International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 6 | ISSN : 2456-3307

A Review on Intelligent Data Mining and Soft Computing Techniques for Effective Intrusion Detection

Suma S G*, Dr. Ganapathy Sannasi
SCSE, VIT, Chennai, Tamil Nadu, India

ABSTRACT

The world of computer today is increasingly dependent on interconnections between computer systems. Internet usage is increasing rapidly that the security against real time attack is a major concern. An intrusion detection system is a key defensive mechanism against the network attacks. Various approaches to intrusion detection are being used currently. This paper discusses about the various feature selection and intelligent classification techniques that are used to detect intrusions effectively. In addition to this, intrusion detection systems based on intelligent soft computing techniques are also discussed. Finally, a new intrusion detection model based on artificial immune system and fuzzy rough set based feature selection is suggested for effective and dynamic intrusion detection.

Keywords : Intrusion Detection System, Neural Networks, Artificial Immune System, Fuzzy System, Particle Swarm Optimization.

I. INTRODUCTION

A **distributed system** is a set of networked computers which communicate with each other by passing message. They are required for handling very large datasets. The main advantage of distributed system is its scalability. Recently, the internet technology has rapidly grown. The networked systems are inevitably exposed to a large number of threats from internal and external hackers which may lead to loss of critical data and unavailability of services. To defend the legitimacy of the systems and to protect the network from malicious attacks and misuse, information security systems are required.

Intrusion refers to a set of actions or events that compromise the integrity, availability or confidentiality of a system or a network [1].Conventionally Firewalls were used as first level protection layer to end the intrusion attempts. Firewall inspects only the data packet header which is inadequate for safekeeping from all kinds of attacks. They do not inspect the entire contents of the packet. Therefore we need an intrusion detection system (IDS) which can analyze the payload of the packet to detect all types of attacks. An IDS is one that monitors the network and reports any malicious activity or violation of the policies. An IDS protects the network from known attacks, unknown attacks or against network and system overload (denial-of-service attacks).

Based on where the detection takes place IDS can be classified as host based and network intrusion detection systems. HIDS monitors the activity information of a particular system. NIDS detects an attack by monitoring the traffic from the network itself rather than collecting the information from each separate host. Based on the detection method used, IDS can be broadly classified as signature based and anomaly based IDS. Signature based IDS looks for specific patterns to detect a known attack. A database of previous attack signatures and known vulnerabilities is maintained and hence the false alarm rate is low. The signature database must be continuously updated and it is unable to detect unknown attacks for which no signature is available. Anomaly based IDS are used to detect unknown attacks by comparing the behaviour of a network against the baseline behaviour. An alarm is triggered when there is a deviation from the baseline.

Anomaly based IDS can detect novel attacks but it has high false alarms.

Data mining techniques can be applied for intrusion detection. Data mining is the process of extracting the interesting knowledge from a large data store such as database, data warehouses or other data repositories.[2] Intrusion detection can be considered as data analysis process. Data mining approaches can be used to build intrusion detection models. The most popular dataset used for intrusion detection is the Knowledge Discovery and Data Mining –KDD'99,and was developed in 1999 by Massachusets Institute of Technology(MIT).Data mining is used for classification where a data item can be mapped into several predefined categories. Data mining techniques are of great use to researchers in the field of intrusion detection as it makes it easier to analyze huge volume of data and optimizes the detection rules. Traditional IDS has many limitations. For improving the performance of the IDS, data mining techniques such as classification, clustering, association rules can be applied on network traffic. Below is figure for IDS based on data mining.

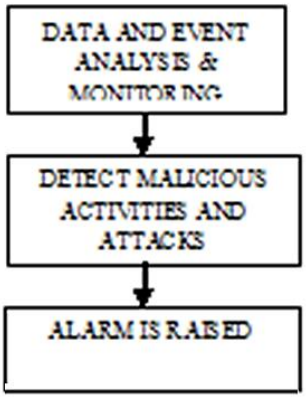

Figure 1. IDS based on data mining

Feature selection and classification algorithms are the data mining techniques that can be used for analysis of the data.

Intelligent IDS are computer programs that are developed using intelligent techniques to provide better detection rate. They analyze the environment and can be situated in a host or a network. They are capable of decision making and a set of agents can be used for decision making. Intelligent IDS is developed by using intelligent techniques for pre-processing and classification. Reprocessing or feature selection is the process of selecting the relevant features that can effectively describe the given problem. The data is reduced by removing the redundant or recurring features without any information loss. Feature selection returns a subset of features in the dataset. The key advantage of feature selection is to reduce the training time and increasing the modelling accuracy. There are three main methods for feature selection: filter methods, embedded methods and wrapper methods. Filter methods are based on general features of the training data and selects variables independent of the model. In wrapper method, a subset of variables is only evaluated. Embedded methods combine the advantages of both the methods. Regularization methods are the common types of embedded methods.

Classification is used to identify the group in which each data instance is related in a given dataset [3]. This is done by learning a model called classifier from a set of trained labels. Anomaly detection based on classification techniques consist of training and testing phase. In the training phase a classifier is learnt using labelled training data available. In the training phase, a test instance is classified either as normal or anomalous, using the classifier. The most important classification approaches are: neural network, Naïve Bayes, fuzzy sets and decision trees[4].

Neural networks: Neural Networks are those systems that are modelled from understanding and emulating the human brain to copy the human abilities and can be used in various fields to categorize the data as intrusive or normal. Generally neural networks consist of a set of interconnected nodes where each node has a weight associated with it. To predict the correct class label of the input, the network learns in the learning phase by adjusting the weights. An artificial neural network consists of connected set of processing units where each connection has a weight which determines how one unit will affect another. The subsets of these processing units act as input and output nodes, and the remaining nodes forms the hidden layer. Neural network does a functional mapping of input values to output values by the activation of each of the input node and allowing them to propagate through the hidden layer nodes to the output nodes.

Naive Bayes: Naive Bayesian classifiers [5] use Bayesian theorem to classify the new instances of a data sample. Bayes theorem provides a way of calculating the posterior probability. Each instance is a

set of attribute values which can be described by a vector,X =(x1, x2,…,xn). Consider there are m classes .The sample X is assigned to the class C_i if and only if $P(X \mid C_i)\, P(C_i) > P(X \mid C_j)\, P(C_j)$ for all i and j in (1, m) such that j <> i. The sample then belongs to the class having maximum posterior probability.

Fuzzy sets: Fuzzy logic is an intelligent technique that has been successfully employed in many IDSs. Fuzzy sets [6,7] is one of the key methodology for representing and processing uncertain information. Uncertainty arises in different forms such as imprecision, inconsistency, non-specificity, etc. Fuzzy sets make the system complexity manageable by exploiting uncertainty. Fuzzy sets are an approach that helps in developing uncertain models of the data that provide smarter and smoother performance than traditional systems.

Decision Trees: A decision tree [8] is a tree where a test or a decision on a data item is represented by each non-terminal node in the tree. The choice of a certain branch depends upon the outcome of the test. In order to classify a particular data item, the decision tree algorithms start at the root node and follow the assertions down until it reaches a terminal node. When a terminal node or a leaf node is approached, a decision is made. Decision trees are also taken as a special form of a rule set, characterized by their hierarchical organization of rules. Using decision trees which gains insight for intrusion detection.

KDD'99 is the most widely used dataset for the evaluation of different intrusion detection methods. The KDD'99 cup dataset was built based on the data captured in DARPA'98 evaluation program. Each connection record is represented and described by 41 attributes in this dataset. The attributes consists of both continuous and discrete variables. The table below shows four categories of simulated attacks:

Table 1. Attack categories

Attack Type	**Description**
DoS	Denial of Service makes a resource too busy to serve a legitimate request.(e.g. Syn flood).
U2R	Unauthorized access to super user or root functions.(e.g. buffer overflow attacks).
R2L	Unauthorized access from a remote machine(e.g. guessing password).
Probe	Surveillance and scanning the machine to determine any vulnerabilities in the system(e.g. port sweep, mscan).

KDD'99 features are classified into three categories: basic features, content features ad traffic features.

Basic features are attributes extracted from a TCP/IP connection. The features are extracted from the packet header. Content features looks for suspicious behaviour in the payload portion. The payload of the original TCP packet is evaluated. Traffic features are features that are computed with respect to a window interval. Traffic features are categorized into same host features and same service features. In "Same host" features, the connections in the past two seconds that have the same destination host are only examined. In "Same service" features, the connections in the past two seconds that have the same service. Both types are referred to as "time based traffic features".

The table below lists the set of features defined for connection records.

Table 2. List of KDD features

Basic features	**Content features**	**Traffic features**
Duration	Hot	Count
protocol_type	num_failed_logins	serror_rate
service	logged_in	rerror_rate
src_bytes	num_compromised	same_srv_rate
dst_bytes	su_attempted	diff_srv_rate
Flag	root_shell	srv_count
Land	num_root	srv_serror_rate
wrong_fragment	num_file_creations	srv_rerror_rate
Urgent	num_shells	srv_diff_host_rate
	num_access_files	
	num_outbound_cmds	
	is_hot_login	
	is_guest_login	

The rest of paper is organized as follows. In **Section 2**, presents a review of the literature, feature selection & classification technique are discussed. Recent works on

intrusion detection system are discussed. In **Section 3**, Comparative analysis is shown. The performance of different IDS approaches is compared. **Section 4**, the proposed work for intrusion detection is discussed and **Section 5** summarizes the conclusion.

II. LIETRATURE SURVEY

2.1 Works on Intelligent IDS:

The recent computational intelligence approaches for IDS includes neural network based IDS, fuzzy and rough sets, genetic algorithms, particle swarm intelligence and soft computing based methods.

Neural Network Based IDS: Basant Subba et al[9] proposed an IDS based on artificial neural network model which used various optimization techniques to minimize the overall computational overhead involved in training and execution. The proposed artificial neural network based IDS used the feed forward and the back propagation algorithms along with other optimization techniques to maintain a high performance level. The proposed ANN model used only a single hidden layer and hence its computational cost is comparatively less than other SVM or C4.5 based IDS.

Fuzzy Set based IDS:

Another approach for intrusion detection is based on **fuzzy rule based systems** (FRBCS) to correctly identify all type of attacks [10].Salma Elhag et al [11] proposed a combination of genetic fuzzy systems and pair wise learning for improving detection rates on Intrusion Detection Systems. This approach is based on a divide and conquers strategy, in which the original problem is divided into sub problems. The binary sub problems are independently learned by different base classifiers and their outputs are then combined to classify an instance. This approach makes use of a Genetic Fuzzy System (GFS) [12] for intrusion detection. GFS provides a better separation of different types of alarms and higher interpretability of the rule-set. This makes use of one versus one (OVO) methodology in which the usage of pair wise learning is an efficient way of overcoming multi class problem [13]. This method is a combination of GFS and OVO for detecting intrusions.

The approach makes use of a new FRBCS which is called as Fuzzy Association Rule-Based Classification for High-Dimensional problems (FARC-HD). The main aim of the FARC-HD algorithm is to obtain an accurate and compact fuzzy rule-based classifier [14]. The model consists of three processes:

[1] Extraction of fuzzy association rule for classification: The fuzzy association rules for classification are generated by employing a search tree and limiting the depth of the branches to find short fuzzy rules.
[2] Pre-Screening candidate rule: Here the most interesting rules are pre-selected by means of a subgroup discovery mechanism.
[3] Selection of Genetic rules and tuning is the final process.

Classification is done via binarization techniques. OVO approach is also known as pair wise classification in which the original m-class problem is transformed into m.(m-1)/2 binary subsets. A query is submitted to all binary models for classifying the instances and the output of these models are combined by constructing a score-matrix to decide the final class.

The performance of the approach was evaluated in terms of Accuracy, Mean F-Measure, Average accuracy, Attack accuracy, Attack Detection Rate and False Alarm Rate. The FARC-HD OVO algorithm has best trade-off among the different performance measures, especially regarding the average accuracy, attack accuracy and false alarm rate. Higher accuracy rate indicate better identification of individual classes of the problem. The use of divide and conquer strategy improved the individual accuracy of different classes and was reflected as a high value for average accuracy metric.

Artificial intelligence researchers proposed some optimization techniques called Swarm intelligence which was inspired from animal's collective behaviour such as bird flocking, fish schooling and ant colonies. In particle swarm optimization [15], each individual of population is called as a particle. Here, each particle updates its velocity and position in all iterations properly. A fitness function is defined on a particle's location. The optimization problem is to find the best position, one that minimizes the fitness function.

Genetic algorithm is an evolutionary computational technique inspired by natural evolution.GA process

uses the selection operator, crossover operator and mutation operator to evolve the chromosomes. The process initially consists of a randomly generated population of chromosomes which represents all possible solution of a problem. A fitness function is used as an evaluation function. The effectiveness of the algorithm is affected by: the fitness function, GA parameters and the representation of individuals.

Ketan Sanjay et al [16] proposed an evolutionary approach for feature selection. Feature selection is based on mathematical intersection principle using genetic algorithm (GA). Feature selection done by the proposed method was compared with different feature selection (FS) techniques like CFS, IG and CAE. The proposed method selected the minimum number of features from the NSL KDD data set which improves the Naïve Bayes classifier accuracy along with reduced time complexity.

2.2 Recent Works on IDS:

Soo-Yeon Ji and Bong-Keun Jeong[17] proposed an approach for **multilevel intrusion detection method** which consist of 3 main steps: rule generation, determining the attack categories and also a visual analysis to provide transparent reasons by integrating a visual analytics tool. The study used the NSL-KDD dataset to verify the approach and achieved 96% accuracy in detecting the attack categories. Rule based model was designed using the Classification and Regression Tree (CART).CART captures the patterns of the input data and creates a decision tree. In the study three attack categories were considered: Denial-of-service(DOS),Probe and R2L.Feature extraction and selection was done using Discrete Wavelet Transform(DWT).Once the features are extracted, the relevant features are only selected to generate the classifier. The classifier can be used to detect the exact category of attack. Visual analysis provides detailed understanding and analysis on the network traffic data.
The visual analytics tool was used to showcase and analyses the intrusions identified and to find their reasons efficiently. iPCA tool provides dimension contribution analysis to identify the features that represents different attacks. Clear separation among attack categories were identified when DWT features were used.

DOS attacks mainly aims at disrupting the network service operations by wasting the processing and storage capacities on unwanted tasks. Ognjen Joldzic et al [18] proposed a distributed and scalable solution for the detection of Denial-of-Service (DOS) attacks. The approach mainly focused on providing a transparent system that can detect and mitigate the DOS attacks.

The solution relied on the benefits provided by the Software Defined Networking (SDN). For high speed networks, the performance of SDN was not effective. This problem was solved by the use of a load balancing algorithm. A load balancing algorithm was used to equalize the load on each of the active processors and for capacity scaling by including additional processors.

Omar Y. Al-Jarrah et al [19] proposed a **botnet** intrusion detection method built on randomize data partitioned learning model [RDPLM].Most of the existing botnet IDS are rule-based. Previously known botnet signatures are used to compare the signatures of the network traffic. This is inefficient since the rules needs to be updated and it is difficult to update the rules due to the increase in the network traffic. In this approach a novel feature ranking and Voronoi-clustering-based data partitioning techniques were used to remove redundant features. To detect the botnet intrusions, the network flow characteristics were used rather than the content of the payload. The experimental results in the study show that it has very high detection accuracy and a reduced computational cost.

Using an **ensemble** approach, the classification accuracy can be improved by combining the multiple classifiers. Abdulla Amin Aburomman and Mamun Bin Ibne Reaz [20] proposed a novel SVM-kNN-PSO ensemble method for intrusion detection .This approach focuses on an ensemble construction method that used PSO generated weights to create ensemble of classifiers to detect intrusion. To find the behavioural parameters for PSO; LUS method was used as a meta-optimizer. Three approaches were used to combine the expert opinion: POS, meta-optimized PSO and WMA. Using the KDD99 dataset, the comparison of the three approaches was done. The experimental results obtained in this approach shows that the best results were obtained with PSO.

Parham Moradi and Mehrdad Rostami [21] proposed a **feature** selection method based on the approach of **graph clustering and ant colony optimization**. This

approach involves three main processes: It firstly represents the entire feature set as graph. Secondly, the features were divided into several clusters using a community detection algorithm. Then, the selected final subset of features using the modified ant colony optimization based search process. The proposed approach can be classified as filter-based feature selection method, as the feature subset is evaluated by means of a seperability index matrix and doesn't need any learning models. The proposed method was evaluated in terms of classification accuracy, size of subset selected features and execution time. Classification accuracy was evaluated on different classifiers like SVM, DT, NB, KNN and RF. The evaluation reported the best result for the NB classifier. The proposed method not only resulted in better accuracy but also the execution time of the proposed method was comparable with other feature selection methods.

AIS is one of the **evolutionary** methods that is used for developing classification approaches.AIS approaches have been successfully applied in many areas such as classification, clustering, pattern recognition and optimization[22].The performance of the AIS approach is better when compared to other approaches such as artificial neural network, genetic algorithms, fuzzy systems.

Meng-Hui Chen et al [23] proposed a population-based incremental learning approach with artificial immune system for network intrusion detection. This proposed an approach where PBIL integrated into AIS classification for intrusion detection. The PBIL algorithm was introduced by Baleya and Caruana in 1995.It is a type of genetic algorithm which maintains the statistics contained in any problem. PBIL algorithm consist of following steps: 1.Population is generated from a probability vector. 2. Each member is evaluated and ranked based on its fitness. 3. The population probability vector is updated based on the fittest individuals. 4. Mutation5. Repeat 1 through 4.Collaborative Filtering: For a given set of items and set of users the user may rate a subset of items and the system predicts a missing user rating for an item. Collaborative filtering techniques can be classified into two based on their processing approaches: model based and memory based. Memory based considers a group of similar users or items. It predicts the missing user rating for items or users with similar profile.

Meng-Hui Chen et al [23] proposed an approach to develop a classifier named as PBIL-AIS ^{CF} classifier, to solve classification. The study focuses on the generation of multiple types of antibodies to solve classification problem. This approach consists of three phases: Data collection and pre-processing phase, evolution phase or applying PBIL-AIS for affinity calculation, creation of new antibodies and detection phase where detection rules of proposed classifier are employed. The performance of the AIS evolution effect can be improved by the PBIL process for the creation of new antibodies. Other methods like negative selection mechanisms can be implemented into AIS for improvement of the performance. The performance of the model was evaluated against KDD99 Cup dataset. The performance of the proposed approach was compared with other algorithms such as SVM, Naïve Bayes, decision trees, AIRS and CLONALG.The accuracy of the model was evaluated in terms of TP, TN, FN and FP.

Accuracy = TP + TN/ TP+TN+FN+FP .The approach was compared with other methods and the PBIL-AIS ^{CF} yielded a better average accuracy ratio.

III. COMPARATIVE ANALYSIS

Basant Subba et al[9] compared the performance of proposed Artificial Neural Network basedIDS with the Naive Bayes model, SVM and C4.5 based models on NSL-KDD dataset.

Table 3. Comparison of various intrusion detection models on NSL-KDD dataset

Feature Selection	Accuracy	Detection Rate
SVM	99.63	99.16
Naïve Bayes	97.16	91.65
C4.5	99.25	96.98
ANN	98.86	95.77

Meng-Hui Chen et al [23] proposed a population based incremental learning approach. The performance of PBIL-AIS algorithm in the sampling dataset of KDD99 Cup is shown in Table. The PBIL-AIS model is average in accuracy.

Table 4. Performance on KDD99 Cup Dataset

	Min	Average	Max

	Accuracy	Accuracy	Accuracy
Training set	95.10	97.03	98.21
Testing set	92.66	93.2	94.01

Table 5: Comparison of algorithms on KDD'99 Cup dataset

Algorithm	Accuracy
SVM	98.70
KNN	98.91
CLONALG	75.52
Immunos-81	95.20
PBIL-AIS	97.03

The performance of the algorithm was compared with other popular algorithms on KDD99 Cup Dataset.The results shows that the AIS algorithm yields a much better average accuracy than other evolutionary algorithms like CLONALG and Immunos-81.

Table 6. Summary and comparison of existing works on intrusion detection

Paper	Approach	Feature selection	Advantage	Disadvantage
A Novel SVM-kNN-PSO Ensemble Method For Intrusion Detection System	Combines responses produced by multiple classifiers into a single response.	No	Classification accuracy is improved by combining opinions from multiple experts into one.	It requires longer running time
A Transparent And Scalable Anomaly-Based DoS Detection Method	Mechanism that aims to calculate a score(entropy) for each packet and decide if it is legal or not by considering the score value.	No	Better accuracy and attack prevention efficiency for known and unknown attacks.	Memory and communication overhead issues.
Feature Selection In Mixed Data: A Method Using a Novel Fuzzy Rough Set-Based Information Entropy,	Fuzzy rough set-based information entropy for feature selection in a mixed data set.	Filter-wrapper based	Filter-wrapper method has used to achieve the classification accuracy.	-
Data Randomization and Cluster-Based Partitioning for Botnet Intrusion Detection	Eliminates the redundant features and reduces the dataset volume for botnet intrusion	Feature ranking and clustering based	high detection accuracy and low computational cost	Memory usage limitations.
Integration of Graph Clustering With Ant Colony Optimization For Feature Selection	Features are divided into several clusters andACO method was developed to select the final subset features.	Filter based	Minimizes the redundancy and classification accuracy was increased .	The algorithm has several parameters which needs to be tuned.
A Neural Network Based System for Intrusion Detection and Attack Classification	Used feed forward and the back propagation algorithms	No	computational cost is comparatively reduced due to single hidden layer	Accuracy and detection rate is less than SVM. Processing time is a concern
A population-based incremental learning approach with artificial immune system for network intrusion detection	A classifier using an artificial immune system (AIS) combined with population-based incremental learning	No	Have the properties of self-learning and adaptability.	Average accuracy is slightly less than SVM

IV. PROPOSED MODEL

The proposed model is to develop a classifier using artificial immune system combined with fuzzy rough set feature selection. The proposed model is a new intrusion detection system which has the properties of self-learning and adaptability.

Feature selection is the process of removing redundant features without decreasing the prediction accuracy. A fuzzy rough set based feature selection algorithm is proposed. Fuzzy rough set can be used to deal with real valued data and also to process mixed data. Fuzzy rough sets are based on fuzzy relations which are allowable to be defined for different kinds of attributes to measure the similarity between objects.

Proposed System Architecture

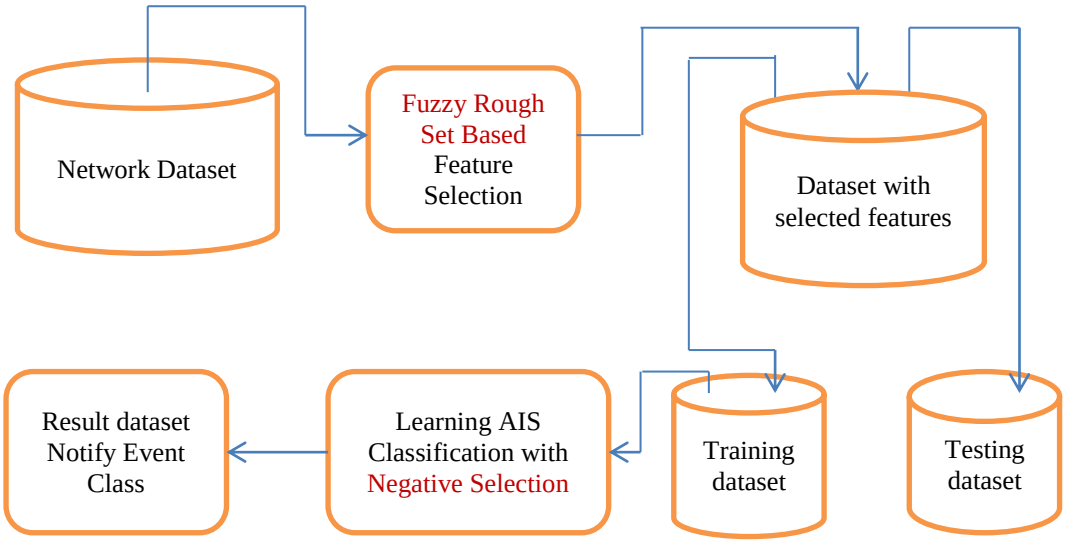

Figure 2. Proposed System Architecture

A distance function is employed to define the fuzzy relations for the real-valued attributes. The fuzzy relations and fuzzy rough set theory provide a new framework for dealing with mixed data. To introduce a new classification algorithm based on artificial immune system. Fuzzy Rough set based feature selection method removes the redundant and irrelevant features in order to improve the performance of the classifier and reduces the computation time. The AIS approach classifies the records into normal and anomaly categories. New features can be added to find novel attacks. The Negative selection mechanism can be implemented into the AIS improve the performance of the method.

AIS Algorithm:

1. Input training data.
2. Initial antibody pool.
3. Calculate the afffintiy for each antibody with the incoming antigen.

   $\text{Affinity}_j = \frac{1}{hj+ej}$ where j=1,2,3….

   Affinity_j is the affinity value for j^{th} antibody with current antigen, h_j is the net distance as nominal data, e_j is the net distance as numeric data.
4. If detected correct class, then

   Clonal expansion mechanism

   $W_t^{new} = W_t^{old} + \Delta\tau t$

   Antibody hierarchy mechanism

   Else

   Create new antibody
5. Repeat step 3 to 4 till all antigens have been processed.

V. CONCLUSION

The most important issue in computer network security is the detection of intrusion attacks. This paper is a survey on various feature selection and classification techniques that can be used for intrusion detection in a network. The advantages and disadvantages of various approaches towards intrusion detection have been discussed. Intrusion Detection System based on neural networks, genetic algorithm, particle swarm

optimization, fuzzy systems and artificial immune system have been included in the discussion. In addition, a new classifier using artificial immune system and fuzzy rough set feature selection has been proposed which has the properties of self-learning and adaptability. The performance of AIS can further be improved by implementing negative selection algorithm. The proposed approach aims at improving the detection accuracy and reduces the computational time.

VI. REFERENCES

[1]. William Stallings, "Cryptography and Network Security Principles and Practices", Prentice Hall, Upper Saddle River, 2006.

[2]. Krishna Kant Tiwari, Susheel Tiwari, Sriram Yadav, "Intrusion Detection Using Data Mining Techniques", International Journal of Advanced Computer Technology, Vol.2, No. 4, pp.1-3, 2010.

[3]. Sagar S.Nikam, "A Comparative Study of Classification Techniques in Data Mining Algorithms", Oriental Journal of Computer Science and Technology, Vol.8, No.1, pp.13-19.

[4]. Sannasi Ganapathy, Kanagasabai Kulothungan, Sannasy Muthurajkumar, Muthusamy Vijayalakshmi, Palanichamy Yogesh and Arputharaj Kannan, " Intelligent feature selection and classification techniques for intrusion detection in networks:a survey", EURASIP Journal on Wireless Communications and Networking, Vol.271, pp.1-16, 2013.

[5]. S Chebrolu, A Abraham, P Johnson, Thomas, "Feature deduction and ensemble design of intrusion detection systems", Computers & Security, Vol.24, No.4, pp.295-307, 2005.

[6]. W Zhang, S Teng, H Zhu, H Du, X Li, "Fuzzy Multi-Class Support Vector Machines for Cooperative Network Intrusion detection", Proc.9th IEEE Int.Conference on Cognitive Informatics (ICCI'10), pp.811-818, 2010.

[7]. L Zadeh, "Role of soft computing and fuzzy logic in the conception, design and development of information/intelligent systems, in Computational Intelligence:Soft Computing and Fuzzy-Neuro Integration with Applications", Proceedings of the NATO Advanced Study Institute on Soft Computing and its Applications, Vol.162, pp.1-9, 1998.

[8]. SS Sivatha Sindhu, S Geetha, A Kannan, "Decision tree based light weight intrusion detection using a wrapper approach", Expert Syst.Applications, Vol.39, pp.129-141, 2012.

[9]. Basant Subba , Santosh Biswas, Sushanta Karmakar , "A Neural Network Based System for Intrusion Detection and Attack Classification", Twenty Second National Conference on Communication (NCC), pp.1-6, 2016 .

[10]. Ishibuchi, H and Yamamoto, T."Rule weight specification in fuzzy rule-based classification systems", IEEE Transactions on Fuzzy Systems, Vol.13, pp.428-435, 2005.

[11]. Salma Elhag, Alberto Fernande, Abdullah Bawakid , Saleh Alshomrani , Francisco Herrera"On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems", Expert Systems with Applications, Vol.42, pp.193-202, 2015.

[12]. Cordon, O., Gomide, F., Herrera, F., Hoffmann, F., & Magdalena, L, "Ten years of genetic fuzzy systems:Current framework and new trends.Fuzzy Sets and Systems", Vol.141, No.1, pp.5-31, 2004.

[13]. Hastie, T., & Tibshirani R, "Classification by pairwise coupling", The Annals of Statistics, Vol.26, No.2, pp.451-471, 1998.

[14]. Alcala-Fdez, J., Alcala, R., & Herrera, F, "A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning", IEEE Transactions on Fuzzy Systems, Vol.19, No.5, pp.857-872, 2011.

[15]. Seyed Mojtaba Hosseini Bamakan, Huadong Wang, Tian Yingjie , Yong Shi , "An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization", Neurocomputing, Vol.199, pp.90-102, 2016.

[16]. Ketan Sanjay Desale, Roshani Ade"Genetic Algorithm based Feature Selection Approach for Effective Intrusion Detection System", International Conference on Computer Communication and Informatics, pp.1-6, 2015.

[17]. Soo-Yeon Ji, Bong-Keun Jeong, Seonho Choi, Dong Hyun Jeong, "A multi-level intrusion detection method for abnormal network behaviours", Journal of Network and Computer Applications, Vol.62, pp.9-17, 2016.
[18]. Ognjen Joldzic, Zoran Djuric, Pavle Vuletic, "A transparent and scalable anomaly-based DoS detection method", Computer Networks, Vol.104, pp.27-42, 2016.
[19]. Omar Y.Al-Jarrah, Omar Alhussein, Paul D.Yoo, Sami Muhaidat, Kamal Taha, Kwangjo Kim, "Data Randomization and Cluster-Based Partitioning for Botnet Intrusion Detection", IEEE Transactions on Cybernetics, Vol.46, No.8, pp.1797-1801, 2016.
[20]. Abdulla Amin Aburomman, Mamun Bin Ibne Reaz, " A novel SVM-kNN-PSO ensemble method for intrusion", Applied Soft Computing, Vol.38, pp.360-372, 2015.
[21]. Parham Moradi, Mehrdad Rostami, "Integration of graph clustering with ant colony optimization for feature selection", Knowledge-Based Systems, Vol.84, pp.144-161, 2015.
[22]. Alatas, B., Akin, E., 2005."Mining fuzzy classification rules using an artificial immune system with boosting", Adv.Databases Inf.Syst.3631, pp.283-293, 2005.
[23]. Meng-Hui Chen, Pei-Chann Chang, Jheng-Long Wu, "A population-based incremental learning approach with artificial immune system for network intrusion detection", Engineering Applications of Artificial Intelligence, Vol. 51, pp. 171-181, 2016.