International Journal of Enhanced Research in Management & Computer Applications
ISSN: 2319-7471, Vol. 6 Issue 6, June-2017, Impact Factor: 1.544

Effective Analysis of Intrusion Detection System using Rough Set Theory

Sagar Deshapande[1], Prof. S.L Deshpande[2]

[1]Department of Studies in of Computer Network Engineering, VTU Belagavi, Karnataka, India
[2]Professor Department Studies in Computer Network Engineering, VTU Belagavi, Karnataka, India

ABSTRACT

The principle behind intrusion detection techniques is to identify the malicious or intrusive behavior present in data packets of network. But the problem associated with the intrusion detection technique is that, it considers all the features present in the network packet which reduces the performance of system. The proposed model makes use of Rough Set Theory (RST) to generate reduct data set. This reduct data set contains required features which are sufficient to represent the whole data set. Time required to train and test dataset is analyzed from experimental results. Proposed system is compared with different intrusion detection techniques, result of comparison shows that proposed system gives better performance.

Keywords: Intrusion Detection System (IDS), Rough Set Theory (RST), Decision Tree classifier

1. INTRODUCTION

Intrusion detection technique can be considered as one of the core technologies of network security. The objective of IDS is to detect malicious behavior present in network [1]. The excessive use of internet leads to a huge amount of data generation over the application networks. This data may be personal data, organizational data or financial data. These data are more vulnerable to cyber attacks, so it is challenging task to provide security to such large amount of data.

There are basically two types of IDS, one is misuse based and another is anomaly based. In Misuse based intrusion detection the behaviors of the systems are monitored and matched with the intrusive behaviors or signatures to detect attacks. In anomaly based intrusion detection system, normal behavior of system is observed to find out threats in network [2].

Misuse based intrusion cannot detect unknown attacks and anomaly based increases false alarm rate and these method considers all the features of data set which reduces performance of system. Proposed model tries to reduce these limitations. Presented method of Intrusion detection system contains three phase's namely pre-processing, feature extraction and classifier. In pre-processing, text data will be converted into numerical value which will help to attribute reduction and classify them as normal and anomaly.

Thus to address the problem of detecting intrusions, the proposed system will detect intrusions which uses rough set theory for feature extraction and decision tree as classifier. Brief explanation of proposed system is given in fallowing section.

2. LITERATURE SURVEY

As discussed in above section two basic methods of Intrusion detection systems are signature based intrusion detection and anomaly based intrusion detection. Signature based intrusion detection requires update in pattern library to detect unknown attacks. But it has advantage of low false alarm rate as compared to anomaly based intrusion detection method [3]. Anomaly based intrusion detection system analyze normal behavior of system and it is more flexible because it has ability to detect new or unknown attacks. It analyzes only normal behavior of system so it requires less memory for storage [4].

These disadvantages involved in intrusion detection method leads to hybrid models which combines the advantages of intrusion detection methods and at the same time overcomes the disadvantages involved in intrusion detection methods [5].

Feature selection is the one of major issue in intrusion detection system. Consideration of all the features of dataset reduces the performance of system therefore rough set theory is used as feature selection method. It increases performance of system and accuracy by generating reduct of dataset which gives reducted features, those are necessary for detecting intrusive behavior present in network [9][10].

Classifier analyzes features to classify dataset as normal and attacked. Over the period of time different classifiers are tried to detect malicious behavior present in network and compared to get better results [11].

3. METHODOLOGY

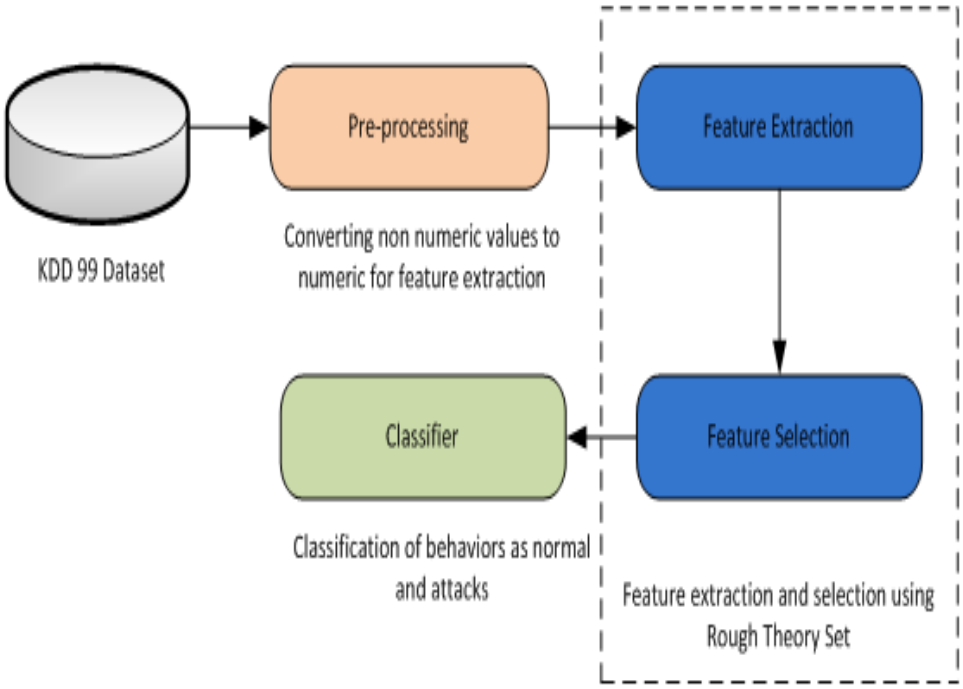

Figure 1: Architectural Block Diagram

Above figure 1 shows architectural block diagram of proposed model which helps to detect intrusive behavior present in network. It consist of phases such as pre-processing, feature selection& feature extraction and classifier.

KDD 99 Dataset: KDD dataset is given as input to system which will help to test system performance and it contains all the features of dataset.

Pre-processing: The KDD dataset will be pre-processed. Dataset may contain some text or non numerical values which have to be converted to numerical values for feature extraction and classify them as normal and anomaly data.

Feature extraction and selection using Rough Set Theory: Feature selection is an important process in intrusion detection. In proposed model RST is used for feature selection and it will produce reduct dataset. This reduct dataset reduces dimensionality of training dataset. Rough set theory is expected to increase efficiency by faster and more accurate detection rate [1].Data set may contain indistinguishable, imperfect and incompatible data. Theory of Rough Sets has especially intended to handle these types of situations.

Decision Tree Classifier: Decision tree is machine learning algorithm used as classifier which will analyze each reduced attributes from rough set theory. It will tend to increase the accuracy of system. Decision tree will build a trained data set based on reduced attributes. The testing data set will be compared with the trained data set which is generated by the Decision tree. By comparing both sets, testing data set will be classified as normal or anomaly [6].Decision Tree provides the accurate result, easy to represent and understand, takes less memory to handle noisy data.

4. IMPLEMENTATION

KDD dataset is given as input to feature collection. Rough set theory is used to obtain reduct data set which contains required features of dataset. Then some features are given for training and some for testing.

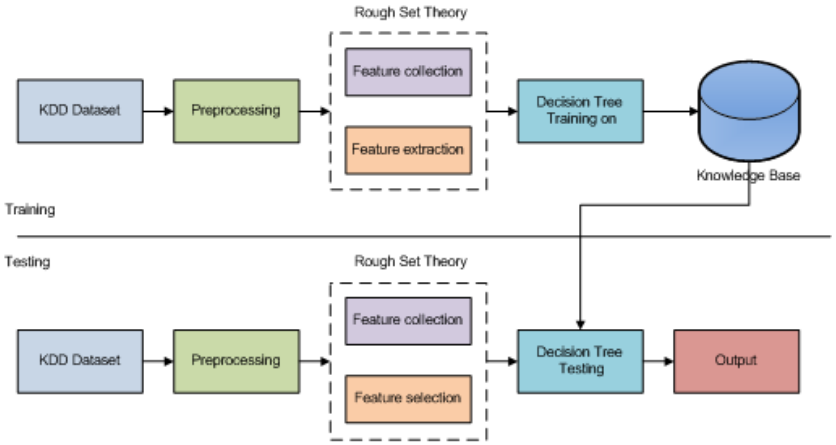

Figure 2: Architecture Flow of Proposed Method

The training and testing of the dataset selected is subjected to pre-processing, and features are selected and extracted using Rough Set Theory and the output of RST will be reduct feature set and is applied for Decision Training and Testing to classify and display the classification result as attack or normal as shown in Figure 2.

5. RESULTS

An entire process of intrusion detection here depends upon the rough set theory and decision tree classifier. Time taken for train and test the dataset using rough set theory and decision tree is considered to plot the graph which helps to analyze results.

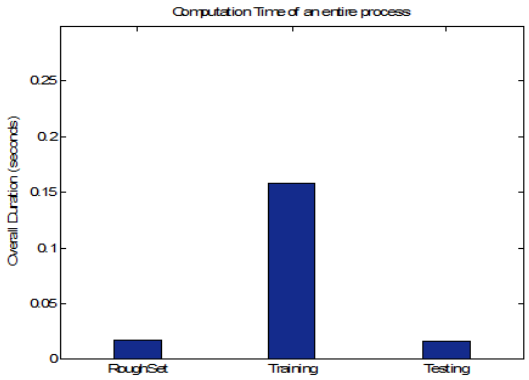

Figure 3: Plot of Rough Set Theory, training and testing process durations (in seconds)

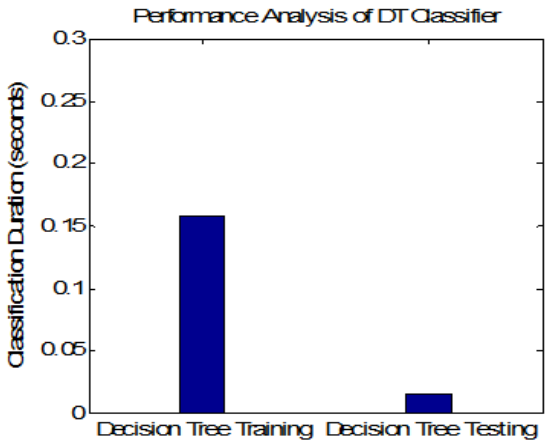

Figure 4: Plot of training and testing of DT classifier periods (in seconds)

It is observed that there is no change in time taken without reduction of features and with reduction of features for training and testing dataset. But feature reduction increases accuracy which is around 99% and reduces memory usage when processing of dataset takes place.

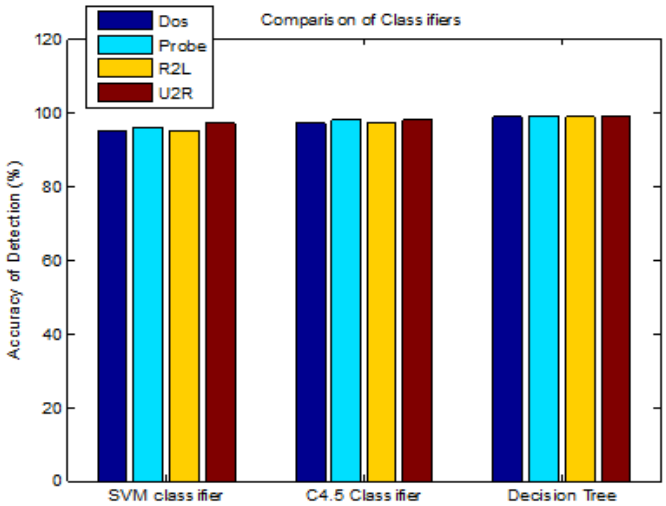

Figure 5: Comparison of three classifiers

Above comparison graph shows that SVM classifier has greater accuracy for detecting user to root attack which is around 97% and C4.5 classifier has greater accuracy for probe& user to root attack which is around 98%. Decision tree has greater accuracy that is approximately 99% which is irrespective of different attacks.

CONCLUSION

Proposed system uses rough set theory that reduces the features of KDD Dataset, this system can also reduces memory usage while processing dataset and the Decision tree provides greater accuracy as compared to other classifiers and it has greater accuracy irrespective of different types of attacks.

REFERENCES

[1]. Wa'el M. Mahmud, Hamdy N. Agiza, and Elsayed Radwan, "Intrusion Detection Using Rough Sets based Parallel Genetic Algorithm Hybrid Model", WCECS, 2009.
[2]. Shailendra Kumar Shrivastava, Preeti Jain, "Effective Anomaly Based Intrusion Detection using Rough Set Theory and Support Vector Machine", International Journal of Computer Applications, Volume 18, Issue 3, March 2011.
[3]. Howard E. Poston" A brief taxonomy of intrusion detection strategies" IEEE Conference Publications, pp 255-263,2012.
[4]. Priya U. Kadam, Prof. Manjusha Deshmukh "Various Approaches for Intrusion Detection System: An Overview", International Journal of Innovative Research in Computer and Communication Engineering Vol. 2, Issue 11, pp 6894-6902 2014.
[5]. Farid Lawan Bello; Kiran Ravulakollu; Amrita " Analysis and evaluation of Hybride intrusion detection system models", 2015 International Conference on Computers, Communications, and Systems (ICCCS),pp 93 – 97, 2015.
[6]. Kajal Rai, M. Syamala Devi, Ajay Guleria, "Decision Tree Based Algorithm for Intrusion Detection", International Journal Advanced Networking and Applications, Volume 7, Issue 4, 2016.
[7]. Sayali D. Jadhav, H. P. Channe, "Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques", International Journal of Science and Research (IJSR), Volume 5, Issue 1, 2016.
[8]. Nandita Sengupta, Jaydeep Sen, Jaya Sil, Moumita Saha, "Designing of on Line Intrusion Detection System using Rough Set Theory and Q-learning Algorithm", ELSEVIER, 2013.
[9]. Nikita Gupta, Narender Singh, Vijay Sharma, Tarun Sharma, Aman Singh Bhandari, "Feature Selection and Classification of Intrusion Detection System using Rough Set", International Journal of Communication Network Security, Volume 2, Issue 2, 2013.
[10]. Prasanta Gogoi and Dhruba K. Bhattacharyya, "A rough set-based effective rule generation method for classification with an application in intrusion detection", International Journal of Security and Networks, Volume 8, Issue 2, 2013.
[11]. Meghana Solanki and Vidya Dhamdhere, "Intrusion Detection System Using Means of Data Mining By Using C 4.5 Algorithm", International Journal of Application or Innovation in Engineering & Management, Volume 4, Issue 5, 2015.