International Journal of Research in Advent Technology (IJRAT)
Vol. 1, No. 2, September 2013, ISSN: 2321–9637

# SMART NETWORK INTRUSION DETECTION SYSTEM USING HYBRID APPROACH

Nivedita P. Chaudhari [1] , Dr. Leena Ragha [2]
[1 2] *Computer Department , RAIT, Mumbai*
[1] *M.E. Student, RAIT, Nerul, Navi Mumbai*
[2] *Head of Computer Department, RAIT, Nerul, Navi Mumbai*
[1] niveditachaudhari88@gmail.com

**ABSTARCT:**
**As network-based computer systems play increasingly vital roles in modern society, they have become target of enemies and criminals. Therefore, we need to find the best ways possible to protect systems. Intrusion detection is a challenging area of research in current scenario. Many studies and research has been done already in network intrusion detection system, but still there is need to get a lot of attention. Our object is we affect a "smart" method for intrusion detection using k-mean clustering and probabilistic classification based on fuzzy GNP hybrid rule mining. K-mean clustering is task of grouping objects or known structures in data that are in some way or another similar, without using known structures in data. In our method, k-mean algorithm is used in another fashion such as for filtration purpose to remove irrelevant data. The aim is to make system time efficient. Genetic Network Programming (GNP) is one of the fields from biological computation uses directed graph structure. To deal with both discrete and continuous attributes, fuzzy set theory of combination of triangular and triangular membership functions. Genetic operations such as crossover, mutation-1, mutation-2 and fitness function are used to generate more number of strong hybrid rules. Rather than using conventional measurement methods for hybrid rules, new technique called $\chi^2$ evaluation is used. Genetic operations changes fuzzy membership parameters changes. Michalewicz's operator is used to overcome this problem in our methodology. At the end we have strong and robust hybrid rules without loss of information, which is used for classification purpose to match data using proposed classification algorithm. Proposed method flexibly can be applied to anomaly and misuse detection in network intrusion problem. Experimental results with KDD99CUP dataset shows our system provides high accuracy and detection rate with low negative false rate as compared with other machine-learning techniques and hybrid approach.**

***Keywords:*** *Data mining, Evolutionary Computation, Intrusion Detection, Fuzzy Class Association Rule Mining*

## 1. INTRODUCTION

As computer communication technologies develop, damages caused by intrusion and crimes related to computer systems have been increasing. Network based computer systems play vital role in modern society as use of internet is increasing widely. They are becoming target for enemies and criminals. So there is need to build the system which can protect computer systems. The term security comprises when intrusion takes place. Intrusion can be defined as any action that compromises integrity, confidentiality and availability of data or resources. Intrusion detection system (IDS) [1] monitors network traffic and malicious activities. Alert system and separation of abnormal data from normal data are main key functions of IDS. The problem of IDS can be defined in two parts; extraction of useful data from history and organization of such useful information for accurate classification. IDS deals with two kinds of intrusion; one is anomaly intrusion and another is misuse intrusion.

Though many studies and research [8] have presented their work to resolve network intrusion detection problem, there is need smart model which can detect any intrusion to network as well as predict and handle new types of intrusion. When term "smart" comes, we need technology that can detect new intrusions based on previously known intrusions. That's why data mining has become vital part of network intrusion.

K-mean clustering algorithm [9] is one data mining techniques. This algorithm follows simple procedure to classify data through certain number of data. In our work, k-mean algorithm is used as filter algorithm. According to procedure, it will cluster data having similar characteristics. From such clusters we will remove data from such cluster having less number of records. The purpose of such filtration is to make system time efficient and generation of hybrid rules will become easy.

An evolutionary algorithm such as genetic algorithm [10] [13] and genetic programming [11][13] represent their solution using gene structure and tree structure respectively. Genetic network programing (GNP) uses directed graph structure to represent solution giving advantage of reusability of nodes. By combining fuzzy set theory with GNP proposed system can handle both discrete and continuous attributes. In order to get more hybrid association rules, genetic operations such as selection, crossover, mutation-1 and mutation-2 are used. Support, confidence and $\chi^2$ factors are evolved for measurement of generated hybrid rules. In addition, new fitness function that provides the flexibility of mining more rules and mining rules with higher accuracy is given in order to adapt to different kinds of detection. Proposed non-uniform mutation of fuzzy membership function is used to adjust parameters of fuzzy membership function after genetic operation.

After extraction of class-association rules, these rules are used for classification. In proposed method, probabilistic classification is used. For detection, the normal-pattern rules and intrusion-pattern rules are extracted from training dataset. Probabilistic classification is basically designed for anomaly and misuse detection, in order to classify new data correctly. The probability distribution function of average matching distribution of data with rules can be calculated. Based on this, new connection data can be classified into normal activity, misuse and anomaly intrusions.

The paper is organized as follows. Section 2 gives idea about k-mean algorithm and how it is used for data preprocessing. The basic concept of GNP and class association rule is explained in section 3. GNP based class association rule mining with evolutionary operation and Michalewicz's operator overviewed in section 4. Section 5 described proposed classification method for misuse and anomaly detection based on rules from both normal and attack rule pool. Simulation results with KDD99CUP [6] are given in section 6. Conclusion is mentioned in section 7.

## 2. K-MEAN CLUSTERING ALGORITHM AND DATA PREPROCESSING

Clustering is a method of grouping data into specified number of clusters, such as data from same clusters are quite similar whereas data from different clusters are quite different from each other. K-means is well known unsupervised learning algorithm used for clustering. The procedure is simple and easy to group a given data through certain number of clusters, let say k number of clusters. The mean value of numerical data contained within each cluster is called as centroid. In proposed algorithm, k must be greater than number of attacks.

The algorithm is composed of following steps [9]:

(1) Place K number of records, represented by data that are being clustered, as initial centroids.
(2) Assign each record to the cluster that has closest centroid.
(3) When all records have been assigned, recalculate the position of K centroids.
(4) Repeat step 2 and 3 until centroids no longer move. This results into separation of records into respective group.

The main goal of K-means clustering algorithm is to group data into specified number of clusters. The approach is generally used in intrusion detection system to classify data into normal and attack instances. In proposed algorithm, the approach of k-means clustering as classification has been changed to filtration. The referred dataset contains both discrete and continuous attribute. These both attributes before applying to k-

means algorithm are converted into numerical format to deal with clustering using numeric conversion and vertical division method. The data records then grouped into k clusters. The clusters containing minimum number of records are filtered out so as dataset contain only similar type of records. Because of filtration, from remaining dataset strong and efficient number of rules can be generated with less time required as compared with existing systems.

## 3. BASICS OF GNP AND CONVENTIONAL CLASS ASSOCIATION RULE MINING

In this section, basic idea of genetic network programming and conventional class association rule mining is given.

### 3.1. *Structure of Genetic Network Programming*

A large number of studies been conducted on evolutionary optimization techniques. Genetic Algorithm and Genetic Programming are typical evolutionary algorithms. GA evolves strings and mainly applied to optimization problem.GP was devised later which uses tree structure for solution. GNP [15] is one of the evolutionary algorithm which uses directed graph structure instead of trees and strings. Directed graph structure of GNP is represented by gene structure which consists of judgment nodes and processing nodes [2]. Function of judgment node is to examine attribute of tuple and return a judgment result either "Yes" or "No". Then, corresponding branch of judgment node is connected to next node. $J_1$, $J_2$… $J_N$ *are* judgment functions by judgment nodes showing attributes to be examines where $N$ is the total number of judgment functions. On the other hand processing nodes are set of $P_1$, $P_2$… $P_{M,,}$ where $M$ is total number of processing functions which works as action/processing functions. Execution starts from start node, then next node to be executed is determined according to the connection between nodes and a judgment result of current activated nodes. Fig. 1 shows the gene of a node in GNP individual. Fig. 2 is the genotype expression of GNP in GNP individual. 0 is start node, 1 is processing node, 2 is judgment node. $ID_i$ works as identification number. $C_{i1}$, $C_{i2}$… $C_{ij}$ means the nodes connected from node $i$, firstly, secondly and so on.

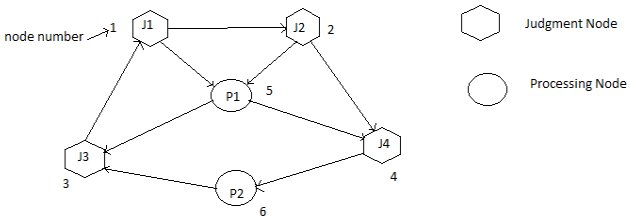


Fig. 1. Basic Structure of GNP

| | $NT_i$ | $ID_i$ | $C_{i1}$ | ..... | $C_{ij}$ |
|---|---|---|---|---|---|
| Node 1 | 2 | 1 | 2 | 5 | |
| Node 2 | 2 | 2 | 4 | 5 | |
| Node 3 | 2 | 3 | 1 | | |
| Node 4 | 2 | 4 | 6 | | |
| Node 5 | 1 | 5 | 3 | 4 | |
| Node 6 | 1 | 6 | 3 | | |

Fig. 2. Genotype Expression of GNP

### 3.2. *Conventional Class Association Rule Mining*

Association rule mining find correlation between features or attributes used to describe a data set [20]. Let $I = \{A_1, A_2, ..,A_N\}$ be set of attributes and each tuple $T$ has a set of attributes satisfying $T \subseteq I$. Association rule can be represented as $X \Rightarrow Y$ . It means, tuple satisfying $X$ are likely to satisfy $Y$, where $X$ is antecedent and $Y$ is conclusion of the rule. *Support*($X$) = $x$ shows the fraction of tuples which contains $X$ in the database equals $x$. The measure of strength of rule is called confidence defines as ratio of *support* ($X \cup Y$) /*support* ($X$). Let *support* ($X$) = $x$, *support* ($Y$) = $y$, *support* ($X \cup Y$) = $z$ and number of database tuples equals $N$. If event $X$ and $Y$ are independent, then *support* ($X \cup Y$) = $xy$. Then, $\chi^2$ of rule $X \Rightarrow Y$ can be calculated as [19]

$$\chi^2 = \frac{N(z - xy)^2}{xy(1 - x)(1 - y)} \quad (1)$$

Important association rule mining rule should satisfy following conditions:

$$\chi^2 > \chi^2_{min} \quad (2)$$

$$\text{support} \geq \text{sup}_{min} \quad (3)$$

$$\text{confidence} \geq \text{conf}_{min} \quad (4)$$

where, $\chi^2_{min}$ , $\text{sup}_{min}$ , $\text{conf}_{min}$ are minimum values given in advance. The main aim of class association rule mining is to find all rules from the database.

Let $A_i$ be the attribute in a database and $k$ be the class labels. Then class association rule can be represented as

$$(A_i = 1) \wedge \ldots \wedge (A_j = 1) \Rightarrow (C = k) \quad (5)$$

Judgment nodes examine values of attributes of database tuples and processing nodes calculates measurements of association rules in GNP. Judgment result of Yes or No by judgment nodes determine next node in graph structure corresponding to Yes-side or No-side. Yes-side of judgment node is connected to next judgment node whereas No-side is connected to next processing node for further processing.

How class association rules are generated is shown in fig. 3. $P_1$ is processing node from where transition starts. $A_1$, $A_2$, $A_3$ denotes the function of judgment nodes. $N$ is the number of total tuples. *a, b, c, a(1), b(1), c(1)* are number of tuples moving to Yes-side at each judgment node and number of tuples moving to Yes-side at each judgment node under condition of class 1, respectively. Number of rules can be derived in form like equation 5.

## 4. GNP BASED HYBRID RULE MINING

In this paper, hybrid rule mining is used for extraction of rules. Hybrid rule mining utilizes both discrete and continuous attributes in one single rule.

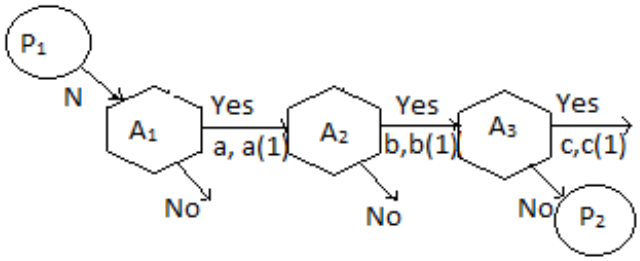


Fig. 3. Node Transition in Conventional Class Association Rule Mining

### *4.1 Fuzzy Membership Function for Continuous Attributes*

In this paper, we utilize advantage of fuzzy theory to have every continuous attribute value in [0, 1]. Fuzzy theory allows complex system to have linguistic description [16]. In the paper, each continuous attribute in the database is transformed into 5 linguistic terms (Vey low, Low, Middle, High, Very high). By using 5 linguistic terms for single continuous attribute, we will get more accurate membership value for corresponding continuous attribute.

The linguistic terms are defined by the combination of trapezoidal and triangular membership functions. The parameters *x1*, *x2*, *x3*, *x4* and *x5* are also evolved along with the evolution of GNP. Each continuous attribute have its own membership value. The parameters for each continuous attribute are initialized by analyzing the distribution of data. During evolution process, the parameter value of fuzzy membership function should be adjusted generation by generation. Fuzzy membership values are used to determine the transition in GNP individuals while searching for association rules.

Table I shows example of small database with two continuous attributes. Fig. 4 and fig. 5 show the fuzzy membership function for attribute $A_1$ and $A_2$, respectively.

### *4.2 Hybrid Rule Mining based on Fuzzy GNP*

The conventional representation of class association rule based on GNP is shown in Fig. 3. Hybrid rule mining is combination of fuzzy GNP and conventional class association rule mining. Hybrid rule mining have the advantage of utilizing both discrete and continuous attributes in one single rule. Fig. 7 describes an example of hybrid rule mining representation. Rule extraction starts from processing node $P_1$. The first judgment node examines the fuzzy membership value of continuous attribute $A_1$, the second judgment node examines the value

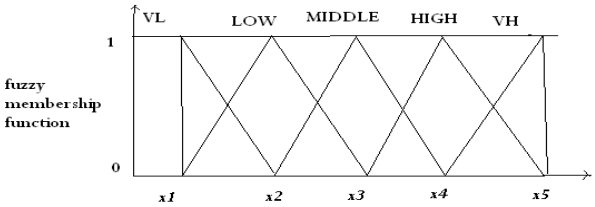

Fig. 4. Fuzzy Membership Function

Table I Example of Small Database

| TID | $A_1$ | $A_2$ |
|---|---|---|
| 1 | 100 | 10000 |
| 2 | 200 | 8000 |
| 3 | 300 | 6000 |
| 4 | 400 | 4000 |
| 5 | 500 | 2000 |

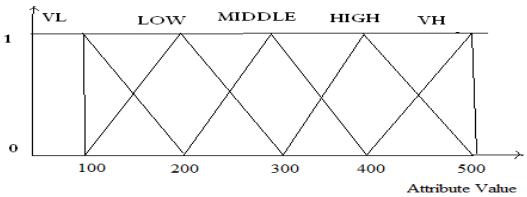

Fig. 5. Membership Function for Attribute $A_1$

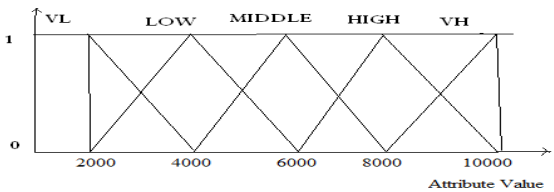

Fig. 6. Membership Function for Attribute $A_1$

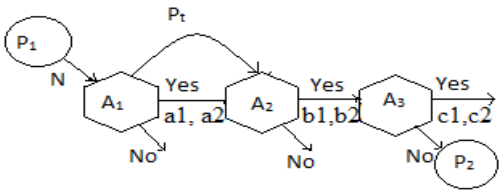


Fig. 7. Node Transition in Hybrid Rule Mining based on Fuzzy GNP

of discrete attribute $A_2$, and the third judgment node examines the value of symbol attribute $A_3$. $N$ is the number of total tuples. $P_t$ is the probability moving to Yes-side. When random value selected in (0, 1) is smaller than the certain probability $P_t$, GNP selects Yes-side and goes to the next judgment node. Probability $P_t$ is the fuzzy value of corresponding attribute. Otherwise, transition starts from next processing node to find new rule.

### *4.3 Extraction of rules using GNP with Fuzzy Membership Function*

The training dataset contains both normal connections and intrusion connections for misuse and anomaly detection. Initially, GNP examines all tuples of connections in database. In fig. 4, N is total number of tuples in the database. a1, b1, and c1 are number of tuples moving to yes side at judgment nodes with normal class $C = 1$ and a2, b2 and c2 are those with intrusion class $C = 2$. Criteria given in table are taken into consideration to pick up the rules to be stored in two rule pool namely normal rule pool and intrusion rule pool.

### *4.4 Updation of rules in rule pool*

Strong and robust rules are extracted by GNP are stored in the rule pool with its support, confidence, $\chi^2$ value with fuzzy parameter of fuzzy membership function. Fuzzy rule which is already in rule pool may extract again. In that case, membership functions and $\chi^2$ value might be changed. Fuzzy rule with higher $\chi^2$ value replaces remaining rule. Its fuzzy parameters adjusted using Michalewicz's operator . Therefore, the pool is updated every generation and only important fuzzy hybrid rules are stored.

### *4.5 Fitness and Genetic Operations*

Fitness function of GNP individual as

$$F = \sum_{r \in R} \left\{ \chi^2(r) + (1 - \mu_R(r)) + n(r) \right\} \qquad (6)$$

where, $R$ is set of suffixes of extracted important rules; $\chi^2(r)$ is $\chi^2$ value of rule $r$; $n(r)$ is the number of attributes in the antecedent of the rule $r$. $\mu_R(r)$ is crisp set containing all rules in rule set called $R$ denoted as $R = \{ r_1, r_2, \ldots, r_m \}$ where $m$ is total number of rules.

$$\mu_R(r) = \begin{cases} 1 & if\_ r \in R \\ 0 & if\_ r \notin R \end{cases} \qquad (7)$$

Every generation, individuls are replaced with the new one by following genetic operations. Selection, crossover and mutation are three kinds of genetic operands.

(1) Selection: In this method, rules are extracted on the basis of fitness function.
(2) Crossover: Two parents exchange selected nodes and their connections with probability $P_c$ generate two new offspring.
(3) Mutation-1: Each node branch is selected from one individual with probability $P_{m1}$ generate new individual.
(4) Mutation-2: Each node function is selected from one individual with probability $P_{m2}$ generate new individual.

### *4.6 Michalewicz's Non-uniform Mutation Operator*

Each continues attribute has its own fuzzy membership function. The parameter *x1*, *x5* for each continuous attribute $A_i$ is initialized by analyzing the distribution of data. During the whole evolution process, the parameters *x1, x2, x3, x4 and x5* of fuzzy membership function should be adjusted generation by generation, so as to get more appropriate distribution of parameters. From the second generation the parameter are selected by non-uniform mutation with probability of $P_m$. Since this part is based on real coding scheme, Michalewicz's non-uniform mutation operator [21] can be used. Let $x_k$ be a parameter selected for mutation in the *k*th generation, then next generation *t* as follows.

$$x_t = \begin{cases} x_k + \Delta(k, UB - x_k), & when \quad \varepsilon is\ 0; \\ x_k - \Delta(k, x_k - LB), & when \quad \varepsilon is\ 1; \end{cases} \qquad (8)$$

where, *UB* and *LB* are the lower and upper bounds of the variable $x_k$ and ε is a random binary values in {0,1}. The function $\Delta(k, y)$ returns value in (0,y), where $\Delta(k, y)$ approaches 0 as *k* increases.
Such property causes the operator to search the space uniformly at first and very locally in the later generation. The actual $\Delta(k, y)$ can be calculated by the following equation [17]:

$$\Delta(t, y) = y(1 - r^{(1-k/T)^b}) \qquad (9)$$

where, *r* is a unifrom random number in (0,1), *T* is the maximal generation number and *b* is a system parameter determining the degree of dependency on the iteration number. Different *b* means different non-uniformality in algorithm.

## 5. INTRUSION DETECTION WITH PROBABILISTIC CLASSIFICATION

Classifier is constructed to classify new connection data into normal, misuse and anomaly intrusion correctly. Classification is done after extraction of important class association rule including normal and intrusion.

### *5.1 Probability Distribution Function*

Probability distribution function is constructed according to the matching probability of all data with rules of normal and intrusion. The matching probability of data with a rule is defined as follows.

$$Match\_\Pr_{k,r}(d) = \frac{N_{k,r}(d)}{N_{k,r}} \qquad (10)$$

where, $N_{k,r}(d)$ is the number of matched attributes of data *d* with antecedent part of rule *r* in class *k*. $N_{k,r}$ is the number of attributes in the antecedent part of rule *r*.
Then, the average matching probability of the data with all the rules in class *k* is calculated as follows.

$$Avg\_Match_k(d) = \frac{1}{R_k} \sum_{r \in R_k} Match\_\Pr_{k,r}(d) \qquad (11)$$

where, $R_k$ shows set of rules in class *k*.

Finally, average probability distribution function can be created by distribution of average matching probability of all the training data $d \in D_{train}$ with the rules $r \in R_k$, where $D_{train}$ is set of training data.

### *5.2 Classification using Probability Distribution Function*

The probability that $P_k(d)$ that new connection data i.e. testing data belongs to class *k* can be calculated using average probability distribution function. The probabilities of belonging normal class *k* = 1, known intrusion class *k* = 2 and unknown intrusion class *k* = 0 are calculate by Eq. (12), (13) and (14) as follows.

$$P_{k=1} = (1 - Avg\_Match_2(d)) \times (Avg\_Match_1(d)) \qquad (12)$$

$$P_{k=2} = (Avg\ \_\ Match\ _2(d)) \times (1 - Avg\ \_\ Match\ _1(d)) \quad (13)$$

$$P_{k=0} = 1 - \sum_{k \in C} P_k(d) \quad (14)$$

where, $C$ = {1,….., $k$,….$K$} is set of classes. $K$=2 is used in this paper. Based on calculation of these probabilities, $d$ is assigned to class with highest probability. Positive false rate (PFR) increases when normal data labeled as intrusion and negative false rate (NFR) increases when intruded data labeled as normal. In order to balance PFR and NFR, medication of probability $P_k$ ($d$) is introduced.

Fig. 8 show flowchart for proposed method. Proposed method overcomes crisp boundary problem and deals with multi-data type database. Time required with comparison of conventional class association rule mining method is less as we are using K-means clustering algorithm for filtration purpose.

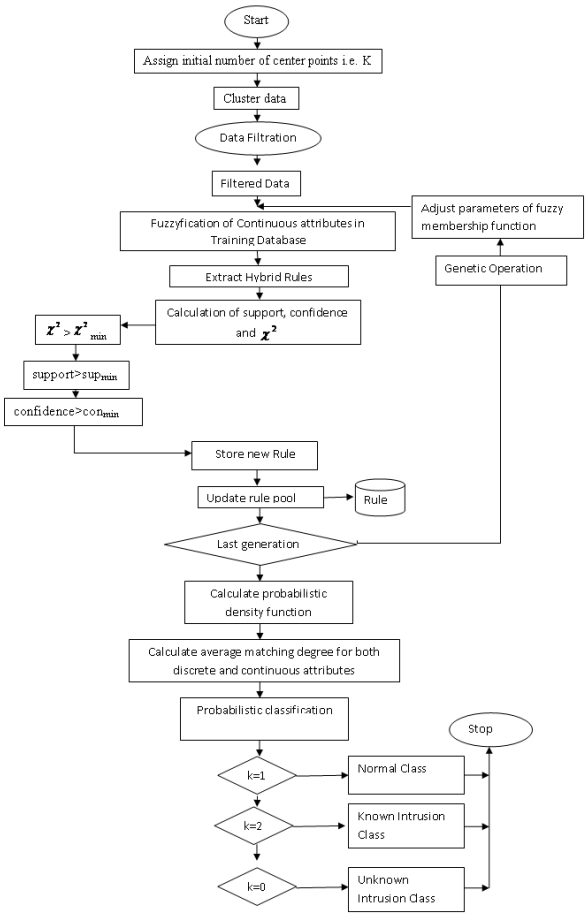


Fig. 8 Flow Chart of Proposed Methodology

## 6. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

This section describes the experimental results and performance evaluation of proposed system. The performance of system is evaluated accuracy, detection rate and false alarm rate. For experimental evaluation, we have taken KDD99CUP dataset. The parameters setting for proposed system is given table 2.

Table 2. Parameter Setting

| Parameter | Value |
|---|---|

| | |
|---|---|
| Population Size | 120 |
| Generation | 1000 |
| Crossover Rate | 1/5 |
| Mutation-1 Rate | 1/3 |
| Mutation-2 Rate | 1/3 |
| $\chi^2$ | 6.63 |

### *6.1 Comparison with previous finding*

This section describes comparison with some existing intrusion system. Table 3 gives comparison with KNN and KNN-DS theory [3] system on the basis of accuracy parameters

Table 3. Comparison with KNN and KNN-DS Theory

| System | Accuracy |
|---|---|
| KNN | 89.9 |
| KNN-DS | 95.2 |
| k-means+ Hybrid rule mining | 99.1 |

Table 4 gives comparison of proposed system with NB [5], KM-NB [5] using accuracy, detection rate and false alarm rate.

Table 4. Comparison with NB and KM-NB

| Method | NB | KM-NB | k-mean + Hybrid rule Mining |
|---|---|---|---|
| Accuarcy | 83.19 | 99.6 | 99.65 |
| Detection Rate | 94.7 | 99.8 | 98.11 |
| False Alarm | 19 | 0.5 | 0.3 |

Table 5 gives comparison with K-mean [4], Pre-M [4] and MDKM [4] using detection rate and false alarm rate.

Table 5. Comparison with K-mean, Pre-M, MDKM

| | K-Mean | Pre-M | MDKM | Proposed |
|---|---|---|---|---|
| Detection | 95 | 96.1 | 96.8 | 97.3 |
| FalseAlarm | 3.8 | 1.9 | 1.7 | 0.7 |

## 7. CONCLUSION

With the increasing use of internet, the security threats have multiplied many folds. Along with all conventional methods, intrusion detection system has come a long way to fight against security vulnerabilities. The proposed hybrid algorithm is combination of both K-means algorithm and probabilistic classification using fuzzy GNP hybrid rule mining. K-means algorithm is basically used for filtration of training database after clustering. Filtration of database is on basis of relevancy of data contained in each cluster. After processing of clustering and filtration data is passed for further process i.e. rule extraction and classification. Hybrid rule mining algorithm based on GNP combines Fuzzy-based class association rule mining and probabilistic classification in order to extract more important rules from the database. The use of genetic algorithm in intrusion detection system is particularly useful as it considers both temporal and special information. Moreover use of fuzzy logic can help in detecting anomalies which cannot be discretely deemed as normal and anomalous. The proposed algorithm is time efficient compared with existing system giving the advantage of formation of robust rules. Accuracy, detection rate and false alarm rate can be increased by using proposed algorithm.

**References**

[1] Simon Edwards, "Network Intrusion Detection Systems: Important IDS Network Security Vulnerability", Technical Evangelist, September 2002.
[2] Ramesh Agrawal, Mahesh Joshi, " PN-Rule: A New Framework for Learning Classifier Models in Data Mining", Technical Report, Department of Computer Science and Engineering, March 02, 2000.
[3] Deepika Dave and Prof. VineetRichhariya, " Intrusion Detection with KNN Classification and DS-Theory", International Journal of Computer Science and Information Technology and Security, ISSN:2249-9555, Vol.2, No.2, April 2012.
[4] LI Han, "Using A Dynamic K-Means Algorithm to Detect Anomaly Activities", 7th International Conference on Computational Intelligence and Security, 2011.
[5] Z.Muda, W. Yassin, M.N. Sulaiman, N.I. Udzir, "Intrusion Detection based on K-Means Clustering and Naïve Bayes Classification", 7th International Conference on IT in Asia (CITA), 2011.
[6]KDD1999data[online].Available: kss.ics.uci.edu/databases/kddcup99/kddcup99.html
[7] TheodorosLappas and KonstantinosPelechrinis, " Data Mining Techniques for (Network) intrusion Detection Systems", Department of Computer Science and Engineering,2007.
[8] Reema Patel, AmitThakkar, AmitGanatra, " A Servey and Comparative Analysis of Data Mining Techniques for Network Intrusion Detection Systems", International Journal of Soft Computing and Engineering, ISSN:2231- 2307, Vol. 2, Issue-1, March 2012.
[9] Satinder Pal Singh, "Data Clustering using K-Mean Algorithm for Network Intrusion Detection", Master's Thesis, Dept of Computer Science and Engineering, Lovely Professional University, Jalandhar, May 2010.
[10]Vivek K Kshirsagar ,Sonali M. Tidke and Swati Vishnu, "Intrusion Detection System using Genetic Algorithm and Data Mining: An Overview", International Journal of Computer Science and Informatics, 2231 –5292, Vol-1, Iss-4, 2012.
[11] Wei Lu and IssaTraore, "Detecting New Forms of Network Intrusion using Genetic Programming", Computational Intelligence, Volume 20, 2004.
[12] D.E. Goldberg, Genetic Programming in Search, Optimization, and Machine Learning.
[13] J.R. Koza, Genetic Programming on the Programming of Computer by Means of Natural Selection.
[14] Kaoru Shimada and Kotaro Hirasawa, "Exceptional Association Rule Mining using Genenetic Network programming", In proceeding of: Proceedings of The 2008 International Conference on Data Mining, DMIN 2008, July 14-17, 2008
[15] Eloy Gonzales, Bun TheangOng, Koji Zettsu, "Optimized Class Association Rule Mining using Genetic Network Progeamming with Automation Termination", The second International Conference on Advance in Information Mining and Management, 2012.
[16] J. G.-P. A. El Semaray, J. Edmonds, and M. Papa, "Applying data mining of fuzzy association rules to network intrusion detection," presented at the IEEE Workshop Inf., United States Military Academy, West Point, NY, 2006.
[17] Sergey Brin, Rajeev Motwani, Craig Silverstein, "Beyond Market Baskets: Generalizing Association Rules to Correlations", acm 0-89791-911-4/97/0005, 1997.
[18] T. Anitha Devi and K. RubaSoundar, "An Efficient Model for Network Intrusion Detection System based on an Evolutionary Computational Intelligence Approach", International Conference on Recent Trends in Computational Methods, Communication and Controls, 2012.
[19] Eloy Gonzales, Shingo Mabu, Karla Taboada, Kaoru Shimada and KataroHirasawa, " Class Association Rule Mining with Chi-squared Correlation Measures using Genetic Network programming", ICROS-SICE International Joint Conference 2009, August 18-21, 2009.
[20] Flora S. Tsai, "Network Intrusion Detection using Association Rules",International Journal of Recent Trends in Engineering, Vol.2 No.2, November 2009.
[21] O. Cordon, F. Herrera, L. Megdalena, P. Villar, "A Genetic learning process for the scaling factors, granularity and context of the fuzzy rule based system data base", Information Sciences, Volume 136, Issues 1–4, Pages 85-107, August 2001.