Intrusion Detection using Classification and Feature Selection of NSL-KDD data sets

Roshni Suryawanshi[#1], Prakash Kashyap[#2], Santosh Kushwaha[#3]

1Department of CSE, Rajiv Gandhi Prodyogiki Vishwavidyalaya, Bhopal, India
1yogeshlnct65@gmail.com
3santosh4mf@gmail.com
roshni.sist@gmail.com

Abstract— **Intrusion Detection System (IDS) is a valuable tool for the defense-in-depth of computer networks. However, Intrusion detection systems faces a number of challenges.One of the important challenge is that, the input data to be classified is in a high dimension feature space. In this paper, we effectively proposed PSO-DT(Iterative Particle Swarm Optimization-Decision Tree) intrusion detection system. Where, Particle Swarm Optimization (PSO) is used as a feature selection algorithm to maximize the CART (Classification And Regression Trees) Decision Tree classifier detection accuracy and minimize the timing speed. To evaluate the performance of the proposed IPSO-DT IDS several experiments on NSL-KDD benchmark network intrusion detection dataset are conducted. The results obtained demonstrate the effectiveness of reducing the number of features from 41 to 11, which leads to increase the detection performance.**

Keywords— **Wireless Sensor Networks, Clustering, Energy Efficiency, Stable Election, Network Lifetime.**

I. INTRODUCTION

An intrusion detection system (IDS) is a device or software application that monitors a network or systems for malicious activity or policy violations. Any detected activity or violation is typically reported either to an administrator or collected centrally using a security information and event management (SIEM) system. A SIEM system combines outputs from multiple sources, and uses alarm filtering techniques to distinguish malicious activity from false alarms.

There is a wide spectrum of IDS, varying from antivirus software to hierarchical systems that monitor the traffic of an entire backbone network.[citation needed] The most common classifications are network intrusion detection systems (NIDS) and host-based intrusion detection systems (HIDS). A system that monitors important operating system files is an example of a HIDS, while a system that analyzes incoming network traffic is an example of a NIDS. It is also possible to classify IDS by detection approach: the most well-known variants are signature-based detection (recognizing bad patterns, such as malware) and anomaly-based detection (detecting deviations from a model of "good" traffic, which often relies on machine learning). Some IDS have the ability to respond to detected intrusions. Systems with response capabilities are typically referred to as an intrusion prevention system.An IDS is relatively new technology for intrusion detection methods that have emerged in recent years. Intrusion Detection System's (IDS's) main role in a network is to help computer systems to prepare and deal with the network attacks. IDS collecting knowledge, from several different sources and compares this information with pre-existing patterns of differentiation as to whether there are assaults or weaknesses. Means example coordinating must be in precise way generally discover characterization may not be exact and the outcomes will be influenced. Lately, numerous specialists are centering to make use of information digging ideas for Intrusion Detection [1]. This is a procedure to remove the understood data and information. Interruption identification is the procedure of noxious assault in the framework and system when we are presently correspondence or deductive information in the constant environment [2][3]. Since its creation, interruption discovery has been one of the key components in accomplishing data security. It goes about as the second-line protection which supplements the entrance controls. It manage all through the procedure [3][4]. IDS manage administering the episodes happening in PC framework or system situations and inspecting them for indications of conceivable occasions, which are encroachment or inevitable dangers to PC security, or standard security hones (IDS) have developed to identify activities which imperil the trustworthiness, secrecy or accessibility of all sources as a push to give an answer for existing security issues [5].

1.1.1 Intrusion detection functions include:

1. Monitoring and Analyzing the possible threats..
2. Analyzing framework setups and vulnerabilities.
3. Assessing framework and record it.
4. Ability to perceive assaulted examples.
5. Analysis of strange action designs.
6. Tracking client strategy infringement.

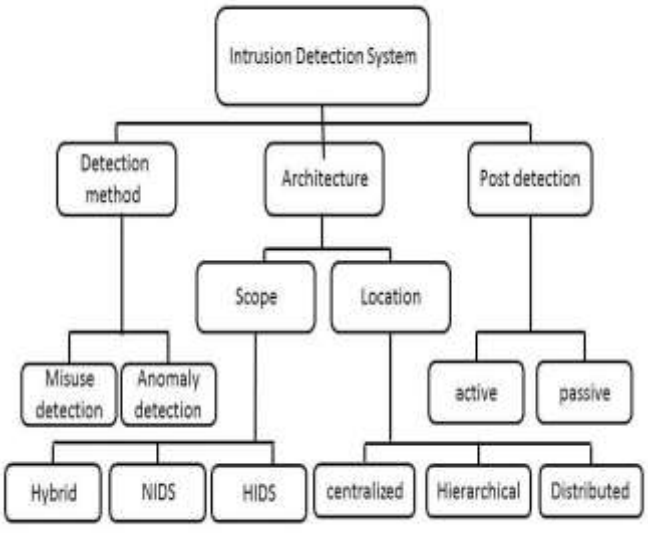


Figure 1.1: IDSs Classification Dimension

II. BACKGROUND

In 2014, Kebina Manandhar et al. [1] Detection of Faults and Attacks Including False Data Injection Attack in Smart Grid Using Kalman Filter. Author present s,Kalman filter generates estimates for state variables using the mathematical model for the power grid and the data obtained from the sensor network deployed to monitor the power grid. A $\chi 2$-detector can then be employed to detect the discrepancies

between the estimated data and the measured data, and trigger alarms. The $\chi 2$-detector can effectively detect attacks, such as the DoS attack and random attack, even though the states of the system do not remain constant at various time periods. However, the study shows that the $\chi 2$-detector cannot detect the statistically derived False Data Injection attack.

In 2012Mohammad Sazzadul et al. [2] prescribe An Implementation Of Intrusion Detectionnsystem Using Genetic Algorithm: Author's present an Intrusion Detection System (IDS), by applying genetic algorithm (GA) to efficiently detect various types of network intrusions. Parameters and evolution processes for GA are discussed in details and implemented. This approach uses evolution theory to information evolution in order to filter the traffic data and thus reduce the complexity. To implement and measure the performance of our system we used the KDD99 benchmark dataset and obtained reasonable detection rate. GA for solving various problems three factors will have vital impact on the effectiveness of the algorithm and also of the applications [19]. They are: i) the fitness function; ii) the representation of individuals; and iii) the GA parameters. The determination of these factors often depends on applications and/or implementation.

In 2014, Gideon Creech et al. [3] proposed A Semantic Approach to Host-based Intrusion Detection Systems Using Contiguous and Discontiguous System Call Patterns: Authors present new host-based anomaly intrusion detection methodology using discontiguous system call patterns, in an attempt to increase detection rates whilst reducing false alarm rates. The key concept is to apply a semantic structure to kernel level system calls in order to reflect intrinsic activities hidden in high-level programming languages, which can help understand program anomaly behaviour. Excellent results were demonstrated using a variety of decision engines, evaluating the KDD98. The ELM methodology has been verified as applicable to the IDS problem, with potential synergies uncovered due to the rapidity of decision engine training possible using this scheme. Portability between different versions of the same operating system has been investigated, and promising results suggest that the semantic approach introduced by this paper is extremely applicable to the task.

In 2014, Vahid Golmah et al. [4] proposed An Efficient Hybrid Intrusion Detection System based on C5.0 and SVM :Authors develop a hybrid method of C5.0 and SVM and investigate and evaluate the performance of our proposed method with DARPA dataset. The motivation for using the hybrid approach is to improve the accuracy of the intrusion detection system when compared to using individual SVM and individual SVM. The data set is first passed through the C5.0 and node information is generated. Node information is determined according to the rules generated by the C5.0. All the data set records are assigned to one of the terminal nodes, which represent the particular class or subset. This node information (as an additional attribute) along with the original set of attributes is passed through the SVM to obtain the final output. The key idea here is to investigate whether the node information provided by the C5.0 will improve the performance of the SVM.

In 2013, V. Jaiganesh et al. [5] Mangayarkarasi proposed a back-propagation approach to detect intrusion :First the input and its corresponding target are called a Training Pair is generated. Then the training pair is applied to the network. Detection rate and false alarm rate are the performance measure used for evaluation of proposed method. The detection rate for DoS, Probe, U2R, R2L attack is below 80%. Poor detection of attackers if some hidden attackers are present is one of the issues.

In 2011 Sufyan T. et al. worked on anamoly based intrusion detection in [6]: They have developed anamoly based IDS based on BPN and used packet behaviour parameter for experiment. The proposed model first detects normal-abnormal traffic then abnormal events are classified into four attack typtes (DOS, PROB, U2R, or R2L) and then detailed classification of abnormal events into 29 subattack types . 22 features of KDD99 dataset is used for experiment. 5 preliminary, 7 secondary,

10 less important features are categorized. They faced several issues which are as follows:Large amount of training data requires to train ANN an to get accurate results.There is little compromise between increasing the classification levels and the percentage of detection

In 2016, Ibraim M. Ahmed [7] Enhancement of Network Attack Classification using Particle Swarm Optimization and Multi Layer-Perceptron:

In this study, a proposed system has been developed that achieves classification technique by using hybrid soft computing technique which is Multi Layer-Perceptron (MLP) with Particle Swarm Optimization (PSO). The PSO has been used to improve the learning capability of the MLP by setting up the linkage weights in an attempt to enhance classification accuracy of the MLP. Simulation results conducted over three forms of experiments show that the proposed system gives high classification compared with other methods. The experiments results show that the proposed system gives high classification result compared with other methods. The results also show that increasing the numbers of practices and numbers in the training step will enhance the accuracy of classification. Future suggestion can be in the design and development of the proposed network intrusion system on a real time environment.

In 2015, MAHMOD S. MAHMOD et al. [8] Hybrid Intrusion Detection System Using Artificial Bee Colony Algorithm and Multi-Layer Perceptron:

In this study, hybrid Artificial Bee Colony (ABC) algorithm and Multi-layer Perceptron (MLP) were proposed to build an efficient network IDS. The MLP was utilized as a classifier to distinguish the normal and abnormal packets in the network traffic. The structure of MLP has been created relying on the features of (NSL-KDD 99) dataset. In addition, ABC algorithm is employed for training MLP by optimizing the values of linkage weights and bias. Training and Testing were performed by means of using NSL-KDD Dataset, which is the improved version of KDD99 dataset. The experiments results showed that the proposed method provides a high detection accuracy which is about (87.54%) and with (0.124%) error rate.

In 2014, Ralf C. Staudemeyer et al. [10] presents a Extracting salient features for network intrusion detection using machine learning methods:

This process is supported by detailed visualisation and examination of class distributions. Distribution histograms, scatter plots and information gain are presented as supportive feature reduction tools. The feature reduction process applied is based on decision tree pruning and backward elimination. This paper starts with an analysis of the KDD Cup '99 datasets and their potential for feature reduction. The dataset consists of connection records with 41 features whose relevance for intrusion detection are not clear. All traffic is either classified 'normal' or into the four attack types denial-of-service, network probe, remote-to-local or user-to-root. Using our custom feature selection process, we show how we can significantly reduce the number features in the dataset to a few salient features. We conclude by presenting minimal sets with 4–8 salient features for two-class and multi-class categorisation for detecting intrusions, as well as for the detection of individual attack classes; the performance using a static classifier compares favourably to the performance using all features available. The suggested process is of general nature and can be applied to any similar dataset.

In 2012, Ahmed A. Elngaret al. [9] A Fast Accurate Network Intrusion Detection System:

In this paper we proposed a Fast accurate anomaly network intrusion detection system (PSO-DT). Where, PSO algorithm is used as a feature selection method and then classify the reduced data by C4.5 decision tree classifier. The NSL-KDD network intrusion benchmark was used for conducting several experiments for testing the effectiveness of the proposed PSO-DT network intrusion detection system. Also, a comparative study with applying GA feature selectionwith C4.5 decision tree classifier was accomplished.The results obtained showed the adequacy of the proposed PSO-DT IDS of reducing the number of features from 41to 11 which leads to enhance the detection performance to 99.17% and decreasing the timing speed to 11.65 sec.

Table 2.1 summary of Different work based on technique used and the issue related to these work,

S. No.	Paper	Title	Technique Used	Issues
1	An Implementation Of Intrusion Detectionn system Using Genetic Algorithm	Mohammad Sazzadul et al.	applying genetic algorithm (GA) to efficiently detect various types of network intrusions. used the KDD99 benchmark dataset and obtained reasonable detection rate.	Doesn't provides efficient results for large datasets.
2	A Semantic Approach to Host-based Intrusion Detection Systems Using Contiguous and Discontiguous System Call Patterns	Gideon Creech et al.	apply a semantic structure to kernel level system calls in order to reflect intrinsic activities hidden in high-level programming languages which can help understand program anomaly behaviour	Can be applied for other domains such as warehousing.
3	An Efficient Hybrid Intrusion Detection System based on C5.0 and SVM	Vahid Golmah et al	a hybrid method of C5.0 and SVM and investigate and evaluate the performance with DARPA dataset	data set records are assigned to one of the root nodes
4	a back-propagation approach to detect intrusion	V. Jaiganesh et al	First the input and its corresponding target are called a Training Pair is generated. Then the training pair is applied to the network. Detection rate and false alarm rate.	Can't be applied for missing attributes.
5	Enhancement of Network Attack Classification using Particle Swarm Optimization and Multi Layer-Perceptron	Ibraim M. Ahmed .	developed that achieves classification technique by using hybrid soft computing technique which is Multi Layer-Perceptron (MLP) with Particle Swarm Optimization (PSO)..	Learning approach and hence error rate needs to be reduced for better results.
6	a Extracting salient features for network intrusion detection using machine learning methods	Ralf C. Staudemeyer et al.	feature reduction process applied is based on decision tree pruning and backward elimination with an analysis of the KDD Cup '99 datasets and their potential for feature reduction.	Less efficient and contains more classified tree..
7	On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems	Jun Wang, Xu Hong, Rong-rong Ren, Tai-hang Li	a new methodology based on GFS and pairwise learning for the development of a robust and interpretable IDS. Concretely, this approach is based on the FARCHD algorithm,	Provides complex system for the detection of intrusioms.

III. OPTIMIZATION METHOD

Particle Swarm Optimization (PSO)

PSO is initialized with a group of random particles (solutions) and then searches for optima by updating generations. In every iteration, each particle is updated by following two "best" values. The first one is the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called pbest. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the population. This best value is a global best and called gbest. When a particle takes part of the population as its topological neighbors, the best value is a local best and is called lbest.

After finding the two best values, the particle updates its velocity and positions with following equation (a) and (b).

v[] = v[] + c1 * rand() * (pbest[] - present[]) + c2 * rand() * (gbest[] - present[]) (a)
present[] = persent[] + v[] (b)

v[] is the particle velocity, persent[] is the current particle (solution). pbest[] and gbest[] are defined as stated before. rand () is a random number between (0,1). c1, c2 are learning factors. usually c1 = c2 = 2.

Decision Tree (DT)

Decision tree learning uses a decision tree as a predictive model which maps observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves).

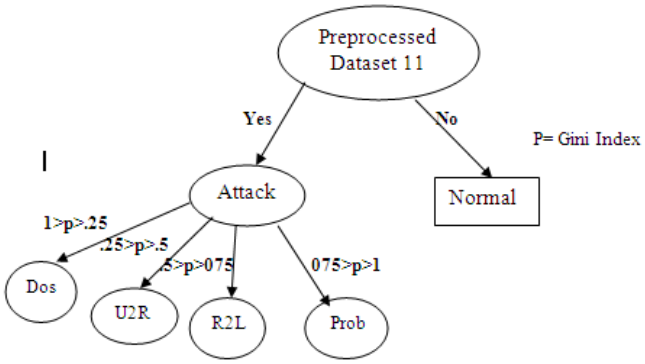
Classification and Regression Trees (CART)

CART algorithm (Breiman et. al., 1984) for finding the coefficients of the available features is a step-wise procedure, where in each step, one cycles through the features x1, x2, …, xf doing a search for an improved linear combination split. If there are symbolic features, they are converted to numeric features. Each instance is normalized by centering each value of each feature at its median and then dividing by its interquartile range. CART algorithm is very similar to ID3.

IV. PROPOSED WORK

4.1 Proposed Method

The proposed anomaly intrusion detection systemis using the advantages of PSO feature selection in conjunction with CART DT classifier to detect and classify the network intrusions into five outcomes: normal and four categories of intrusions. It consists of the following three fundamental building phases:
(1) Preprocessing ,
(2) Feature selection based PSO and
(3)Classification using CART DT. Figure 4.1 shows the overall architecture of the proposed IPSO-DT intrusion detection system.



CART tree Classification

Preprocessing phase: The following three pre-processing stages has been done on the NSL-KDD dataset:

1)Symbolic features are converted to numeric value.

There are three futures in each packets have characters values (protocol type, Service, Flag),which must converted to numeric value by compute number of time each feature is repeated, then ascending feature according to its repeated time, like 1 give to the feature have a greater number of repeated time, 2 for the feature have less frequently, … etc.

2) Each Attack name is converted to its category, 0 for Normal, 1 for DoS (Denial of service) , 2 for U2R (user-to-root), 3 for R2L (remote-to-local), and 4 for Probe

3) Normalization is implemented since the data have significantly varying resolution and ranges. The features values are scaled to be within the range [0, 1], using the following equation:

$$Xn = \frac{X - Xmin}{(Xmax - Xmin)} - 1 \qquad \text{(4.1)}$$

where, Xmin , Xmax are the minimum and maximum value of a specific feature. Xn is the normalized output.

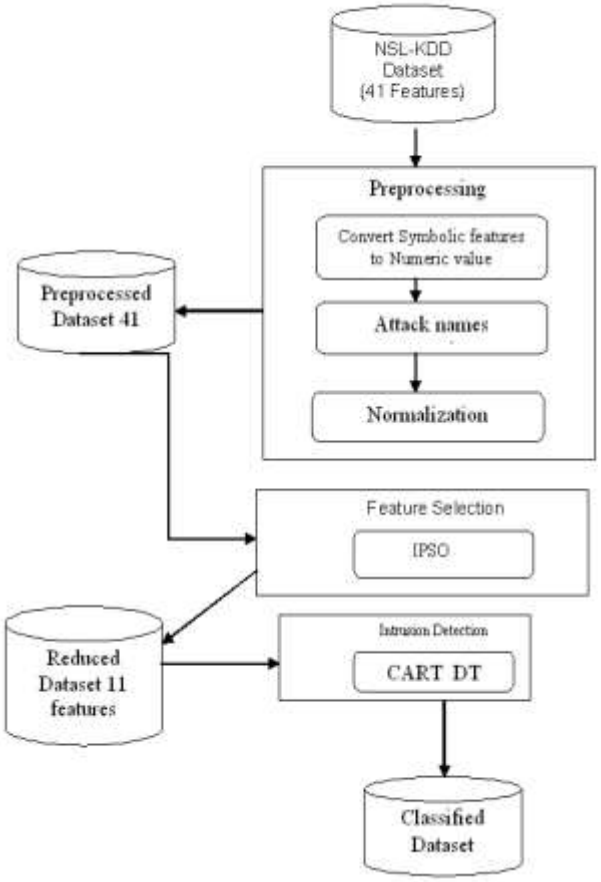


Figure 4.1. The overall architecture of the proposed IPSO-DT intrusion detection system

PSO Feature Selection Phase:

In this paper, PSO algorithm [23] has been used as a feature selection method to reduce the dimensionality of the NSL-KDD dataset. PSO efficiently reduces the NSL-KDD dataset from 41 features to 11 features, which reduces 73:1% of the feature dimension space. At every iteration of the PSO algorithm, each particle Xi is updated by the two best values pbest and gbest. Where, pbest denotes the best solution the particle Xi has achieved so far, and gbest denotes the global best position so far. Algorithm 2 shows the main steps of the PSO algorithmbased feature selection.

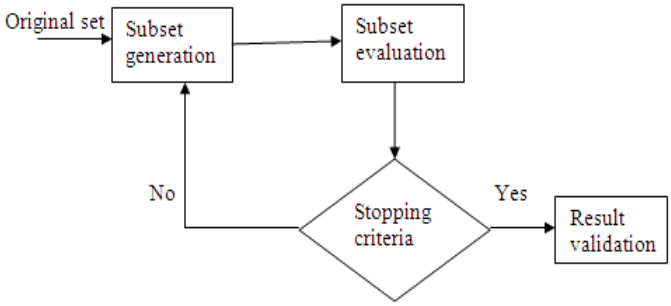


Figure 4.2 Main steps of feature selection process

Algorithm IPSO algorithm-based feature selection

Input: n: the swarm size.
a1 , a2 : positive acceleration constants. w: inertia weight.
MaxGen: maximum generation(selected dataset).
MaxFit: fitness threshold(classified features).
Output:
Global best position (best features of NSL-KDD dataset)

1: Initialize a population of particles with random positions and velocities on d=1,...,41
NSL- KDD features dimen-sions pbesti=0, Gbest=0, Iter=0.
2: while Iter < MaxGen or gbest < MaxFit do
3: for i = 1 to number of particles n do
4: Fitness (i) =Evaluate (i)
5: if fitness (i) > fitness (pbesti) then
6: fitness (pbesti) = fitness(i)
7: Update pid = xid
8: end if
9: if fitness(i) > Gbest then
10: Gbest = Fitness (i)
11: Update gbest = i
12: end if
13: for each dimension d do
14: Update the velocity vector.
15: Update the particle position.
16: end for
17: end for
18: Iter= Iter+1
19: end while
20: Return the Global best position.

CART-DT classification Phase:

A decision tree classifier is built using the CART algorithm [26].Then the reduced 11 features output from the PSO where passed to the CART decision tree classifier to be classified to one of the five categories: Normal, Dos, U2R,R2L and prob.

Gini purity of a node p(1-p)

- where p = relative frequency of data items.

Entropy of a node -Σplogp

- -[p*log(p) + (1-p)*log(1-p)]
- Max entropy/Gini when p=.5
- Min entropy/Gini when p=0 or 1

Performance evaluation

The detection effectiveness of the proposed PSO-DT IDS are measured in term of TP Rate, FP Rate and F-measure; which are calculated based on the confusion matrix. The confusion matrix is square matrix where columns correspond to the predicted class, while rows correspond to the actual classes. Table I gives the confusion matrix, which shows the four possible prediction outcomes [33]

Table. 4.2 Confusion Matrix

Actual Class	Normal	Attack
Normal	TN	FP
Attack	FN	TP

where,

True negatives (TN): indicates the number of normal events are successfully labeled as normal.

False positives (FP): refer to the number of normal events being predicted as attacks.

False negatives (FN): The number of attack events are incorrectly predicted as normal.

True positives (TP): The number of attack events are correctly predicted as attack.

$$TP\ Rate = \frac{TP}{TP+FN} \quad \text{.....(4.2)}$$

$$FP\ Rate = \frac{FP}{FP+TN} \quad \text{.....(4.3)}$$

$$F - measure = \frac{2*TP}{(2*TP)+FP+FN} \quad \text{.....(4.4)}$$

4.4 About the Dataset

NSL-KDD is a data set suggested to solve some of the inherent problems of the KDD'99 data set. Although, this new version of the KDD data set yet tolerate from some of the problems discussed by McHugh et al. and may not be a perfect representative of current real networks, because of the deficit of public data sets for network-based IDSs, we believe it yet can be applied as an imposing benchmark data set to auxiliary researchers comparison splits intrusion detection methods.

Table 4.1: NSL-KDD Dataset

class	Attacks in the training data
probe	Ipsweep, Nmap, Portsweep, Satan
DOS	Back, Land, Neptune, Pod, Smurf, Teardrop
U2R	Buffer_overflow, Loadmodule, Perl, Rootkit
R2L	Ftp_write, Guess_passwd, Imap, Multihop, Phf, Spy, Warezclient, Warezmaster

V.RESULT ANALYSIS

In our approach we are considering remaining normal data set for final pruning. For final pruning we consider those data which are not receiving normal. Then we find the Support Count Base on the field value which is shown in Table 5.1, Table 5.2 and Table 5.3. Based on the below tables we are arranging it in six different labels. It is T1,T2… T6 as shown in table 5.4. Then we apply the PSO Iterations on the associated terms find from the previous filtration. The support filtration is 50 %. If the value obtained by the suspicious node which qualifies the equivalent optimum threshold value then it will be added into the attack database. The final suspicious node database is then created.

Table 5.1: Features1 (10-22)

0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0

Table 5.2: Features2 (23-31)

1	1	0.00	0.00	1.00	1.000	0.01	0.06	0.00
1	1	0.00	0.00	0.00	0.000	1.00	0.00	0.4

Table 5.3: Features3 (32-41)

1	1	0.00	0.06	0.00	0.00	0.00	0.00	1.00	1.00
1	1	1.00	0.00	0.01	0.03	0.00	0.00	0.00	0.00

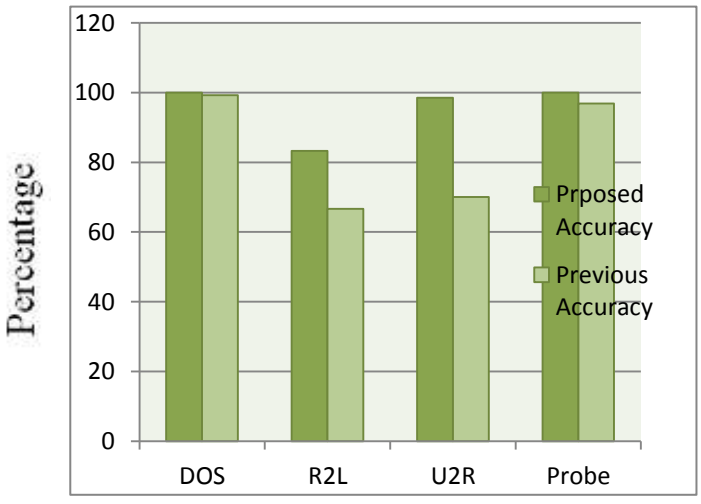
Table 5.5: Detection Result Total

Model	Accuracy
Proposed Approach	**95.5 %**
Fuzzy Ensemble	**93 %**
Random Forest [22]	**92.93 %**
JRip [23]	**92.30 %**
SVM [24]	**92.18 %**

Table 5.6: Attack Type Comparison

Attack Type	Proposed Accuracy	Previous Accuracy[33]
DOS	**100**	**99.25**
R2L	**83.3**	**66.66**
U2R	**98.5**	**70**
Probe	**100**	**96.88**

Graph 5.1 Classification of accuracy



VI. CONCLUSION

In this dissertation, we have applied data with associated items by the use of decision tree (CART) and Intrusion Detection Using iterative particle swarm optimization. For this we are using NSL-KDD dataset. In this approach the normal and attack nodes are separated first. Then it is examined for malicious behavior. For classification, we have used decision tree technique name CART, which is applied to form the associated items for the next preprocessing in the next pruning. We examine the feature selection value for the detection of different intrusion types defined. If the value obtained after iteration passed the threshold assigned, then it will be categorized as the specific intrusion and type will identified. Denial of Service (DoS), User to Root (U2R), Remote to User (R2L) and Probing (Probe) attacks is pondered in this dissertation for intrusion detection. The results show the improvement in detection as compared to the previous method.

VII. REFERENCES

1.Kebina Manandhar, Xiaojun Cao, Member, IEEE, Fei Hu, Member, IEEE, and Yao Liu, "Detection of Faults and Attacks Including False Data Injection Attack in Smart Grid Using Kalman Filter", Ieee Transactions On Control Of Network Systems, Vol. 1, No. 4, December 2014.

2.Mohammad Sazzadul Hoque1 , Md. Abdul Mukit2 and Md. Abu Naser Bikas3 ,"An Implementation Of Intrusion Detection System Using Genetic Algorithm", International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.2, March 2012 .

3.Gideon Creech, Student Member, IEEE, and Jiankun Hu†, Member, IEEE, "A Semantic Approach to Host-based Intrusion Detection Systems Using Contiguous and Discontiguous System Call Patterns", IEEE Transactions on Computers, 2014.

4.Vahid Golmah, "An Efficient Hybrid Intrusion Detection System based on C5.0 and SVM", International Journal of Database Theory and Application Vol.7, No.2 (2014), pp.59-70.

5.V. Jaiganesh, Dr. P. Sumathi, S. Mangayarkarasi,"An Analysis of Intrusion Detection System using back propagation neural network" IEEE Computer Society Publication -2013.

6.Sufyan T. Faraj Al-Janabi, Hadeel Amjed Saeed "A Neural Network Based Anomaly Intrusion Detection System" 2011 Developments in E-systems Engineering ,IEEE Publication - 978-0-7695-4593-6/11 , DOI 0.1109/DeSE.2011.19.

7.Ibraim M. Ahmed, "Enhancement of Network Attack Classification using Particle Swarm Optimization and Multi Layer-Perceptron", International Journal of Computer Applications (0975 – 8887) Volume 137 – No.12, March 2016.

8.Mahmod S. Mahmod, Zakaria A. Hamed Alnaish, Ismail Ahmed A. Al-Hadi, "Hybrid Intrusion Detection System Using Artificial Bee Colony Algorithm and Multi-Layer Perceptron", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 13, No. 2, February 2015.

9.Ahmed A. Elngar, Dowlat A. El A. Mohamed, Fayed F. M. Ghaleb, "A Fast Accurate Network Intrusion Detection System", (IJCSIS) International Journal of Computer Science and Information Security,Vol. 10, No. 9, September 2012

10.Ralf C. Staudemeyer_, Christian W. Omlin†, " Extracting salient features for network intrusion detection using machine learning methods", Research Article — SACJ, Submission, 2014.

11.Salma Elhag a , Alberto Fernández b,⇑ , Abdullah Bawakid c , Saleh Alshomrani c , Francisco Herrera c,d, "On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems", 0957-4174/ 2014 Elsevier Ltd. All rights reserved.

12.Meng Jianliang,Shang Haikun,Bian Ling,"The Application on Intrusion Detection Based on K-means Cluster Algorithm", International Forum on Information Technology and Applications, 2009.

13.Lundin, E. and Jonsson, E. "Survey of research in the intrusion detection area", Technical Report, Department of Computer Engineering, Chalmers University of Technology, Göteborg, Sweden. January 2002.

14.Li Tian,Wang Jianwen,"Research on Network Intrusion Detection System Based on Improved K-means Clustering Algorithm", International Forum on Computer Science-Technology and Applications, 2009.

15.S.Devaraju, S.Ramakrishnan:,"Analysis of Intrusion Detection System Using Various Neural Network classifiers, IEEE 2011.

16.Moriteru Ishida, Hiroki Takakura and Yasuo Okabe," High-Performance Intrusion Detection Using OptiGrid Clustering and Grid-based Labelling", IEEE/IPSJ International Symposium on Applications and the Internet, 2011.

17.Prakash Ranganathan, Juan Li, Kendall Nygard, "A Multiagent System using Associate Rule Mining (ARM), a collaborative filtering approach", IEEE 2010, pp- v7 574- 578.

