Chapter 7

Conclusions, Suggestions and Scope for future research

7.1. Introduction

This chapter gives an idea about whole research work. This research work is carried out in order to develop data mining framework for detection of intrusion attack on computer network security.

7.2. Conclusions

1. Intrusion based security attacks are challenging for Pune IT industrial units.

2. Unauthorized access of computer network is most severe security threat to Pune IT industrial units.

3. Network administrators of Pune IT industrial units, considers ‘accuracy of intrusion detection’ as most important parameter for selection of IDS.

4. Pune IT industrial units face problem of false alarm generation with existing intrusion detection system.

5. In this study, attempts have been made to use Data Mining techniques with the aim of detecting intrusion based security attacks in the computer network. Decision tree classification technique used by SIDDM framework is capable and usable for intrusion detection. This technique with appropriate parameter and feature reduction is able to better classify network activity and recognize whether it is valid or not.
6. This study proposes the supervised data mining approach SIDDM for detection of intrusion based security attack on the computer system. Supervised models are constructed from large storage of network data and once model is built it can be used to predict attack in unknown network data. It proposes method to identify threats which may serious harm to computer.

7. The proposed model has prognostic capability, for unknown network data it significantly identifies attacks. Model will offer the advantage of considering those unlabeled records. This model is used to classify the network data samples as anomalous behaviour data or the normal behaviour data. Thereby the proposed model can be greatly deployed for intrusion detection in IT industrial units.

8. This research proposes anomaly based intrusion detection through SIDDM framework. This systematic framework is developed to detect intrusion based security attack using data mining whereas most of commercial IDS do this by statistical analysis. In this empirical work, experiments are performed using supervised classifiers on benchmarked KDD network data collection.

9. From empirical results of this research experiments, it is concluded that decision tree-J48 classifier is best and stable classifier for organizations concerned with overall correct classification of intrusion detection. The experimental results on KDD benchmark dataset evident that proposed algorithm achieved high detection rate on different types of network attacks. Comparison of all the performed experiments result shows, that the classifier with 10-fold cross validation using the J48 decision tree algorithm with the appropriate parameter values showed the best classification accuracy. This classifier model has a prediction accuracy of 99.742% on the training datasets. If J48 is used with ensemble method it takes long time for detection therefore without ensemble methods are faster and accurate for intrusion based security attack detection.

10. In summary, the results from this study contribute towards improving the networking security and give solution for detection of intrusion based security attack.
Proposed framework for intrusion detection

SIDDM (Systematic Intrusion Detection using Data Mining) model is developed in this research work. Computer network security always demands, superior methods for intrusion detection. IDS are to detect all intrusions at first effectively.

This framework proposes

- To construct intrusion detection system based on supervised data mining model and step to built model are as follows.

  Construct training model using available labeled network history data, apply data preprocessing and remove unnecessary attributes from data set. On preprocessed dataset, apply supervised attribute selection to select only most relevant feature; this list is given in chapter 5 in section 5.6.1. On reduced dataset apply 10 fold cross validation, training of model using J48 decision tree classifier. Once model is trained test using unclassified and unlabeled network data; now model is ready for intrusion detection for new data.

- Alternatively rules constructed by this model can be incorporated in existing intrusion detection script or rule engine. these rules are given in annexure 2.

Features of framework

- SIDDM framework offers anomaly based intrusion detection for identification and categorization attacks.
- SIDDM framework uses supervised data mining method for construction of model.

Steps to use framework

Intrusion detection system using the concept of SIDDM framework uses following steps to identify whether the network data is normal or attack.

1. Collect network data using packet sniffer software.
2. Apply data preprocessing.
3. Classification the network data based on rules generated by model.

4. Attack with attack type identification will be done by model.

Advantages of SIDDM Model

- SIDDM Model strengthens computer network security by providing data mining based framework to find abnormality in data.

- SIDDM Model provides a method to construct intrusion detection system. Method proposed is useful for construction of highly accurate intrusion detection system.

- SIDDM Model detects intrusion with high accuracy. This model generates less amount of false positive and false negative.

- Intrusion attacks are classified with high efficiency i.e time taken by model construction is also less.

- Intrusion detection and Classification rules are generated. These rules are available in annexure 2.

- Rules generated by SIDDM for detection and classification of intrusion attack can be incorporated into tools like Snort (commercial intrusion detection tool), firewalls, or detection scripts to identify intrusion attack.

This thesis research contributes to both the network security and the data mining field. Below is the summary of contributions:

- This thesis presents a systematic analysis of several steps involved in a data mining process, providing both theoretical and realistic contributions.

Data mining techniques are used to specify the kind of patterns to be found in data mining tasks. Various data mining techniques are surveyed using experiments. before construction of predictive data mining model, supervised techniques like decision tree, rule based classification and bayes net are surveyed for their applicability in intrusion detection. Predictive: to perform inference on data and to make predictions. Prediction model discovers the
relationship between dependent and independent variables. Data mining showing how particular attributes within the data will behave in future

- The Network Intrusion predictive model, which is developed in this study, generated various patterns and rules.

- Thesis determines characteristic analogy for intrusion based security attacks.

- Thesis categorizes the types of losses due to detected attacks.

- Presents, a survey of the various data mining techniques that have been proposed towards the enhancement of IDSs.

- Demonstrates, the effectiveness of supervised classification techniques in detecting anomalies.

- Analyses, components of computer network security.

- Elaborates, Intrusion Detection System, which is one of most important component of computer network security.

Intrusion Detection Systems (IDS) are the second layer of defense. It detects the presence of attacks within traffic that flows in through the holes punched into the firewall. Intrusion detection is the process of which monitors the events occurring in a computer system or network to analyze them for signs of intrusion. An Intrusion Detection System (IDS) constantly monitors actions in a certain environment and decides whether they are part of a possible hostile attack or a valid use of the environment. Following are the types of intrusion detection system.

The conclusion, of this study has shown that the data mining methods generate interesting rules that are crucial for intrusion detection and prevention in the networking industry.

This thesis attempts to address the problem of intrusion attack detection with the use of data mining supervised model. In summary, the results from this study can contribute towards in improving the networking security.
7.3. Suggestions

For computer network security in IT industries, it is imperative that IT industries must enhance security to detect intrusion based security attack. Following are the suggestions for obtaining effective complete network security.

1. Implement second layer of security compulsorily.

   Pune IT industries must apply second layer of security. First layer of security antivirus and firewall are not sufficient. Compromise with security causes serious consequences. As antivirus and firewall do not offer completely secure network therefore along with firewall and antivirus other security components are must. Therefore second layer of security needed to implement.

2. Install Anomaly based intrusion detection system as second layer of security. data mining based anomaly detection

3. Implement SIDDM Model (Data mining framework ) developed in this thesis for accurate intrusion detection

   Data mining framework suggested in this research gives higher accuracy of intrusion detection. This reduces problem of false alarm and gives accurate intrusion detection. This framework gives efficient alternative way for intrusion detection.

7.4. Future scope of work

   These experiments and their results provide reliable guidelines for future research in applying supervised classifiers for field of intrusion detection and expose some new avenues of research .Many improvements can be added to the intrusion detection system developed in this thesis.

   The study of intrusion detection systems is quite new relative to many other areas of systems research and it stands to reason that this topic offers a number of opportunities for future exploration. There are a variety of research directions that can be further developed using part of this thesis.
Chapter 7- Conclusions, Suggestions and Scope for future research

- Developing, An Intrusion detection tool (Software), using method proposed through this research.

- Investigating, applicability of unsupervised data mining techniques for intrusion detection.

- Developing, a system that operates with a more global scope may be capable of detecting distributed attacks or those that affect an entire enclave. Development of such a system would be a valuable contribution to the study of intrusion detection.

- This study was carried out using supervised data mining techniques. Supervised Classification algorithms such as J48 decision tree, rule based One R and bayes net algorithms. So further investigation needs to be done using other classification algorithms such as Neural Networks and Support Vector Machine plus using association rule discovery.