Journal of Theoretical and Applied Information Technology
15th February 2023. Vol.101. No 3
© 2023 Little Lion Scientific

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INVESTIGATION OF INTRUSION DETECTION SYSTEM USING RANDOM FOREST, CART AND PROPOSED SECURE RANDOM FOREST ALGORITHMS (SRFA)

P.KANAGAVALLI[#1], DR.V.KARTHIKEYANI[*2]

[1#1]*Research Scholar, Periyar University, Salem, Tamilnadu,India.*
[2]*Assistant Professor, Department of Computer Science, Government Arts and Science College, Komarapalayam,Tamilnadu,India.*

E-mail: [1]kavisekar7826@gmail.com, [2]drvkarthikeyani@gmail.com

ABSTRACT

Wireless Sensor Network (WSN) are composed of low cost sensor nodes and commonly deploy in open and unprotected area, which make security the principal mission in this form of network, due to their traits WSN is susceptible to various types of attacks and intrusions, the place it require security mechanisms to protect towards these attacks. Intrusion detection machine (IDS) is one of the major and efficient protecting strategies against intrusion and attacks in WSN. In this paper, a novel feature extraction algorithm, specifically Correlation Based Feature Extraction (CFS) algorithm is proposed with an aim to limit the training time and to decorate the lifetime of the system. The trust level node is estimated by using utilizing the behavior analysis and residual energy level of nodes. Thus, we have proposed a new Trust Algorithm (TA) to compute the trust level of nodes in the network. Finally, SRFA primarily based classifier is used to classify the nodes into a trustworthy, untrustworthy or malicious node based totally upon the measured trust stage of the nodes. The results absolutely confirmed that the proposed intrusion detection system extensively reduces the false positive rate, thereby proving that the proposed approach is capable of identifying anomalies in network better than different current system.

As cyber threats develop in sophistication, network defenders need to use every device in the defensive arsenal to defend their networks. Data mining techniques, such as decision tree analysis, provide a semi-automated approach to detect adversary threats. In this paper, a novel feature extraction algorithm, specifically Correlation Based Feature Extraction (CFS) algorithm is proposed with an intention to limit the training time and to beautify the lifetime of the system. The trust level node is estimated via using the behavior analysis and residual energy level of nodes. Thus, we have proposed a new Trust Algorithm (TA) to compute the trust level of nodes in the network. Finally, SRFA based classifier is used to classify the nodes into a trustworthy, untrustworthy or malicious node based totally upon the measured believe degree of the nodes. This proposed method is compared with SVM and C4.5 and evaluates the overall performance the usage of KDD99 dataset. Simulation results prove that the proposed SRFA can successfully mitigate malicious node and gives higher effects when in contrast to SVM and C4.5. In previous years, a dramatic enhancement issues in the number of attacks, intrusion detection field and it becomes the mainstream of data assurance. Sensors nodes are used in WSN with the onboard processors that manages and monitors the environment in the a particular area. They are connected to the Base Station which acts as a processing unit in the WSN System. In WSN, data mining is the process of extracting model and pattern that are application oriented with possible accuracy from rapid flow of data. Because of their special characteristics, and limitations of the WSNs, the traditional data mining approaches are not directly applicable to WSNs. A widespread analysis of different pre-existing data mining techniques adopted for WSNs are examined with different classification, evaluation approaches in this paper. Finally, a few research challenges to adopt data mining methods in WSNs are also pointed out. A general concept of how traditional data mining techniques are improved to attain better.

Keywords: *Decision Tree, Intrusion Detection System, SVM, C4.5, CART, Secure Random Forest Algorithms*

1. INTRODUCTION

This As wireless networks discover outstanding use in emergency situations, their security is a primary concern. It is notably possibly that unauthorized network activities known as intrusions can appear which absorbs resources intended for authorized users of the network. Hence the security of networks beneath attack is affected [1, 2]. This creates the want for Intrusion Detection Systems (IDSs) which identify malicious traffic and protects the networks from attacks. Data mining approach is newly used in intrusion detection. Data mining is well known for "Data Retrieval method that is retrieved from the massive series of data. It is used to retransform it into a statistically vast constructions and events in data. There are many exclusive kinds of facts mining strategies such as K-Means,ID3,NB Tree etc [3]. that has to keep track of classification, link analysis, clustering, association, rule abduction, deviation analysis, and sequence analysis. Data Mining offers an Intrusion Detection Model inclusive of these facts mining strategies via extracting know-how from the massive datasets and by examining them. Recently most of the advantageous intrusion detection systems remember on the machine learning system, where the machine learning mechanisms are very purposeful so that it offers excessive opportunity of detecting the intrusions in the network. The important purpose for the ability of the machine learning system in the detection of intrusions is that it uses helps vector machines, neural networks, and all these chances are based totally on decision trees that have environment friendly extensive schemes in anomaly detection structures [4]. Thus, it enhances the classification overall performance and accelerates the velocity of processes. However, the existing intrusion detection techniques, which are proposed via more than a few researchers for misuse and anomaly detection [5], are usually no longer ample to provide the required security to WSNs due to the fact they have restricted power and tiny structure.

The attacks are carefully designed with the aid of the attackers and subsequently the utility of the present intrusion detection methods causes a excessive false positive rate. Moreover, the existing Intrusion Detection methods are successful of detecting only known intrusions considering they classify situations by way of the policies they have received primarily based on coaching from past data. However, it is integral to construct wise IDS with effective learning abilities in order to secure the network from each inner and external attacks [6]. In this work we have proposed a new intrusion detection system based totally on trust mechanism. The proposed system consists of the three modules. The first module is responsible for the feature extraction process. Moreover, a novel algorithm Correlation Coefficient Based Feature Extraction (CFS) algorithm for enhancing the feature extraction process is proposed. The second module is responsible for trust computation process. Moreover, a new algorithm Trust Algorithm (TA) algorithm is proposed for the purpose of trust computation process. Moreover, by using the third module that utilizes the Secured Random Forest Algorithm (SRFA) based Classifier malicious nodes are identified and that deals about decision regarding the classified nodes in the network.

DECISION TREE ALGORITHMS FOR IDS

Decision tree induction algorithms have been used for classification in many software areas such as medicine, manufacturing, and production, Financial Analysis, astronomy, and molecular Biology. Decision tree are the fundamental of Several Commercial rule induction System [7]. Many decision tree algorithms have been proposed by way of more than a few researchers for high-quality decision making [8]. In this research work we will use CART and RF data mining based decision tree algorithm for comparative result.

RANDOM FOREST IN INTRUSION DETECTION SYSTEM

Random forest technique works on the rule divide – and – conquer scheme which is used in the classification mission. As it is an ensemble process, it amalgamates a crew of fragile learner to produce well-built leaner which can categorize the data precisely. The bagging scheme and random selection of aspects are united in it. N number or tresses are produced in random forests. Each tree represents regular and distinct malicious classes. A giant range of data sets are effortlessly managed via a random forest algorithm.

```
Algorithm: Random forest modeling for network
IDS
Input: NSL-
KDD dataset
Output:
Detection Rate
Step 1: Load
the dataset
```

Step 2: Apply pre-processing technique Discretization Step 3: Cluster the dataset into four datasets.

Step 4: Partition the data set into training and test

Step 5: Select the best set features using feature subset selection measure Symmetrical uncertainty (SU) Symmetrical uncertainty compensates information gain

$$SU(X, Y) = 2[IG(X/Y)/H(X) H(Y)]$$

Step 6: Data set is given to Random forest for training

Step 7: The test data set is then fed to random forest for classification

Step 8: Calculate accuracy, Detection rate, False alarm rate

In this section, experiments results analysis is discussed. All experiments were conducted using platform of Windows with configuration of Intel® core™ i7 CPU 2.70 GHZ, 8 GB RAM. KNIME tool was used to evaluate the method and perform feature selection. KNIME software [10] is a machine learning tool that consists of series of machine learning algorithms. During classification, the parameters of KNIME are set to its default values. All experiments are carried out on NSL-KDD datasets [11]. NSL-KDD is a refined version of the KDD'99 dataset. It overcomes some inherent problems in the original KDD dataset. Redundant records in the training set have been removed so that the classifiers produce unbiased results. There is no duplicate data in the improved testing set. Therefore, the biased influence on the performance of the learners has been significantly reduced. Each connection in this dataset contains 41 features and 2 classes, labeled as either attack or normal. The attacks in the dataset are grouped into DoS, Probe, R2L and U2R; with these attacks divided into training set and test set. Researchers in this work carry out the experiments using the KDDTrain and KDDTest data.

Random Forest algorithm has been used in many research articles for performing intrusion detection. In the works [9] and [10], a hybrid Network Intrusion Detection System (NIDS) using the combination of anomaly and misuse-based detection has been built using Random Forest algorithm.

CART in IDS

It is a binary recursive splitting method which can process nominal and continuous attributes as targets. The data is handled as such in its original form. Starting at the root node, data are partitioned to produce two children and each of them is partitioned to produce their children and so on. This process continues till no more splits are possible due to the lack of data. The tree is pruned backwards to the root by cost-complexity pruning. Next the split which contributes the least to performance on training data is pruned. CART generates a set of nested pruned trees and the optimal tree is selected by examining on testing data [11]. Many research articles have implemented CART algorithm for detecting intrusions in networks. In one such article [12], CART and fuzzy logic methods have been employed to find out intrusions using KDD Cup '99 dataset. The CART algorithm is described below.

Algorithm: CART for IDS **Input:** Clusters of data, **Output:** Decision tree- classified data
Step 1: Create a node N

Step 2: If tuples in clusters (C) are of same class M then return N as leaf node labeled with class M.

Step 3: Otherwise, apply Attribute selection method (Gini Index) to find the best splitting criterion.

Step 4: Label node N with the splitting criterion j.

Step 5: If splitting attribute is discrete-valued and multi way splits allowed then attribute list = attribute list - splitting attribute.

Step 6: Let Cj be the set of data of cluster that satisfy outcome of j. If Cj is empty then attach leaf labelled with majority class of C to node N Else repeat step 3.1 to 3.5.

//Partiton tuples and grow decision tree for each partition. Step 7: Return N.

Step 8: Classify the instance as normal or anomaly.

2. PROPOSED METHODOLOGY

In order to increase the detection ability of IDS and prevent the service providers from attack, we propose an efficient algorithm for IDS. This section provides the details about the proposed scheme. The proposed scheme mainly consists of three phase namely: (i) Feature Selection; (ii) Trust Value Computation; and (iii) Classification.

FEATURE SELECTION

The first step of the proposed approach is to select the features from the input data set. This step is important because it involves to identifying those features of the data that may trigger an alarm when an intrusion is suspected. In this paper, a Correlation-based feature selection (CFS) is proposed to optimize the efficiency of the feature selection process and enhance the accuracy of the classification. The main concept of this approach is to evaluate the relevance and the redundancy of the selected feature subset which is searched in the given search space for the optimal solution. CFS is one of classical filter algorithms that choose features according to the result of the heuristic (correlation-based) assessment function. The preference of this function is to select subsets whose features are extraordinarily related with the class but uncorrelated with each other. While insignificant features that show low association with the class ought to be ignored on the grounds, repetitive features are chosen due to high relation with at least one of the rest of features. The acknowledgment of a feature will rely upon the degree to which it predicts classes in territories of the instance space not as of now anticipated by different features. The feature subset assessment function in CFS is as:

In Eq. 1, M_s is the heuristic evaluation for a feature subset s including k features, $\overline{Tcf}$ is the mean correlation degree between features and the category label, and $\overline{Tff}$ is the average inter-correlation degree among features. The evaluation of CFS is a method of correlation based on feature subsets. A bigger $\overline{Tcf}$ or smaller $\overline{Tff}$ in acquired subsets by the method produce a higher evaluation value, and the set of features with the highest value found during the process is utilized to reduce the size of both the training and testing set. The feature selection process of the CFS approach is presented in Algorithm 1. The main parts of the CFS-algorithm can be summarized as follows:

Algorithm 1: CFS approach for feature selection **Input:** Training Dataset and Testing Dataset **Output:** Selected Feature Subset X_{best}

```
Initialize a population of n bats $X_i = (x_{i1},\dots,x_{iD})^T$ (i = 1, 2,…., n) and $v_i$
Initialize frequency $f_i$, pulse emission rate $r^t$ and loudness $^{t}_{i}$
Initialize $f_{it}(X_i)$ (cf. Eq.1) and $X_{best}$
Initialize $fit_{temp}(i)$ and $Xt_{emp}(i)$ for solution storage
while 1 ≤ t ≤ Max no. of iterations do
    for i = 1 to n do
        Generate new $f_i$ using
        Update $X_i$ and $v_i$ using
        if $r^t <$ rand (0,1) then
            Select a $X_i$ from $X_{best}$
            Generate a new $X_{new}$ using
        end if
        Calculate $f_{it}(X_{new})$ (cf. Eq.1)
        if $f_{it}(X_i) \leq f_{it}(X_{new})$ and $N(0,1) < A^t$ Then
            fit
            $X_{temp}(i) \leftarrow X_{new}$
            Decrease $A^t$ and Increase $r^t$ $using$
        end if
        if $fit(X_{new}) \geq$ Max of $fit_{temp}$ then
            $X_{best} \leftarrow X_{new}$
        end if
    end for
    t = t + 1
end while
```

TRUST VALUE COMPUTATION

In this section, the trust value of a node is estimated based upon the behaviour analysis and Residual Energy Level of nodes. For instance, a node say “A” estimates the trust value of another node “B” based on the behaviour analysis of a node A on node B as well as Residual Energy Level of B. In this work, utilized the following parameters as a trust factor to measure the trust of a node and to detect the malicious node from the network.

Algorithm 2: Trust Algorithm Input: Set of nodes n
Output: Set of trustworthy nodes

Packet Dropping Ratio = No of packets transmitted by the node / No of packets received by the node

Step 2: calculate the Packet Misrouting Ratio

Packet Misrouting Ratio = No of packets misrouted by the node / Total no of incoming packets

Step 3: calculate the Packet Falsely
Injected Ratio Packet Falsely Injected
Ratio =No of packets falsely injected by
the node / Total no of incoming packets
Step 4: calculate the Packet Altering
Ratio

Packet Altering Ratio = No of packets altered by the node / Total no of incoming packets

Step 5: calculate the Residual Energy Level and $E_{Threshold}$ Residual Energy Level = E_{curr} / $E_{initial}$

$$E_{Threshold} = \sum_{i=1} REL(n) / n$$

Step 6: if (PDR < d_{min}) and (PMR < m_{min}) and
(PFIR < f_{min}) and (PAR < a_{min}) and (REL <
$E_{Threshold}$)then
Node i is labelled as highly
trusted node Else

Step 7: if ($d_{min} \le$ PDR $\le d_{max}$) and ($m_{min} \le$
PMR $\le m_{max}$) and ($f_{min} \le$ PFIR $\le f_{max}$) and
($a_{min} \le$ PAR $\le a_{max}$) and
Step 8: if (PDR > d_{max}) and (PMR > m_{max})
and(PFIR > f_{max}) and (PAR > a_{max}) and
(REL > $E_{Threshold}$) then
Node is labelled as malicious node
Node i is isolated from the network
End if
End if
End if
End for

CLASSIFICATION

The final combined subset is used as an input to the classification stage. The main objective of this research is to develop an efficient intrusion detection system by improvement in existing methods. Random forest is a widely used method in NIDS. Random forest is one of the most successful classifiers. It is a classification algorithm obtained through extending the decision tree classifiers using ensemble learning techniques. The existing original random forest (RF) algorithm has some challenges in feature selection process, selection of classifiers, selection strategy for random features for various training and also challenges in combination steps. In this research paper we are presenting a network intrusion detection system based on Modified Random forest classifiers. Proposed Secured random forest algorithm (SRFA) is combination of unpruned classifiers and CART (regression tree) with bagging approach. From the selected features, SRFA selects the best features and built the decision tree, a sampling variable and confusion matrix are used for identification of data more accurately and efficiently.

Algorithm 3: SRFA
Input: Data set D with various entries. ($D_{trainging}$= {D_1, D_2……D_n} where $D_{trainging}$= training data set

Output: Resulting data set $D_{resulting}$, with better detection dare and accuracy

Step 1: Create decision tree and learning

- Retrieve data set and upload it on weka

Step 2: Tree bagging by bootstrap method

- Apply bagging or bootstrapping approach on the data set, it divides the data set into different subset of data set with replacement of rows
- Given a training set X = x_1, ..., x_n with responses Y
= y_1, ..., y_n,
- Bagging repeatedly (B times)
- Selects a random sample with replacement of the training set and fits trees to these samples: For b = 1, ..., B:Sample, with replacement, B training examples from X, Y;
- Call these X_b, Y_b.
- Train a decision or regression tree f_b on X_b, Y_b.
- After training, predictions for unseen samples x' can be made by averaging the predictions from all the individual regression trees on x':-
- Select by taking the majority vote in the case of decision trees

Step 3: from Bagging to Random Forest

- Check all the value in each field, column and for each attribute
- Find the best selection
- Applying Splitting criteria ( )
- Generates all the label nodes based on splitting criteria function
- Apply CART ()
- Measure accuracy for each of the Decision Tree

Step 4: Extra Trees

- Check the entire extra tree
- Match with new results

Step 5: Merging

- Attach a leaf labeled with the majority class in D to node N
- Attach the node return by generate decision tree (generate () ;) to node N;

Step 6: Return Tree_N

3. EXPERIMENTAL RESULT

In this section, an experiment has been conducted which shows the comparison of proposed algorithm with different classification algorithms on the same dataset as mentioned in the paper. The classification algorithms used in this paper are Random Forest (RF), Decision Tree (CART). Results show that proposed algorithm (SRFA) has better accuracy than other algorithms. All experiments are performed by using an Intel(R) CORE™ i5- 3210M CPU @ 2.50GHz, Installed 8 GB RAM and 32-bit Operating system. For all simulations, we trained and tested the performance of the classifiers on the NSL-KDD datasets by using knime tool. The Algorithms analysis is done by comparing their accuracy, precision, recall, True Positive Rate (TPR), False Positive Rate (FPR), detection rate and false alarm rate

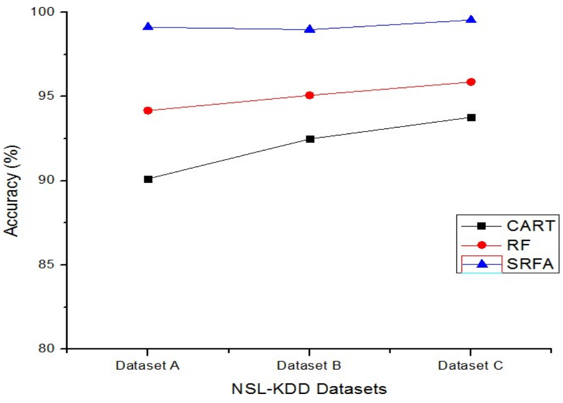

Figure 1: Comparison of Accuracy

Two benchmark datasets, KDDCup 1999 and NSL-KDD, are widely used for simulation of IDSs. The NSL-KDD dataset is an improved version of the KDDCup 1999 dataset where redundant and duplicated data is removed from the training and testing datasets. Every instance of this dataset has 41 features and is labeled as either normal or attack. The attack types fall into the following four classes:

- **Denial of Service (DoS) attacks:** In this type of attack, the intruder tries to keep the network busy by exploiting the bandwidth that results in denial of service for legitimate requests.
- **User to Root (U2R) attacks:** By logging in as a normal user, the intruder tries to access the system with root privilege.
- **Remote to Local (R2L) attacks:** The attacker tries to locate vulnerability to access the system remotely.
- **Probing:** The intruder tries to gather information about a network in order to bypassing its security policy

Accuracy

As shown in Figure 1, all tested classifiers achiev and above in terms of detection accuracy, thus indicating that most of the decision tree based IDS is properly detecting the attacks. Figure 1 also shows that SRFA has the highest detection accuracy rate at 99.54% and outperforms the other classification algorithms. The accuracy of any classifier depends on the type of data, sample size, and data dimension. The result concludes that RF outperforms the other classifier algorithms in term of accuracy detection might be because of a small data size with a small set of features was used. Ahigh number of feature sets can have negative effects on the accuracy of RF, thus performing a feature reduction before using the dataset is highly recommended to evaluate the accuracy of the RF algorithm. But in proposed (SRFA) algorithm produce good accuracy result both small and large size datasets.

False Positive Rate (FPR) and True Positive Rate (TPR)

The graphical representation of performance result is given below the figure 2 and figure 3 displays the TP rate and FP rate respectively three machine learning algorithm with NSL_KDD data set. It shows the true positive rate results using 41 attributes for training data compared to three algorithm SRFA (>99%) with other. And it shows the false positive rate results using 41 attributes for training data compared to other the SRFA algorithm only (<0.03 %).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

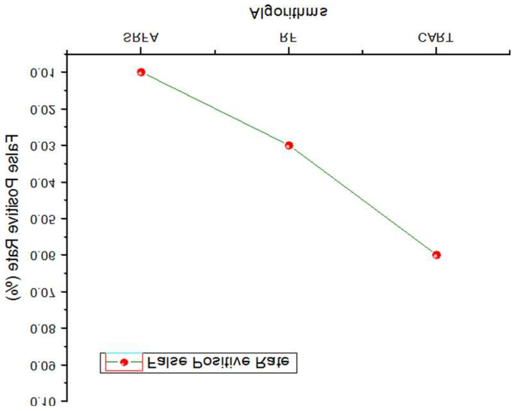

Figure 2: False Positive Rate Of The Proposed Algorithms

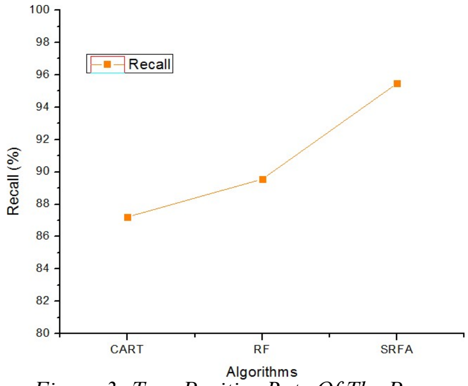

Figure 3: True Positive Rate Of The Proposed Algorithms

False Alarm Rate (FAR)

This metric calculates how often the model is predicting a positive result wrongly. It provides indication of possible error of the model, thus lower value is better. Figures 4 present a comparative study of false alarm rate of the proposed and existing method. The average false alarm rate achieved by proposed algorithm is 0.34 compared to others. For example, in figure 4 the false alarm rate for DOS is 2.59 in the work of CART and 1.02 for RF, but in the proposed work, the value is minimized to 0.27 and for PROBE, U2R, and R2L also the false alarm rate value has decreased drastically.

It eliminates the biasness of records gain when there are many consequence values of an attribute. At first, calculate the achieve ratio of every attribute. The root node will be the attribute whose acquire ratio is maximum. C4.5 makes use of pessimistic pruning to do away with useless branches in the selection tree to improve the accuracy of classification. Intrusion detection algorithm based on C4.5 can be divided into three levels.

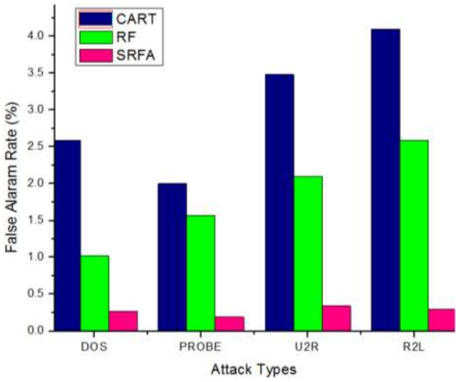

Figure 4: Performance Comparison On False Alarm Rate Of Proposed Work And CART, RF.

Precision and Recall

Precision is outlined because the proportion of positive predictions that created by the classifier that are true. The precision rate straight forwardly influences the performance of the system. The Recall rate are additionally a vital value for estimating the execution of the recognition (detection) system and to demonstrate

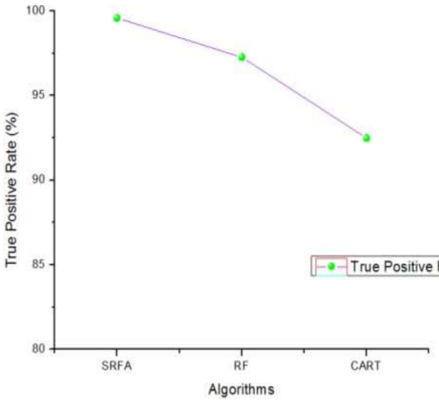

Figure 5, It Is Due To The Normal Traffic Is Big Enough To Be Recognized

the extent of occurrence shaving a place with the positive rate that are effectively anticipated as positive.. Along with the percent of training samples increased, the precision of each algorithm gets better. And the method of SRFA gets the better performance than others. Based on the recall, in figure 6 shows proposed algorithm have the highest recall of about 95.48%, whereas RF have a recall of 89.55%. CART had the lowest recall with 87.72%

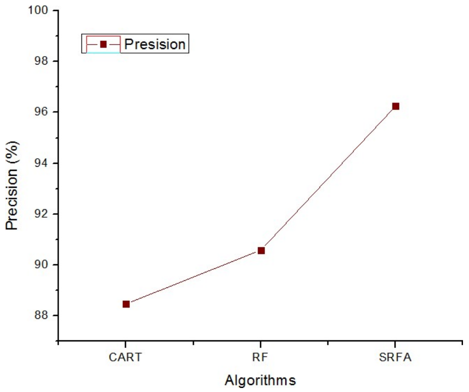

Figure 6: Comparison By Recall

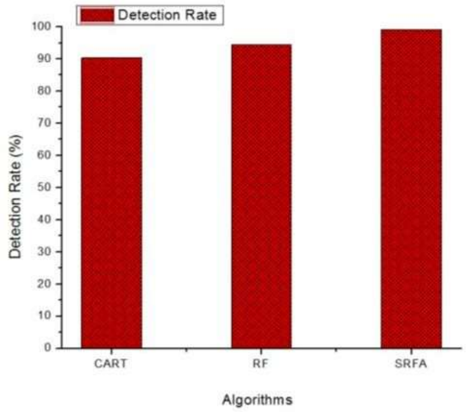

Figure 7: Compare Proposed Algorithm With CART And RF By Detection Rate

PERFORMANCE EVALUATION AND RESULTS

To evaluate our detection system, we applied the evaluation criteria as follows:

- **Detection Rate (DR):** it is the percentage of normal and attack data are classified correctly from the given number of total dataset records.
- **False Positive Rate (FPR):** it is when normal connections are incorrectly classified as intrusions or attacks.
- **True Positive Rate (TPR):** portion of intrusions properly diagnosed.
- **Precision:** is the fraction of retrieved documents that are relevant to the query

4. CONCLUSION

Because of the fast growth of the internet, many techniques are used to avoid intrusions. However, still, there is a need to develop more efficient systems that detect new or unique intrusions. There are different algorithms using which anomaly based intrusion detection can be performed. In this comparative analysis of study, Random Forest has very large classification accuracy but sometimes slow in predictions. It can be concluded that by examine these algorithms, the best-suited techniques according to the network scenarios can be chosen to design efficient Intrusion Detection Systems which can detect attacks with high accuracy. In this paper, proposed a new intrusion detection system using a Secured Random Forest Algorithm (SRFA) is combination of unpruned classifiers and CART

(regression tree) with bagging approach. The experimental results on KDD dataset proposed algorithm achieved high detection rate on different types of network attacks. From the observed results it can be concluded that the Secured Random Forest Algorithm (SRFA) classifier outperforms other classifiers for the considered data-set and parameters.

Intrusion detection is one of the main research problems in computer security. The main goal is to detect infrequent access or attacks to protect internal networks from attacks. In this paper, a novel detection approach called SRFA has been proposed. The proposed IDS reduce the complexity of the system by selecting important features in the datasets. From the simulation results the performance of the proposed system is analyzed with SVM and C4.5. Compared to the single algorithms, combining with multiple algorithms has given much better results. The proposed algorithm outperforms other existing approaches. Simulation results demonstrate that the proposed algorithm is successful in detecting misuse and anomaly intrusion detection system.

REFERENCES:

[1]. Joshua Muscatello,Joshuamartin,"Wireless Network Security",April,2005.

[2]. H. Chan And A. Perrig, "Security And Privacy In Sensor Networks," Computer, Vol. 36, Pp. 103–105, October 2003.

[3].C.K.Marigowda,Manjunathshingadi,"Security Vulnerability Issues Wireless Sensor Networks:Ashortsurvey",InternationalJournal Of Advance Research In Computer And Communication Engineering .Vol.2,Issue 7,July 2013.

[4]. Dr.G.Padmavathi,Mrs.D.Shanmugapriya,"A Survey Of Attacks,Security Mechanisms And Challenges In Wireless Sensor Networks"(IJCSIS) International Journal Of Computer Science And Information Security,Vol.4,NO.1 & 2,2009 .

[5]. Stefan Axelsson, "*Intrusion Detection Systems: A Survey and Taxonomy*", Technical Report No 9, Dept. of Computer Engineering, Chalmers, University of Technology, Sweden, pp. 9-15, 2000.

[6]. Denning D E, "*An Intrusion Detection Model*",. *IEEE Transactions on Software Engineering, Vol. 51, no. 8, pp. 12-26, Aug. 2003.*

[7]. Ektefa M., Memar S., "Intrusion Detection Using Data Mining Techniques",IEEE Trans., 2010.

[8]. Theodoros Lappas and KonstantinosPelechrinis, "Data Mining Techniques for (Network) Intrusion Detection Systems".

[9]. Zhang, Jiong, and Mohammad Zulkernine. "A hybrid network intrusion detection technique using random forests." The First International Conference on Availability, Reliability and Security, ARES, 2006.

[10]. Zhang, Jiong, Mohammad Zulkernine, Anwar Haque. "Random-forests-based network intrusion detection systems." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 38, no. 5, pp. 649-659, 2008.

[11]. Kumar, V., & Wu, X. (Eds.). "The top ten algorithms in data mining." CRC Press.2009.

[12]. Pinem, AsryFaidhulAshaari, Erwin Budi Setiawan. "Implementation of classification and regression Tree (CART) and fuzzy logic algorithm for intrusion detection system." In International Conference on Information and Communication Technology (ICoICT), 2015 3rd, pp. 266-271. IEEE, 2015.