	893	12.45	1020	893	0.00	R2L
Warezmaster	20	20	0.00	20	20	0.00	R2L
Buffer_Overflow	30	30	0.00	30	30	0.00	U2R
Loadmodule	9	9	0.00	9	9	0.00	U2R
Perl	3	3	0.00	3	3	0.00	U2R
Rootkit	10	10	0.00	10	10	0.00	U2R
Total	48,98,431	10,74,992	78.05%	4,94,021	145586	70.53%	

Table.2. Attack Distribution in KDDfull, KDD 10% and KDD Corrected dataset.

Dataset	DoS	U2R	R2L	Probe	Normal	Total
KDD Full	3883370	52	1126	41102	972781	4898431
KDD Full After removing duplicate samples	247267	52	999	13860	812814	1074992
KDD 10%	391458	52	1126	4107	97278	494021
KDD10% After removing duplicate samples	54598	52	999	2133	87832	145586
KDD Corrected	229269	70	16172	4925	60593	311029
KDD Corrected after removing Duplicate samples	22984	70	2898	3426	47913	77291

The KDD cup 99 dataset is huge in size, which offers difficulty in performing the experiment. So, we have used a subset of 10% of KDD cup 99 dataset for our experiment.

B. Evaluation Metric:

The evaluation metrics used in our proposed method are true positive (TP), true negative (TN), false positive (FP) and false negative (FN). Here, true positive indicates the number of correctly classified attack. A true positive is a sign of properly detecting the occurrences of attacks in intrusion detection system. True negative indicates the number of valid records that are correctly classified. A true negative specifies that the IDS have not made a mistake in detecting a normal condition. False positive indicates the records that were incorrectly classified as attacks, whereas in fact they are valid activities. A false positive specifies the wrong detection of a particular attack by IDS. A false positive is often produced due to lost recognition conditions and it represents the accuracy of the detection system. False negative indicates the records that were incorrectly classified as valid activities, whereas in fact they are attacks. A false negative stipulates that the IDS is unable to detect the intrusion after a particular attack has occurred. Based on TP, TN, FP and FN, the performance of our intrusion detection system is evaluated by: a) Accuracy b) Detection Rate (DR), c) Failure Analysis Rate (FAR). The accuracy of our system is obtained by the following expression by the Eq (19).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad \text{Eq (19)}$$

Then, the Detection Rate (DR) is determined based on the expression by Eq (20) given below.

$$DetectionRate(DR) = \frac{TP}{TP + FP} \qquad \text{Eq (20)}$$

Detection rate shows the probability of abnormal data in the test samples in detection. The higher Detection Rate indicates that the algorithm can more accurately reflect the input data anomalies in Eq (21).

$$Failure\ analysis\, rate(FAR) = \frac{FP}{FP + TN} \qquad \text{Eq (21)}$$

Failure analysis rate shows the accuracy of intrusion detection. Lower FAR indicates that the accuracy of detection is high.

C. Comparative Analysis:

In this section, our proposed method is compared with the existing techniques such as FCM+ Bayesian network technique and FB-KFCM+ Bayesian network technique in intrusion detection system. Here, the accuracy of each technique is taken for different cluster size such as 200, 180, 160 and 140. Also, for each cluster size, the accuracy of our proposed LDA-CS+ FB-KFCM+ Bayesian Network is taken. Similarly, the accuracy of the FCM+ Bayesian network technique and the previously used FB-KFCM+ Bayesian network technique are also obtained, which is shown in Table 3. The accuracy is taken based on the sample at a ratio of 9:1, which means that nine samples out of ten are taken for training and the rest 3 is used up for testing purposes.

Table.3.The accuracy is taken based on the sample at a ratio of 9:1

Cluster Size	Accuracy		
	FCM+Bayesian network	FB-KFCM+ Bayesian network	Proposed LDA-CS+FB-KFCM+ Bayesian Network
200	86.7711	93.0023	97.437
180	86.7378	93.4022	97.4373
160	86.7452	91.936	97.4373
140	86.7378	92.6015	97.4373

From Table 3, we have observed that the accuracy of our proposed method for different cluster sizes at 9:1 sample size is higher than the existing method. The accuracy plot of 9:1 for our proposed method is shown in Fig.7.

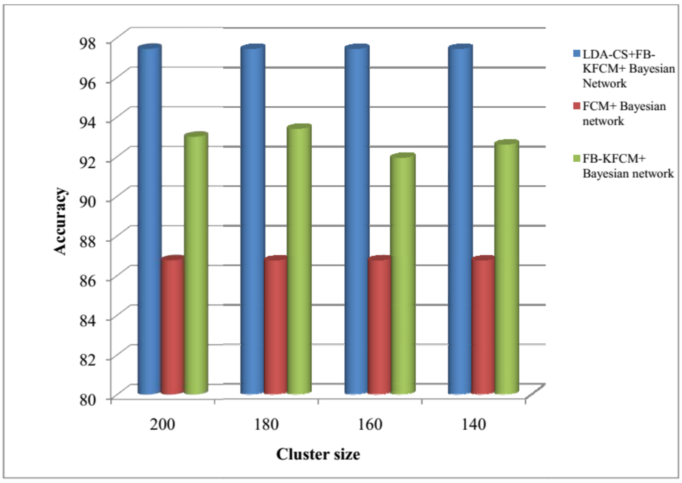
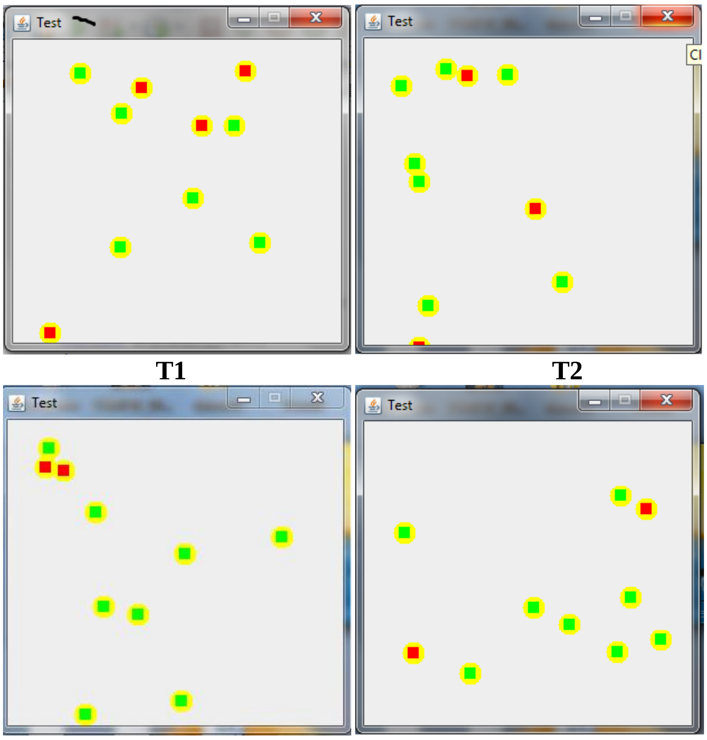


Fig 7: Accuracy plot for 9:1

D. *Application in medical sensor network* :

Our proposed intrusion detection system is applied to medical sensor network in order to detect which of the data are intruded and not intruded. Initially, we have simulated our algorithm using medical sensor networks that consists of totally 8668 data. The whole data of the medical sensor network is trained using Bayesian neural network in our algorithm. After training process, we have used 10 data for testing at each time. In this testing stage, our algorithm will detect which of the data were intruded and not intruded among the 10 data. Here, at time T1 we have used 10 nodes for testing and the simulation result obtained using our method is shown in Fig. 8. In the simulated result, two colours such as red and green was obtained that indicate the data type. The red colour in the result indicates the intruded data and the green colour in the result indicate the not intruded data. Among the 10 data given for testing time T1, 6 are not intruded data indicated by green and the remaining four are intruded data indicated by red colour. Similarly in time T2, another 10 data is given for testing in our algorithm and the simulation result is obtained. From the simulation result obtained at time T2, we have found that 7 of the data among 10 are not intruded and the remaining 3 are intruded.



T1 T2

T3 T4

Fig.8. Simulation Result Obtained for time T1, T2, T3 and T4

Again at time T3, the simulation result is obtained for 10 test data. From the result, we have found that 8 among 10 are not intruded and remaining 2 are intruded. Further at time T4, the 8 data among 10 are not intruded and remaining 2 are intruded, while testing 10 data in our proposed intrusion detection algorithm.

V. Conclusion

In our intrusion detection system, an LDA+CS (Linear Discriminant Analysis + Cuckoo search) is developed by combining LDA and CS. The LDA-CS technique is used in this work for dimensionality reduction and optimal feature selection. Further, the feature reduced dataset is grouped into clusters using Fuzzy Bisector- Kernel Fuzzy C-means clustering (FB-KFCM) method. Then, in the classification step, the centroids from the clusters were taken and trained using the Bayesian Neural Network. For the online identification of intrusion detection node, test data is given to the trained network and tested for obtaining which of the given data is intruded or not. The entire system is applied to medical sensor network to find the intrusion behaviour by simulating the networks in JAVA using KDD CUP 99 dataset. The evaluation metric utilized is the accuracy and the comparative analysis is made against the other techniques. Average accuracy value was found to be 97.43, which wasbetter than the other compared techniques. The high accuracy value shows the efficiency of the proposed technique.

References

1. Zhiyuan Tan, ArunaJamdagni, Xiangjian He and Priyadarsi Nanda, "Network Intrusion Detection Based on LDA for Payload Feature Selection",IEEE GLOBECOM Workshop on Web and Pervasive Security, pp. 1545-1549, 2010.
2. Shailendra Singh and Sanjay Silakari, "Generalized Discriminant Analysis algorithm for feature reduction in Cyber Attack Detection System", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 1, 2009.
3. Li Tian and Wang Jianwen,"Research on Network Intrusion Detection System Based on Improved K-means Clustering Algorithm", International Forum on Computer Science-Technology and Applications, pp.76 – 79, 2009.
4. Hafiz Muhammad Imran, Azween Bin Abdullah, Muhammad Hussain, SellappanPalaniappan and IftikharAhmad,"Intrusions Detection based on Optimum Features Subset and Efficient Dataset Selection", International Journal of Engineering and Innovative Technology (IJEIT), Vol.2, No. 6, 2012.

5. RupaliDatti and Bhupendraverma,"Feature Reduction for Intrusion Detection Using Linear Discriminant Analysis", International Journal on Computer Science and Engineering, Vol. 02, No. 04,pp.1072-1078, 2010.
6. Snehal A. Mulay, P.R. Devale and G.V. Garje,"Intrusion Detection System using Support Vector Machine and Decision Tree",International Journal of Computer Applications, Vol.3, No.3, pp. 0975 – 8887, 2010.
7. Shingo Mabu, Ci Chen, Nannan Lu, Kaoru Shimadaand Kotaro Hirasawa,"An Intrusion-Detection Model Based on Fuzzy Class-Association-Rule Mining Using Genetic Network Programming",IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 41, No. 1, 2011.
8. Gang Wang, Jinxing Hao, Jian Ma and Lihua Huang, "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering", Expert System with Applications, Vol.37, No.9, pp.6225‑6232, 2010.
9. DishaSharma,"Fuzzy Clustering as an Intrusion Detection Technique", International Journal of Computer Science & Communication Networks, Vol.1, No.1,2011.
10. IoannisKrontiris, ZinaidaBenenson, ThanassisGiannetsos, Felix C. Freiling and TassosDimitriou, "Cooperative Intrusion Detection in Wireless Sensor Networks", Lecture Notes in Computer Science, Vol. 5432, pp 263-278, 2009.
11. Sarab M. Hameed, SumayaSaad, and Mayyadah F. AlAni, "An Extended Modified Fuzzy Possibilistic C-Means Clustering Algorithm for Intrusion Detection",Lecture Notes on Software Engineering, Vol. 1, No. 3, 2013.
12. Andrew H. Sung and SrinivasMukkamala, "The Feature Selection and Intrusion Detection Problems", Lecture Notes in Computer Science, Vol.3321, pp.468-482, 2005.
13. Peter Lichodzijewski, A. NurZincir-Heywood and Malcolm I. Heywood, "Host-Based Intrusion Detection Using Self-Organizing Maps",Fac. of Comput. Sci
14. Adebayo O. Adetunmbi, Samuel O. Falaki, Olumide S. Adewale and Boniface K. Alese, "Network Intrusion Detection Based On Rough Set And K-Nearest Neighbour", International Journal of Computing and ICT Research, Vol. 2, No. 1, pp. 60 – 66, 2008.
15. SrilathaChebrolu, Ajith Abraham and Johnson P Thomas, "Hybrid Feature Selection for Modelling Intrusion Detection Systems",Lecture Notes in Computer Science,Vol.3316, pp 1020-1025, 2004.
16. Sumathi M and Umarani R, "Advanced Network Intrusion Detection System Based on Effective Feature Selection",International Journal of Computer Science and Information Technologies, Vol. 4, No.1, pp. 107 – 112, 2013.
17. Mohanabharathi R, T.Kalaikumaran and S.Karthi, "Feature Selection for Wireless Intrusion Detection System Using Filter and Wrapper Model",International Journal of Modern Engineering Research (IJMER), Vol.2, No.4, pp-1552-1556, 2012.
18. Shailendra Singh, Sanjay Silakari and Ravindra Patel, "An efficient feature reduction technique for intrusion detection system", International Conference on Machine Learning and Computing ,vol.3,2011.
19. Son T. Nguyen, Hung T. Nguyen and Philip B. Taylor, "Bayesian Neural Network Classification of Head Movement Direction using Various Advanced Optimisation Training Algorithms" International Conference on Biomedical Robotics and Biomechatronics, pp.1014-1019, 2006.
20. U Aickelin, J Twycross and T HeskethRoberts, "Rule Generalization in Intrusion Detection Systems Using SNORT", International Journal of Electronic Security and Digital Forensics, Vol.1, No.1, pp.101-116, 2007.
21. Thomas G. Dietterich and GhulumBakiri,"Solving Multiclass Learning Problems via Error-Correcting Output Codes", International Journal of Artificial Intelligent research,Vol.2, pp.263-286,1995.
22. Santosh Kumar SahuSauravranjanSarangi and Sanjaya Kumar Jena, "A Detail Analysis on Intrusion Detection Datasets", International Advance Computing Conference, pp.1348 – 1353, 2014.
23. Karthik G and Nagappan A, "Intrusion Detection System Using Kernel FCM Clustering and Bayesian Neural Network",International Journal of Computer Science and Information Technology & Security (IJCSITS), Vol. 3, No.6, 2013.
24. P. B. Taylor and H. T. Nguyen, "Performance of a head-movement interface for wheelchair control," Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 2, pp. 1590 - 1593, 2003.
25. T. Joseph and H. T. Nguyen, "Neural network control of wheelchairs using telemetric head movement," Proceedings of the 20th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, vol. 5, pp. 2731 - 2733, 1998.
26. H.T. Nguyen, L.M. King and G. Knight, "Real-time head-movement system and embedded Linux implementation for the control of power wheelchair", Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4892-4895, 2004.