A Framework for Improving Attack Detection Accuracy using Ensemble Methods

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering

in

Information Security

Submitted By

Cheshta Rani

(Roll No. 801333005)

Under the supervision of:

Dr. Shivani Goel

Assistant Professor

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA – 147004

July 2015

CERTIFICATE

I hereby certify that the work which is being presented in the thesis entitled, "A Framework for Improving Attack Detection Accuracy using Ensemble Methods" in partial fulfilment of the requirements for the award of degree of Master of Engineering in *Information Security* submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of *Dr. Shivani Goel*, and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

Signature:

(Cheshta Rani)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

(Dr. Shivani Goel)
Assistant Professor,
Computer Science and
Engineering Department

Countersigned by

(Dr. Deepak Garg)
Head
Computer Science and Engineering Department
Thapar University
Patiala

(Dr. S. S. Bhatia)
Dean (Academic Affairs)
Thapar University
Patiala

ACKNOWLEDGEMENT

No volume of words is enough to express my gratitude towards my guide **Dr. Shivani Goel**, Department of Computer Science and Engineering, Thapar University, Patiala, who has been concerned and has aided for all the materials essential for the preparation of this thesis report. She has helped to explore this vast topic in an organized manner and provided me with all the ideas on how to work towards a research oriented venture.

I am also thankful to **Dr. Deepak Garg**, Head of Department, Computer Science Engineering Department, and **Ms. Jhilik Bhattacharya**, P.G. Coordinator, for the motivation and inspiration that triggered me for the thesis work.

I would also like to thank the staff members and my colleagues who were always there at the need of the hour and provided will all the help and facilities, which I required, for the completion of my thesis work.

Most importantly, I would like to thank my parents and the almighty for showing me the right direction out of the blue, to help me stay calm in the oddest of the times and keep moving even at times when there was no hope.

Cheshta Rani

ABSTRACT

With the development of internet and communication cyber movement started into new era. Today internet is used by all the organizations and people and they share a lot of sensitive information. A lot of attacks occur through the internet. These attacks exploit the vulnerabilities present in the system and may destroy the sensitive information present in the system. Protection mechanisms need to be provided to protect against these attacks. Firewalls and other basic security measures have been implemented to counter these attacks but these have failed as everyday novel ideas are developed by attackers to attack the system. Thus there is need to develop a system to eradicate these attacks as they damage the confidential information of the organizations. Intrusion detection systems are used for this purpose. Data mining can be used in case of intrusion detection system to differentiate between legitimate and illegitimate connections. Various classification algorithms can be applied that classify the connection either as normal or of specific attack type. Ensemble learning techniques are being currently considered as a new way to detect intrusive activities in systems as they have higher accuracy. The proposed approach is fusion of classification with boosting algorithms. In ensemble learning fusion of two or more techniques is done and accuracy of the combined system is large as compared to the individual techniques. The proposed model is applied on KDDCUP'99 dataset i.e. widely available dataset for intrusion detection systems. The results of the individual classification algorithms are compared with ensembling results. The ensemble learning better classifies the results to their proper category in terms of accuracy and number or instances properly classified.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
CAT	Change Aggregation Tree
CIA	Confidentiality, Integrity and Availability
DCD	Distributed Change-point Detection
DDoS	Distributed Denial of Service
DETER	Cyber Defense Technology Experimental Research
DoS	Denial of Service
GDA	Generalized Discriminant Analysis
HIDS	Host based Intrusion Detection System
HMM	Hidden Markov Model
IDS	Intrusion Detection System
iSVM	Improved Support Vector Machine
KDD	Knowledge Discovery and Data Mining
LAN	Local Area Network
LDA	Linear Discriminant Analysis
NIDS	Network based Intrusion Detection System
R2L	Remote to Local
SVM	Support Vector Machine
U2R	User to Root

Chapter 1

Introduction

1.1 Intrusion

Internet is global public network. With the development of internet and communication system, cyber movement started into a new era. Internet is used by all people and government agencies, for their business activities, personal affairs etc. thus sharing their valuable information. As the usage and complexity of internet has grown, the risk of maintaining the security has also grown. Internet provides an easy way of communication to end users and with this, the risk of threats, malicious activities, intrusion and vulnerabilities is increased.

An intrusion is defined as any set of actions that attempt to compromise the integrity, confidentiality and availability of a resource [1]. An attack is any kind of action that is used to destroy, alter, steal or gain unauthorized access to the system or make unauthorized use of the assets. The attacker can use different methods or skills to exploit the vulnerabilities present in the system, thus posing threat to the confidentiality of information and making the system unreliable to use.

There are different categories of attack such as denial of service (DoS), probe, viruses, Trojans, malware etc. The attackers find the vulnerabilities present in the system application or operating system and use that to exploit the system or to gain access to the system. Most of these attacks occur due to misconfigured software, vulnerable software or open ports present in system.

Lippman et al. provided the taxonomy of the attacks on DARPA 98 dataset in 1999[2]. This taxonomy was used to evaluate the accuracy and performance of the designed intrusion detection systems. There are mainly four categories of attacks namely DoS, U2R, probe and R2L and are discussed below:

1. **Denial of Service (DoS) Attacks**- In this type of attack attackers attempt to make a computer or network resource unavailable by flooding with unwanted traffic,

thereby preventing the normal use of computer or network resources. Different types of DoS attacks are mailbomb, neptune, land, smurf, teardrop etc. Teardrop and ping of death are bugs which create abnormal packets which lead to crashing the system. Mailbomb and neptune are considered as abuses.

2. **User to Root (U2R) Attacks**-In User to Root attack, attacker starts out as normal user on the computer and then exploits the vulnerabilities present in the system and obtains the root or administrator level access to the system. The root level access can be obtained by sniffing passwords, dictionary attack, social engineering attack or buffer overflow where amount of data stored in static buffer is more than its capacity. The programs that are unable to manage temporary files are usually exploited in U2R attacks.
3. **Remote to Local (R2L) Attacks-**In Remote to Local attack, the attacker is not having the account on the local machine but has the ability to send the packets to the system on network. Attacker exploits the vulnerability present to gain local access as a user and can tamper the data. Different types of R2L attacks are ftp-write, guessing password, imap, phf, snmp get attack, sendmail etc. Xlock is a type of social engineering attack while ftp-write, xsnoop and guest try to manipulate the security policies of the system.
4. **Probe Attacks-** Probe attack is an attempt to gather information about computer network such as IP addresses used in the network, operating system type, ports active on the system and vulnerabilities. The attackers use scanning tools to acquire information about the computer network and the vulnerabilities present can be used to launch an attack. Nmap, mscan, ipsweep, satan, portsweep etc. are probe attack types.

1.2 Intrusion detection system

Firewalls and basic security measures are inefficient in detecting the intrusions or attacks discussed above as attacker can use the novel ideas to do the attack. So intrusion detection systems are needed to counteract these vulnerabilities and maintain the CIA (Confidentiality, Integrity and Availability) triad of resources.

CIA triad (Confidentiality, Integrity and Availability) is the most widely used security model and one of the essential principles of information security. These are explained as follows:

- Confidentiality- **:** Confidentiality refers to ensuring that information is only available to the person who has right to access the information or to authorized person. To maintain confidentiality of data various encryption techniques are used applied.
- Integrity-**:** Integrity means the information or data can be modified or deleted by the authorized person using agreed terms and conditions.
- Availability-: Availability means that information and data must be available to authorized user when needed. By ensuring the availability we can prevent Denial-of-service (prevents the normal use of computer network by legitimate user by flooding the target system) attacks.

There are serious consequences to the end users/ parties if any of these principles is breached.

Therefore, intrusion detection system is an automated system which can detect these threats in the computer network and maintain the integrity, confidentiality, and availability of the resources. IDS is a security system that monitors computer systems and network for all inbound and outbound traffic and identifies the entities that attempt to cripple the security controls in place[3].

The intrusion detection systems are classified into two types: network based IDS (NIDS) and host based IDS (HIDS).

- **Network based IDS (NIDS):** NIDS are used to monitor the traffic moving from or to all the devices in the network. They capture the data packets using network tap and work in promiscuous mode. They match the captured data with the database of signatures and trigger the alert if match is found. They are responsible for detecting anomalous, inappropriate, or other data that may be considered unauthorized and harmful occurring on a network. Some advantages of NIDS are:

- They are easy to deploy as it detects the attack in the network area and does not influence the existing infrastructure.
- NIDS works in real time and detects the vulnerabilities as soon as they occur.

- **Host based IDS (HIDS):** Host based IDS are deployed on individual target machine and it uses software that monitors operating system specific logs, system event, and security logs. HIDS monitors the packets moving to or from the target device only and will alert the user if suspicious activity is detected. It works by taking the snapshot of system files and matches it to the previous captured snapshot. If any of the critical system files were modified or deleted, an alert is sent to the administrator or user to investigate. If a match is detected then it triggers/ flags the alarm.
  - HIDS monitors the activities with more accuracy than NIDS and generates fewer false alarms.
  - No extra hardware is required for HIDS as they are installed on the individual machine. These are less costly than NIDS.
  - HIDS monitors the changes such as file access, altered file permissions etc. NIDS does not work like this.

To detect and identify the anomalous behavior intrusion detection system uses two methods. These are misuse detection and anomaly detection.

- **Anomaly detection:** This approach is based upon statistical knowledge of the normal activity of system. This statistical profile consists of legitimate traffic of the network. This approach any activity is considered anomalous if it deviates from the normal profile [2, 4]. This model detects intrusions irrespective of the system type, system vulnerabilities, and type of intrusions provided that normal profile stored in at provides actual normal working condition. This system is able to detect novel and unforeseen attacks. This system has limitations:
  - Audit trails that are used for creating normal profile may contain traces of the malicious activity which are not detected by IDS.
  - Improper normal profile may cause high rate of false alarms.

- **Misuse detection**- Misuse detection is based upon the signatures of attacks which is detailed description of actions performed by attacker to carry out attack. When a malicious user tries to invade, the activity is matched with the signatures stored. If there is match then alarm is triggered. The effectiveness of this system depends upon how frequently the signatures of latest attacks are updated. This methodology is good at detecting known attacks and generates fewer false alarms. The drawback of this method is that it is difficult to maintain and update the signatures as a number of attack variants can be designed to launch the attack.

With all the benefits of the intrusion detection system there are some challenges faced by them. Some of these are:

- Human Intervention-Today the IDSs are mostly the automated systems that facilitate triggering an alarm for security administrator when an intrusive activity is detected, shunning malicious connection or cease the connection by modifying the router's access control list etc. although these automated system provide a lot of facilities but a skilled professional is required to monitor the log files and intrusive activities in recorded by IDS. As intrusion detection system cannot give periodical analysis of intrusions detected.
- False Positives- IDS has the capability of detecting malicious activities and triggering alarms, but this may also creates a lot of false positive alarms. False positive occurs when a normal traffic is considered as malicious. Therefore, IDS is considered as improperly configured.
- Deployment- In design and implementation phase of the IDS. It is very important to plan how to deploy the IDS. Organizations can adopt HIDS or NIDS or a combination of NIDS and HIDS according to their need except that they have sufficient resources for the IDS. Most of the organizations use hybrid approach of HIDS and NIDS [3] and after that choice is to use NIDS as it can monitor multiple systems and there is no need to install software on the production system.
- Encrypted Data- Encrypted data cannot be evaluated by IDS. Therefore, if encrypted data gets into the system. It can release viruses, trojans which could affect the system severely.

- Reactive in nature- With the growth of internet at faster pace, the hackers use new ideas to launch attacks. This a challenge since new attacks and new variants of attacks are currently being developed to launch the attack and time when the signatures of these attacks are publically available new attack variant are developed to have the same effect as the original attack. So, it is necessary to update the signature database as soon as new attack or vulnerability is detected.

Data mining can help in improving the intrusion detection system by overcoming the challenges discussed above.

1.3 Data Mining

Data mining also known as Knowledge-Discovery and Data Mining is the process of discovering the patterns from large volume of data using artificial intelligence, machine learning, database systems or statistics. Data mining process extracts information from the dataset and transforms into an understandable form to be used later. This can frequently be applied to any large-scale data or information processing as well as any application of decision support system, which include machine learning, artificial intelligence and business intelligence.

The actual data mining task can be automatic or semi-automatic. It involves analysis of large volume of data to extract previously unknown patterns such as groups of data records, unusual records, and dependencies (i.e. association rule mining).

The **Knowledge Discovery in Databases (KDD) process** is commonly defined with the stages: Selection, pre-processing, transformation, data mining and interpretation /evaluation. Figure 1 shows different stages of KDD process [5].

- Selection: This stage is concerned with segmenting or selecting data relevant to some criteria.
- Pre-processing: Pre-processing is the stage where unnecessary information is removed or data cleaning is done. When data is taken from several sources, it is quite possible that redundant data is present i.e. same information is present in

different sources in different format. This stage ensures data to a consistent format by cleaning the data.

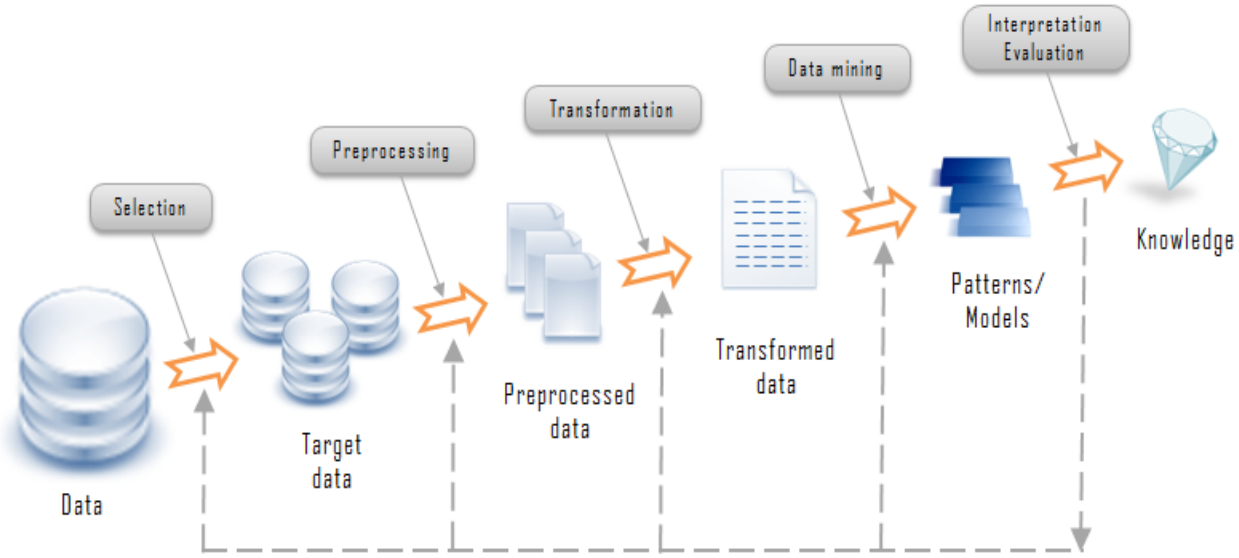

Figure 1.1 Knowledge Discovery Process

- Transformation: In this stage data is transformed in the form suitable for data mining or data is made useful or navigable.
- Data mining: This stage is concerned with actual extraction of result patterns. The *data mining* algorithms like classification, regression, sequence analysis, association rules are applied to process the data to form output which is in the form of patterns or rules.
- Interpretation / Evaluation: The patterns and rules thus formed are converted to new useful knowledge or information which is used in decision making.

1.4 Data mining techniques:

There are two principle goals of data mining: description and prediction [6]. Description refers to finding patterns that describe data and present for user interpretation. Prediction makes use of existing values in the database and predicts the future values of interest. There are several data mining techniques like association rules, classification, clustering, sequential patterns, frequent episodes etc.

- Clustering: Clustering is the method of grouping data into various groups so that data in each group have similar trends or patterns. Clustering is partitioning of data into clusters. Clustering is regarded as a form of unsupervised learning. The objectives of clustering are to uncover the natural groupings, to find the valid and consistent organization of data and to initiate the hypothesis about data. The goal of clustering algorithm is to identify all the clusters or groups in an optimal way. Clustering has advantage over classification technique that it does not require any labelled data for training.
- Classification: Classification technique analyzes and categorises data into classes (known). Entire data is sampled and each sample is labelled with a previously known class label. The input to the classification technique is training set in which different labels are already known. Then it analyzes the training dataset to form a model and the aim of classification technique is to assign a label to future unlabeled dataset. Classification is called as supervised learning as the type of class is known previously known. It is linear regression model.

  Classification based IDS attempt to classify data as normal or any specific attack type. There are various classification models for example decision tree, neural networks, genetic algorithms etc. The main challenge in classification technique is to reduce the number of false positives and false negatives. False positive occurs when any legitimate activity is mistakenly treated as error. False negative occurs when any erroneous activity is taken as legitimate.
- Association rules: Association rules are IF-THEN statements that are used to uncover the relationships between unrelated data in database or any other information store. Association rules have two parts: (1) antecedent i.e. If part and (2) consequence i.e. THEN part. Association rules are in the following form

  $$X => Y$$

  Where X and Y are sets of items.

  It means that transactions that contain X they also tend to contain Y. Association rules use criteria of support and confidence to identify the important relationships. Support means how often the items appear in the database. Confidence indicates how one item is dependent on the other item. Patterns which have high values for

support and confidence are obvious to the users and those of having low value are of no use. Thus, patterns with intermediate values provide the user with previously unknown and interesting information.

- Neural Networks: Neural network is an emulation of the biological neural system. Neural network is a mathematical model which has ability to learn like humans. These have ability to drive meaning from the complicated data and extract patterns that are too complex to be obtained by humans. Neural network has set of nodes i.e. processing elements similar to neurons in human brain. These nodes are interconnected in network an can identify the patterns when they are exposed to data [7]. Neural network is very successful for classification but it has two limitations:
    - It leaves the network as a black box that means no explanation about the result is given.
    - Another problem with neural network is that they have very long learning time and it becomes worse when volume of data is too high.
- Support Vector Machine: Support Vector Machine (SVM) is based on statistical learning theory [6]. The main idea of SVM is to map the dataset non-linearly into the high dimensional feature space and then classify the data using some linear discriminator. SVM is used in the area of regression, classification and decision tree construction. SVM work on risk minimization structure i.e. they minimize an upper bound on error.
- Genetic algorithms: Genetic algorithm is inspired by the Darwin's theory of evolution that states survival of fittest and reproduction is through cross-breeding. Genetic algorithms are best suited for the optimization problems.
    - Initially a population of individuals is created randomly.
    - The pairs of individual are selected based upon fitness value that means one with the higher value of fitness function has higher probability of being selected. Different selection techniques are used like tournament selection, steady state selection and rank selection.
    - These are combined to produce offspring for next generation using crossover.

- Crossover means two chromosomes are combined to produce new offspring with some user defined probability which is 0.9. Crossover can be one point, two point or uniform.
- Mutation is a genetic operator which alters one or more number of genes randomly in chromosomes to maintain genetic diversity. Probability of this is kept low (i.e. 0.1).
- Newly generated population is further used in the algorithm.
- When the termination condition is satisfied, from the current solutions best solutions is selected and returned as optimal solution.

Data mining uses genetic algorithms to minimize the classification errors.

1.5 Ensemble learning

Ensemble learning is use of multiple machine learning algorithms to obtain the performance which is better than any one of the constituting algorithms [8]. It is supervised learning algorithm. The ensemble term is used for the methods that produce multiple hypotheses using one base learner. The term multiple classifier system is also used for the combination of hypotheses (not generated by same base learner). More computation is required for evaluating prediction of an ensemble than prediction of a single model. Ensemble methods are applied to improve the generalization ability and robustness of the single estimator. Ensemble of two techniques that are similar will perform worse than ensemble of more diverse set.

Ensemble techniques are Bayesian model combination, boosting, Bayes model classifier, bootstrap aggregation (bagging), stacking etc.

1.6 Structure of Thesis

The summary of each of the chapter described in the thesis is given below:-

Chapter 2 represents literature review which exhibit the evolution of IDS and the most commonly used dataset of IDS, review of different data mining techniques used in the field of IDS and work done in the field of ensemble learning is evaluated

Chapter 3 illustrates the problem statement and the goals of the proposed work.

Chapter 4 depicts the proposed work and various algorithms used for this.

Chapter 5 discusses the evaluation parameters and the results obtained in proposed work and its comparison of different approaches.

Chapter 6 renders the conclusion, summary and future scope of the proposed work.

Chapter 2

Literature Review

2.1 Evolution of IDS

Intrusion is defined as invasion into organization's important information or person's privacy. Various techniques or methods were previously used to prevent the computer system from unauthorized use, like firewall, encryption techniques etc. These methods were not good enough to prevent the computer systems against intrusion by attackers or hackers. Hackers and attackers are becoming proficient day by day as new methods are being developed by them to breach the security of the system or organization. So the new component called intrusion detection system became important in the field of computer security. Denning in 1987, gave the first intrusion detection model [9]. After that many intrusion detection systems have been developed and these were made to determine the behavior of the network in efficient manner.

Porras and Valdes stated ways to extend both statistical and signature based analysis techniques to monitor the network traffic[10]. They presented the signs of malicious activity, malicious failures and other exceptional events by analyzing TCP/IP packet streams. They presented the standards to evaluate the performance of intrusion detection systems these standards are accuracy, performance and completeness. These are as follows:

- **Accuracy**: It means that attacks should be properly classified and there should be no false alarms.
- **Performance**: It is measured by rate at which data is processed.
- **Completeness**: Completeness is reflected in the number of harmful records that bypass the detection mechanism (false negatives). The system should be such that it detects all the attacks and for this the system should be auto updating and must have information about the attacks.

All the three standards directly depend upon the quality of event stream which is being analyzed.

The taxonomy of intrusion detection system was defined by Debar et. al [11, 12]. The different families of the intrusion detection were classified using their properties in this taxonomy. A brief description of different families of IDS is as follows:

- **Detection method:** It describes the features of the analyzer. Two trends are present in this: (1) The knowledge about different types of attacks and system vulnerabilities is collected in the system. When an attempt to exploit the system is detected then alarm is triggered. This trend is called misuse detection. (2) It detects intrusion by considering deviation from normal or valid behaviour, which is extracted from information collected from various methods. If any deviation is found then alarm is triggered i.e. anomaly detection.
- **Behaviour on detection:** It refers to the behaviour of the intrusion detection system when any intrusive activity happens. If intrusion detection system triggers the alarm and takes corrective or proactive measure to counter the attack then it is active and passive if the system triggers the alarm and does not take any action to control the attack.

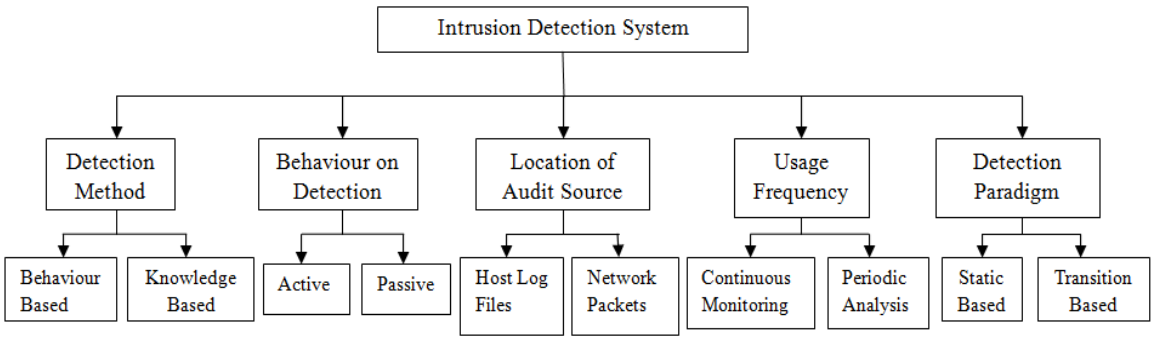

Figure 2.1 Classification of Intrusion Detection System

- **Audit source location:** Intrusion detection systems are distinguished based upon the type of information being analyzed by them. Host based audit sources are syslog, C2 security audit, system sources etc. SNMP information and network packets are network based audit sources.

- **Usage frequency:** It refers to the way detection system performs analysis. Continuous real time analysis is done by dynamic intrusion detection tools in which actions are taken immediately when intrusion happens. Periodic analysis is done in static intrusion detection tools which takes snapshot of the environment and analyze for vulnerable software , configuration errors etc.
- **Detection paradigm:** It refers to the methods used by IDS to detect attacks. It can be either state based (i.e. secure or insecure) or transition based i.e. transition from one state to another for example secure to insecure.

Two techniques are used for detecting attacks in intrusion detection systems: anomaly detection and signature detection. Anomaly detection is able to detect new attacks but has limitation that it has higher rate of false alarms. A new technique called specification based detection was proposed by Sekar et.al that produces lesser false alarms[13]. Method is not good for detecting denial of service and probing attacks. The specification constraints are extracted manually by human experts. Specification can be applied can be applied to host as well as users systems.

Various artificial intelligence, computational and machine learning techniques have been applied on different intrusion detection systems or models to obtain the results. Many issues or problems were faced while applying these techniques like vast amount of network traffic(dataset), adaptation to the changing environment and noisy information.

2.2. Dataset used in Intrusion Detection System

To evaluate the performance of any system or model, datasets are required which can specify whether the system or model is complied with the standards or not. The data for this evaluation can be collected from various sources like data packets, log files etc. Two most commonly used datasets are: DARPA 98 Lincoln dataset and KDDCUP 99 dataset. KDDCUP 99 dataset was created from 1998 DARPA intrusion evaluation dataset by Lincoln lab under contract to DARPA. The data set was created by preprocessing the tcpdump potions of IDS evaluation dataset. Lincoln labs setup environment to collect the data from local LAN. The dataset consists of 9 weeks of training data and 2 weeks of testing data. It consists of millions of TCP connections records and each connection

contains 41 attributes which contain features like source and target IP addresses, protocol used, duration of connection etc. as defined by Stolfo et al. [14]. Each connection is labeled as normal connection or one specific attack type. This dataset was harshly criticized by McHugh in 2000 besides its high usage [15]. Among the raised issues, two important issues were: data rates were too high and no validation was performed to check whether DARPA dataset was actually like real network traffic or not.

2.3 Review of different Data Mining Techniques

Veda et. al tried out different data mining approaches and analyzed the results of these approaches when they were used for anomaly detection[16]. They tried out outlier detection for network based IDS and prediction of system calls for host based IDS. For network based IDS input was 1.5 GB of tcpdump data collected over a span of 24 hours. This data was used as normal data and then preprocessing was done as given in to group data related to one connection or to extract time based, connection based and content based features [17]. Attack data was generated by simulating TCP-SYN flood attack using packet [18].

Different data mining techniques were applied to the data like association rules mining, clustering and classification. The clustering was done by using all of the attributes of the dataset or by selecting some of the specific attributes that were affected by the attack to increase the efficiency. For host based IDS normal data was generated by using stace utility for sshd process. Attack was launched by attempting to connect using ssh or by using different password for this. For host based decision trees, SVM, NaïveBayes and meta learners were used and decision trees gave the best result.

Signature based detection has its limitation in detecting the novel attacks, anomaly based detection overcame this limitation of signature based detection. For evaluating the performance of the intrusion detection systems KDDCUP 99 is most widely used dataset. After some statistical analysis it was found that this dataset suffers from some problems which affect the performance of the system. To solve these problems a new dataset was proposed called NSL-KDD [2]. NSL-KDD is publically available on [19]. NSL-KDD has several advantages over the KDDCUP dataset:

- It does not include any redundant records in train set.
- The test set has no duplicate records.
- The number of records is reasonable in both test and training set.
- The number of records selected from each difficulty level is inversely proportional to the number of records in KDDCUP'99 dataset.

NSL-KDD still suffers from the problems discovered by McHugh but it provides effective benchmark for researchers to evaluating different systems [15].

Classification algorithms used for classification of attacks must do the classification accurately and there should be less number of false positive and negatives. Webb proposes that classification errors can be reduced by using multiboosting which is combination of wagging and adaboost [20].

Salvatore et.al proposed data mining technique like classification, association rules, meta learning, frequent episodes to capture the behavior of the normal and malicious activity [21]. To detect the new attacks, rules in this system can be updated as attackers change their tactics periodically to evade the security system. Meta learning is used as means to combine different intrusion detection models.

DDoS attack causes system crashes, heavy data load, congestion in the communication links, traffic overflow in SYN queues etc. So in order to minimize the effect of the DDoS attack, it is required to detect the attack as soon as possible. Chen et.al using Change Aggregation Tree (CAT), designed Distributed Change-point detection (DCD) [22]. DCD detects abrupt changes in the network traffic. DCD system up to 16 network domains was simulated on the Cyber Defense Technology Experimental Research (DETER) testbed, 220-node PC cluster for emulation experiment at Information Science Institute in Southern California. This method achieved a higher accuracy of 98% with 1% false positive alarms.

Intrusion detection system monitors the network traffic to find any intrusive or malicious activity and provides a layer of defense. It can be done at individual host level i.e. Host based IDS (HIDS) or at network level i.e. network based IDS (NIDS). Kaushik and Deshmukh discussed that there are four major classes of attacks namely Denial of service

(DoS), probe, User to local (U2R) and Remote to user (R2L) [23]. It was found that probe attack exists when there is alteration of file sent by the user or some data is added to the data field sent by sender. R2L attacks are present when the duration of the connection is maximum or above a particular threshold. DoS attack is present when packet or data does not reach the destination. There are a number of ways to detect attacks in Intrusion detection systems and one can make his/her choice according to the need.

A serious network security threat is imposed due to DDoS attack. There is a need to detect the Dos attack at early stage so that damage caused is less and countermeasure to that can be applied much before the damage is done to the system. Most of the methods applied for early detection cannot achieve:

- Detection with fewer false alarms.
- Transfer of packets in real time.

Naguen and Choi proposed a method for proactive detection of DDoS attacks by using k-NN classifier in anti DoS framework [24]. They classified the network into three classes: normal, pre attack and attack.

- Pre attack: it consists of two phases:
    1. Phase 1: selection of handlers and agents.
    2. Phase 2: communication and compromise the victim.
- Attack: it is the actual phase in which attack occurs i.e. phase 3.
- Normal status of the network.

This method is very easy to implement and very less time consuming. K-NN classifier has achieved an accuracy of 91% for early detection of DoS attacks in anti-DoS framework.

IPv6 will replace IPv4 in next generation of internet protocol. For IPv6 there was a need to consider the development of IDS. Liu and Lai developed new intrusion detection model for IPv6 network and was successfully applied in IPv6 experimental network in network[25]. This model was able to detect probe and U2R attacks however this failed to detect R2L and DoS attacks.

The network data provided for IDS are very large and have a lot of ineffectual information and there is a need to remove that information from the original dataset. The

space and time complexities of most of the classifiers are exponential function of the input dimensions. The demand for the number of classifiers samples for classifier's training increases exponentially with feature space dimension. This is called curse of dimensionality. To improve the overall performance of the IDS that are based on classifiers there is a need to reduce the feature set which will decrease the space complexity and time complexity of the system and improve the accuracy.

Most commonly method used for feature reduction is LDA (Linear Discriminant analysis). This technique projects the data to the direction that has largest variance. LDA has some limitations like how many directions one need to choose and it is not good for non-linear datasets or where input space and feature space are related non-linearly. Singh and Silakari proposed a novel technique GDA (Generalized Discriminamt Analysis) for feature reduction[26]. This method reduces the number of input features and increases the accuracy of classification. For selecting the most discriminant features it reduces testing and training time. ANN and C4.5 are used to compare the performance in [26].

2.4 Ensemble Learning in IDS

Ensembling methods can be used in the field of IDS to improve the detection accuracy or detection rate and reduce false alarm rate. A lot of research has been done in this field. Below is work done by some researchers in this field.

Data mining methods can be used in intrusion detection system. There are many data mining algorithms like Support Vector Machine (SVM), clustering, classification tree algorithm (e.g. C4.5). In order to detect the attack efficiently number of false positive must be reduced. Balaji and Varalakshmi used C4.5 classification algorithm and used multiboosting to reduce false positives and errors of C4.5 algorithm [27]. Multiboosting combines features of wagging and Adaboost.

Lappas and Pelechrinis conducted a survey of different data mining techniques that have been applied for the improvement of IDS[28]. These techniques include feature selection, machine learning, statistical approach like Hidden Markov Model (HMM), predictive analysis and ensembling. A number of systems that are implemented using data mining are presented in this like Information Security Officer's Assistant (ISOA), Distributed

Intrusion Detection System (DIDS), Intrusion Detection using Data Mining (IDDM) etc. They also proposed a technique called bi-clustering to analyze the network traffic and improve IDS.

Many IDSs are previously proposed to classify data as normal or intrusive using data mining techniques. One problem with these systems is that intrusion data is very rare and difficult to obtain. To solve this problem novelty detection is used [29]. The novelty detection system use normal class patterns as training examples to generate the model.

Yeung and Chow proposed a non parametric density estimation approach for intrusion detection using normal data only as other parametric density estimation approach fail to detect novelty attacks[30]. This is based upon Parzen-Window estimator with Gaussian kernels. The advantage of the model is it can easily adapt to data changes. Results of this model are compared with winning entry of KDDcup which was submitted by Pfohringer in which ensembling of decision tree with boosting is used. This method gives higher TDR for U2R and R2L attacks which are difficult to detect as they involve very less number of connections. However TAR is slightly less for the proposed approach. This method can perform better under favorable conditions with no intrusion at all.

Two approaches are currently being used for intrusion detection. These are misuse detection and anomaly detection. Both of these approaches are based on "pattern matching" e.g. anomaly detection approach classifies any activity as malicious. If it does not match with normal profile and misuse detection classifies activity as malicious if that matches with the particular attack signature. For IDS system with generalization capabilities which support detection of attacks that are not known previously and have no described pattern. Didaci et.al proposed pattern recognition approach which was based on learning by example[31]. Ensemble learning techniques are effective in providing more reliable results, out of the three fusion rules used, fusion rule based on "belief" function has given better results. The system reduces overall error rate but may also result in reduction of generalization capabilities.

Freitas discussed the usefulness of genetic algorithm and genetic programming in data mining and knowledge discovery[32]. The goal was data mining for classification. Some

pre-processing like attribute selection and post-processing like ensemble of classifiers steps is also discussed. According to him, fitness function adapt for extracting high level of knowledge from genetic operators.

Zhou et.al analyzed ensemble and constituting neural network in the context of regression and classification, which states that better results can be obtained from the ensemble of neural networks instead of the single neural network[33]. An approach named GASEN is presented which trains number of neural networks first and then assigns weights to each network. Then genetic algorithm is used to select some neural networks based upon weights to make an ensemble. GASEN generate the ensemble with very smaller size and stronger generalization ability than boosting and bagging.

Singh et.al [34] proposed a new algorithm called iSVM i.e. improved Support Vector machine for classification of cyber attacks. This is modified or improved version of traditional SVM. Traditional SVM was enhanced by changing the Gaussian kernel that was enlarged to spatial resolution using conformal mapping around margins. It is based upon Riemannian geometry induced by kernel function [35]. The results of iSVM indicate 100% accuracy for classification of DoS and normal class. Testing and training time for this is reduced and high rate of false alarms is improved. Thus, a combination of classification and feature reduction give better results.

Network based computer systems play vital role in today's world hence become target for attackers. To provide high security against intrusion detection a number of machine learning and pattern recognition techniques are proposed. Veerwal et.al proposed the ensemble of classifiers and soft labels for intrusion detection [36]. The performance of fusing Artificial Neural Network(ANN) and SVM classifiers with Average Bayes Combination, Dempster Shafer theory and neural network is proposed. Out of the three fusion with Dempester Shafer theory performs best.

Alhaddad et.al proposed ensemblers with Naïve Bayes and Decision trees as base classifiers [37]. The results show that ensemblers with decision trees perform better as compared to naïve Bayes. The results also show that even the performance of a single classifier decision tree is best which has less training and testing time and accuracy is improved.

Kulkarni proposed ensemble methods which are boosting (AdaBoostM1), bagging and compared the results with J48 decision tree [38]. Weka experimenter was used to classify the results obtained from 10% of KDDCUP 99 dataset. AdaBoostM1 performs best among three as the probability of incorrect classification and error rate is very less.

Patel and Tiwari has done ensembling of two different classifiers with bagging. Firstly individual classifiers SVM and decision trees are applied then combined the result of both classifiers. To the combination bagging technique was applied and results of this were better than ensemble with boosting.

Kumar and Kumar have done a survey of different artificial intelligence techniques that had been applied in IDS [40]. Different ensembles which are based upon AI are compared based upon architecture, approach taken, evaluation metrics taken for comparison etc. It was shown that each method has its own advantages and disadvantages in intrusion detection.

Chapter 3

Problem statement

3.1 Problem Statement

KDDCUP 99 data set has five major classes of attack these are: normal, Denial of Service, Remote to User, User to Root and probe. Out of these 5 classes U2R and R2L have very less training data available and have 11 subclasses of attacks. These 2 classes are not properly trained and they don't give accurate results with test set. This is limitation of the existing IDS.

Classification algorithms are also used to classify any activity as normal or of a particular attack class. Classification algorithms perform well when they are properly trained. These algorithms take more training and testing time. The rate of the false positives and false negatives is also very high. Ensemble learning performs better in that case.

Classification algorithms along with boosting algorithms give better results..

It was analyzed by various researchers that soft computing techniques perform better than other techniques. Self organizing maps have the problem of high dimensionality, high computational overhead and high rate of false alarms. So, one technique alone will not give the best result.

Therefore ensemble of different technique will perform better result in that case. So fusion of classification technique along with bagging or boosting will give better results. The system thus created will contain the complementary features of different techniques and build a robust system.

3.2 Objectives of proposed work

The proposed work has following objectives:

- To build framework that has low or minimum false alarm rate and higher detection rate.
- To design a system that should be highly adaptable i.e. the system adapt itself with the behaviour of user or network and modify itself to perform accurately.
- To design an intrusion detection system that classifies different malicious activities to their true class.

- To design a framework that must be able to differentiate between the legitimate and illegitimate connections in a computer network.
- To design a framework that is highly interpretable and behaviour is easy to understand.

Chapter 4

Proposed work

The proposed approach provides a platform to detect the network intrusion activities and classifies them to one of the five classes according to the signatures using KDDCUP'99 dataset which is the standard data set available for intrusion detection in the network.

4.1 Evaluation of KDDCUP'99 Dataset

KDDCUP'99 dataset was prepared by processing the TCPdump part of DARPA IDS evaluation dataset. This dataset was created and managed by MIT Lincoln labs. The data was collected from intrusion activity simulated in U.S Air Force Local Area Network (LAN). The LAN was operated as it was true military environment. Attacks were generated using some generator or hand-injected in this LAN[41].

Lincoln labs collected nine weeks network traffic. Network traffic of seven weeks was used as training data and consisted of about five million connection records. Network traffic of two weeks was used as testing data that consisted of about two million connection records. Each connection record is a sequence of TCP packets and constitute 41 attributes. Each connection is labeled as either normal or of particular attack type.

KDDCUP'99 consists of about 4,90,000 connection records. Training dataset consists of 24 different types of attack and test set contains additional 14 types of attacks. All of these attacks come under four major classes of attacks: DoS, Probe, U2R and R2L.

The additional attack types of test set have different signatures and checked for whether captured or not by the system and this is done to increase performance.

Below table 4.1 shows different classes present in the dataset and subclasses corresponding to them.

Table 4.1 Classification of attacks

S.No.	Classes	Subclasses
1	Denial of Service(DOS)	Back, land, teardrop, Neptune, pod,smurf
2	Probe	Nmap, ipsweep, portsweep, satan
3	User to Root(U2R)	Buffer_overflow, perl, loadmodule, rootkit,
4	Remote to Local(R2L)	imap, phf, spy, ftp_write, guess_passwd, multihop, warezclient, warezmaster,
5	Normal	NA

Stolfo et.al defined some features of KDDCUP'99 dataset for detecting normal and attack activity and classified into categories [14] like basic, time based features, content based features and content based features:

- **Basic Features-** These are the set of features that can be extracted from each TCP packet or connection. There are 9 attributes that are considered as basic features. Table 4.2 given below represents basic features present in KDD dataset.

Table 4.2 Basic features extracted from individual TCP connection

Feature name	Description	Type of value
duration	Length of connection	Continuous
service	Type of network service on destination e.g. telnet, FTP, HTTP etc.	Discrete
protocol_type	Protocol type used e.g. UDP, TCP etc	Discrete
Src_bytes	Number of bytes sent from source to destination	Continuous
Dst_bytes	Number of bytes from destination to source	Continuous
Flag	Status of connection i.e. normal or error	Discrete
urgent	Number of urgent packets	Continuous
Land	Value is 1 if connection is to or from same host or port, otherwise value is 0	Discrete
wrong_fragment	Number of wrong fragments present	Continuous

- **Traffic Features**- This class of features are extracted based upon time interval. These can be "same host" or "same service" features. These are discussed below.
    - Same Host Features: This explores the connections that have the same destination host as present connection from past two seconds.
    - Same Service Features- This explores the connections that are having same service as the present connection form the past two seconds.

Table 4.3 Time based traffic features

Feature name	**Description**	**Type of value**
Count	Represents number of connections to the same host in past 2 seconds	Continuous
Feature specific to "same host" connections		
Serror_rate	Percentage of connections having SYN error	Continuous
rerror_rate	Percentage of connection with REJ error	Continuous
Srv_count	Number of current connections in past 2 seconds	Continuous
Diff_srv_rate	Percent of connections to different service	Continuous
Same_srv_rate	Percent of connections to same service	Continuous
Features referring to "same service" connections		
Srv_serror_rate	Percent of connections with SYN errors	Continuous
Srv_rerror_rate	Percent of connections with REJ errors	Continuous
Srv_diff_host_rate	Number of connections to different host	Continuous

- **Host Based Traffic Features**- Some attacks cannot be detected using time interval concept as they may use time interval larger than 2 seconds. For those type of attacks window of connections of hosts is used. For example these may use a window of 100 connections of hosts. These types of features are called host based features.

- **Content Based Features**- Stolfo et. al added some features to the data set that look for any malicious activity in data field. This type of features added using domain knowledge. These features are called as content based features. There are 13 features in this category.

Table 4.4 Content based features

Feature name	Description	Type of value
Hot	Number of hot indicators	Continuous
Num_compromised	Count of compromised conditions	Continuous
Logged_in	If successfully logged in then 1, Else 0	Discrete
Root_shell	If root shell is obtained then 1, else 0	Discrete
Num_shells	Number of shell prompts present	Continuous
Num_failed_logins	Number of login attempts that have failed	Continuous
Is_hot_login	If login is from hot list then 1 ; else 0	Discrete
Is_guest_login	If guest login then 1; else 0	Discrete
Su_attempted	1 if super user command attempted; else 0	Discrete
Num_root	Number of root level accesses	Continuous
Num_outbound_cmds	In FTP session, number of outbound commands	Continuous
Num_file_creations	Number of operation for file creations	Continuous
Num_access_files	Number of file access control operations	Continuous

4.2 Proposed Framework

Based on the concepts discussed above, a framework is proposed which identifies various classes of attacks with higher accuracy as shown in figure 4.1.

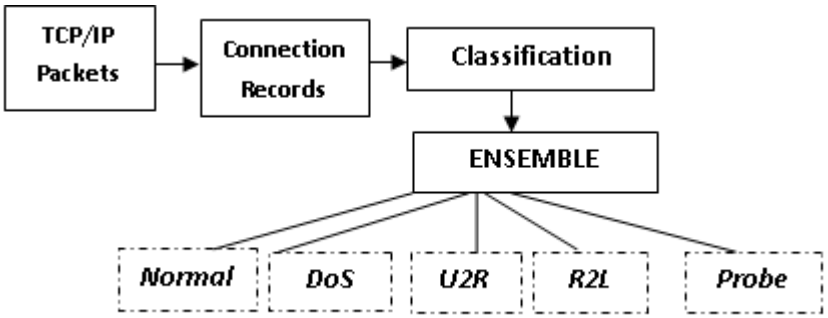

Figure 4.1 Proposed Framework

4.3 Classification

In the proposed framework, records are classified into the existing or known classes. A training dataset is used and from that a model is built. This model is used to classify the unlabeled instances of the dataset later. Thus based upon some past experience instances are classified to existing classes. This is supervised learning algorithm. Classification in this model is done using various classifiers like Naïve Bayes classifiers, random forests, decision tree methods etc.

The working of the classifiers is explained below:

Naïve Bayes Classifiers: These classifiers assign a particular class label to each of the record from dataset. Class labels are assigned from a finite set and each record is represented as vector of features.

For testing the proposed framework 50,000 records are randomly selected from the KDDCUP'99 dataset and each record is having 41 attributes. In Naïve Bayes Classifiers, feature is represented as:

$$y= \{y_1,y_2,y_3,\ldots\ldots\ldots\ldots\ldots\ldots,y_{40},y_{41}\} \quad (1)$$

To each instances to be classified, it assigns conditional probabilities as

$$p = (C_m \,|y_1 , y_2 \ldots\ldots y_{41}) \quad (2)$$

Conditional probalibity is calculated using the following formula

$$p(C_m|y) = \frac{p(C_m)p(y|C_m)}{p(y)} \quad (3)$$

$$p\left(C_m\middle|y_{1,}\ldots\ldots y_{41}\right) = \frac{1}{Z}\, p(C_m) \prod_{k=1}^{n} p(y_i|C_m) \quad (4)$$

Where Z = p(x) is a scaling function which depends on y_1, …… y_n.

Then classifier is combined with the decision rule that chooses the most probable outcome. Bayes classifier function used to assign the class label i.e. z = C_m. as follows:

$$z = argmax_{m\in\{1,\dots m\}}\, p(C_m) \prod_{k=1}^{n} p(y_i|C_m) \quad (5)$$

Random Forests: Random forest is ensemble learning algorithm used for classification and regression. It constructs a model tree based upon training data and then gives mean prediction or mode of class as output. Random forests reduce overfiting of data that is commonly present in decision trees. Random forests work by averaging decision trees that are trained on same training set but on different parts of it. The goal of this method is to reduce variance.

- Bagging or bootstrap aggregating is applied as training algorithm to random forests.
- Given a training set, $Y = y_1, \dots, y_n$ and responses $X= x_1, \dots, x_n$, bagging selects samples randomly with replacement and fits into the model. For c=i, ...C, samples with replacement are drawn say X_c and Y_c and then train decision tree on X_c and Y_c.
- After training prediction for undiscovered samples is done by averaging individual tree on undiscovered samples y`.

$$f = \frac{1}{C} \sum_{c=1}^{C} f_c(y`) \quad (6)$$

Where C = number of samples/ trees.

C is free sample and value chosen for this depends upon nature and size of training set. The value may vary from hundreds to thousands. Optimal value for C can be chosen using cross validation.

4.4 Ensemble Methods

Bagging: Bagging is also called as Bootstrap Aggregating. In this class of algorithms several instances of black-box estimator are created based on random subsets of training datasets. Then predictions from individual estimators are

combined to achieve final predictions. Bagging method reduces variance and overfitting. The algorithm works as discussed below:

1. Given a training set say A of size m.
2. Bagging generates new training set say $A_{i,}$ each of size m` by sampling the training set with uniformity and replacement.
3. As sampling is done with replacement, there may be some instances that are repeated. If m =m` for a large value of m, the number of unique instances *U* in A_i is given by following fraction

$$U = 1 - \frac{1}{e}$$

   The number of unique instances estimates to about 63.2 % of total instnces. Rest all are similar.
4. Then p models are fitted using above p bootstrap samples and these are combined using average of outputs.

Boosting: Boosting is meta-learning algorithm. It works on class of weak learners to convert them into strong. The classifiers that are only a bit correlated with true classification are called weak learners and which are highly correlated with the true classification are called strong learners. The main principle of boosting ensemble is to fit weak learners onto version of data that is repeatedly modified. There are a number of boosting algorithms like AdaBoost, LogitBoost, medaBoost etc. Boosting algorithms are applied to improve the bias and variance.

The working of algorithm is explained below.

1. Initially, weights $M_{1,}$.....M_N are assigned to all the samples (training). The weights assigned to each samples are

$$\boldsymbol{M_i = \frac{1}{N}}$$

2. Firstly very weak learner is chosen out of all the learners and is then trained on the original data.

3. Reweighing of samples is done. The weights of the samples that are incorrectly classified in previous step are increased and that of correctly classified weight is reduced.
4. For each iteration, same procedure is done that means priority is given to incorrectly classified samples and each learner is forced to concentrate on the samples misclassified in previous step.

The results obtained using steps proposed in the above framework are discussed in next chapter.

Chapter 5

Testing and Results

The experiment is done on KDDCUP'99 database to build an intrusion detection model. The network intrusion audit data in contained in this dataset. The proposed framework has been tested on 10% of KDDCUP'99 dataset in which there are total 494,021 connection records. The rules for the proposed framework are generated through training dataset and testing is done on test data to validate the performance of rules.

In the results five attack categories are numbered as:

5.1 Evaluation Parameters

Performance of an IDS is measured using confusion matrix. The information about actual and predicted classification is present in confusion matrix. Confusion matrix is also called as contingency table.

Confusion table consists of following parameters:

- *True positive (TP)*: True positive refers to the number of correct classifications or predictions that indicate them as intrusive or malicious.
- *True negative (TN):* It refers to the number of detected normal instances that are actually normal.
- *False positive (FP):* False positive refers to number of detected attack instances which are actually normal.
- *False Negative (FN)*: False negative is number of detected normal instances which are taken as attack.

Table 5.1 shows confusion matrix and its outcomes

Table 5.1 Confusion Matrix

		Predicted Class	
		Positive	Negative
Actual Class	Positive	TP	FN
	Negative	FP	TN

The performance or accuracy is measured in terms of detection rate and false alarm rate. True positive is when activity detected as normal and is actually normal so it should be kept high. False positive occurs when there is no attack but detected as attack and it should be low.

Various standards are evaluated to measure the performance of the IDS for this confusion matrix. These are:

- Accuracy: Accuracy is defined as ratio of correctly classified instances to the total number of instances. Accuracy is also called as classification rate and its formula is:

$$Accuracy = \frac{correctly\ classified\ instances}{total\ number\ of\ instances}$$

$$Accuracy = \frac{TN + TP}{TN + FP + FN + TP}$$

- Detection rate (DR): It is the ratio of correctly detected attacks to the total number of attacks. This is also called as True Positive Rate or Recall. Formula is:

$$DR = \frac{No\ of\ correctly\ detected\ attacks}{total\ no\ of\ attacks}$$

$$DR = \frac{TP}{TP + FN}$$

- True Negative Rate: It is the ratio of total number of detected normal attacks to the total number of normal instances.

$$TNR = \frac{TN}{TN + FP}$$

- False Positive Rate (FPR): FPR is the fraction of normal instances detected as attack to the total normal instances:

$$FPR = \frac{FP}{TN+FP}$$

Or

$$FPR = 1 - TNR$$

- False Negative Rate (FNR): FNR is ratio of total number of attack instances detected as normal to the total number of attack instances.

$$FNR = \frac{FN}{FN + TP}$$

Or

$$FNR = 1 - TPR$$

5.2 Results

The results obtained by applying different classification algorithms on the KDD dataset are shown in the following figures.

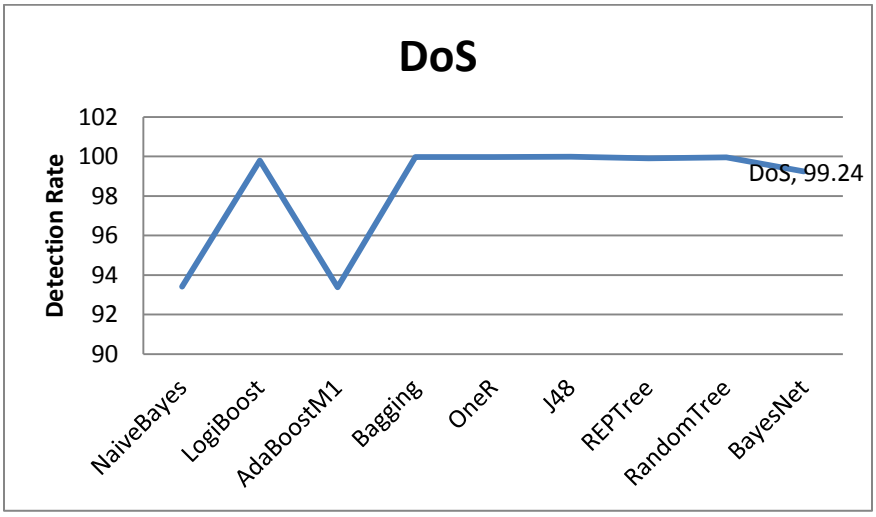

Figure 5.1 Detection Rate of DoS attack

Figure 5.1 shows detection rate of DoS attack using different algorithms. Detection rate of DoS attack is maximum for J48 algorithm which is 99.9%. All the algorithms give good performance for DoS attack detection. AdaBoostM1 has less accuracy as compared to others.

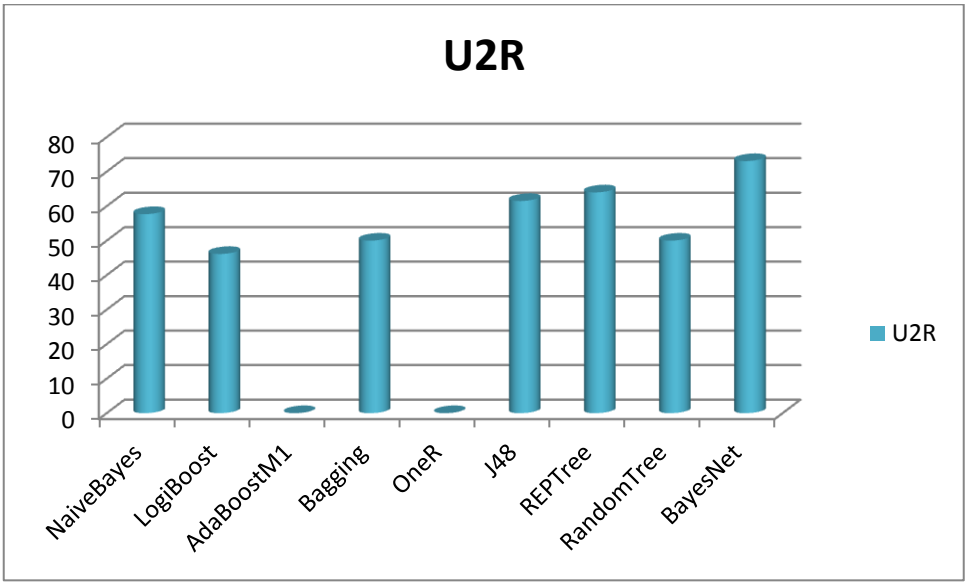

Figure 5.2 Detection rate of U2R attacks

Figure 5.2 shows detection rate of U2R attacks. Most of the algorithms cannot detect U2R attacks properly as the training data available set for this type is very less. In dataset under consideration, U2R attack data is about 0.01% of the total data. Bayesnet algorithm give classification accuracy of 73% in U2R attacks. AdaBoostM1 and OneR are unable to detect U2R attacks.

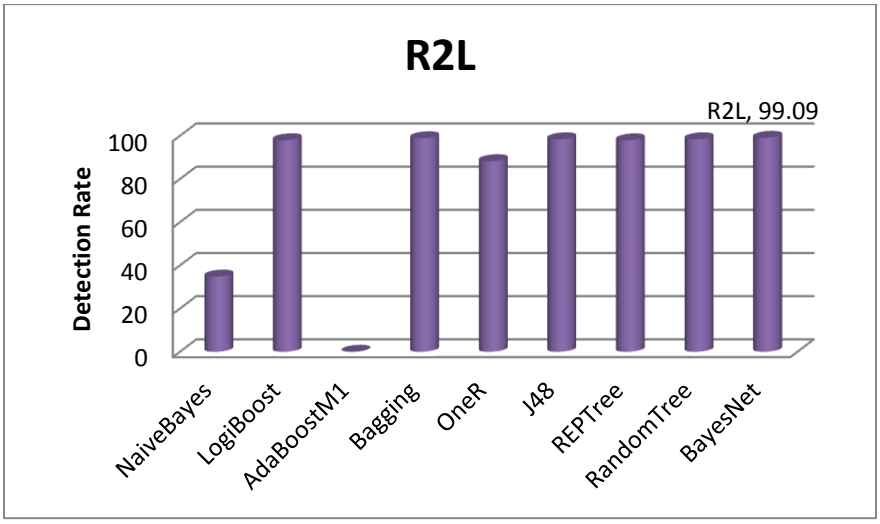

Figure 5.3 Detection rate of R2L attacks

The detection accuracy achieved using different algorithms with R2L attacks is shown in figure 5.3. BayesNet algorithm gives best accuracy with these types of attacks. NaiveBayes and AdaBoostM1 gives poor performance in case of R2L attacks.

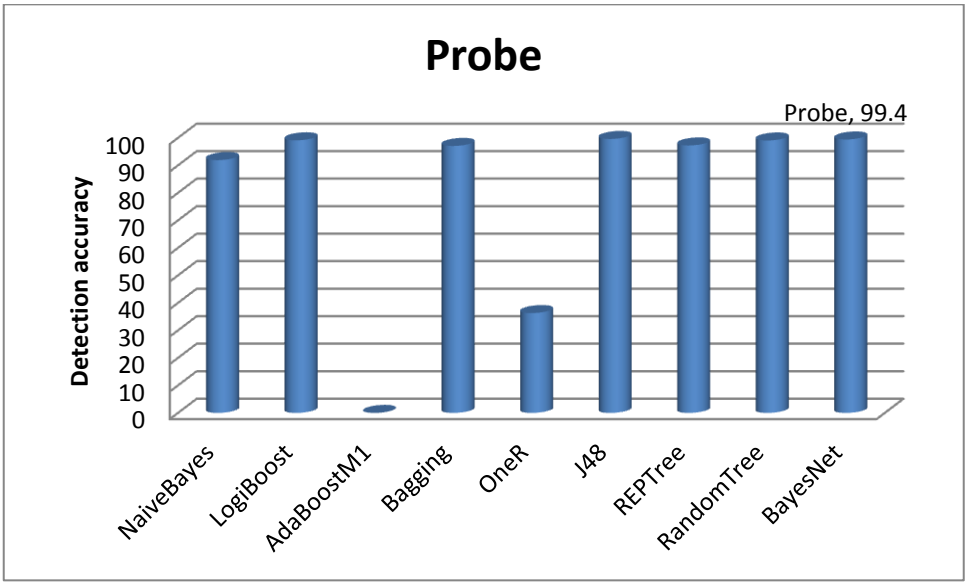

Figure 5.4 Detection accuracy for Probe attacks

Figure 5.4 represents detection accuracy for Probe attacks. For R2L attacks, best accuracy is achieved using BayesNet algorithm. The performance of J48 and BayesNet algorithms is best in these. The accuracy of almost all algorithms is good for normal and DoS attack classes.

Ensemble learning: Ensemble method is applied for J48, NaiveBayes, BayesNet and OneR algorithm. Fusion of these is done with bagging , AdaboostM1 and MultiBoost algorithms.

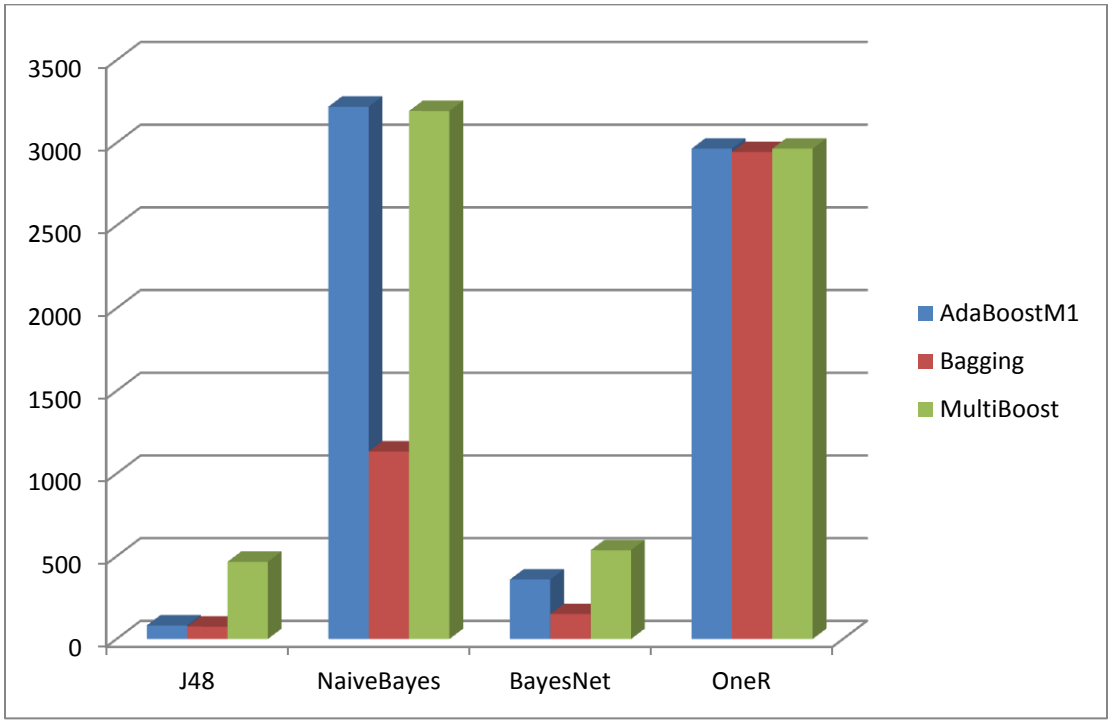

Figure 5.5 Number of incorrectly classified instances using ensemble learning

When the fusion of classification boosting algorithm is done J48 algorithm gives best performance. Figure 5.5 shows number of incorrectly classified instances when ensemble learning is applied. When fusion of J48 and bagging is done then least number of incorrect classifications is present i.e. only 75. In case of NaïveBayes algorithm maximum numbers of incorrect classifications are present.

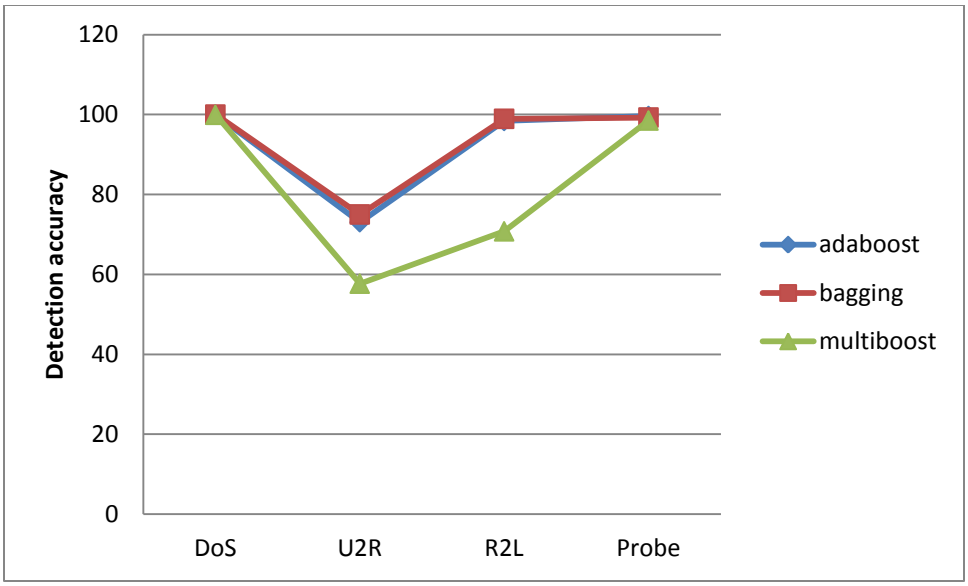

Figure 5.6 Detection accuracy of ensemble learning with J48

Figure 5.6 shown above presents results achieved when applying J48 with other techniques or algorithms. J48 algorithm with bagging gives best results.

Table 5.2 Comparison of classification and ensemble learning

Algorithm	**DoS**	**U2R**	**R2L**	**Probe**
J48	99.18	61.5	98.87	99.4
J48+bagging	99.9	75	98.93	99.26

Table 5.2 shows the comparison of J48 and J48 with bagging. Detection accuracy has improved by applying ensemble learning. In case of U2R attacks detection accuracy has improved from 61.5 % to 75%. Thus with ensembling results are improved.

Chapter 6

Conclusions and Future Scope

6.1 Conclusion

In this report, cyber attack classification system is presented which classifies the connections either as normal or of specific attack type. Two approaches are used in this work. The first approach is classification technique which classifies the attacks to their classes. Different classification algorithms applied for this approach are NaiveBayes, BayesNet, J48, RandomForest etc. The approaches which give high detection rate for two J48 and BayesNet. The individual class results are shown in results part. The second approach is ensemble approach where combination of multiple techniques is applied to improve the results. The combination of multiple techniques gives better result as compared to the single technique. With the classification algorithms, boosting algorithms AdaBoost, bagging and multiboost are used. Adaboost algorithm provides a good bias and variance reduction capability. The intrusion detection system classifies the connections records to the proper class and the results by applying different algorithm are shown in the tables. When bagged boosting is applied with J48 algorithm, best results are achieved. The normal and abnormal behavior of network can be classified with high accuracy using classification rules thus leaving the system with lesser inconsistencies. Therefore, combination of different approaches i.e. ensemble learning gives better performance.

6.2 Future Scope

The major concern of any intrusion detection system is to reduce the false alarm rate and to increase the detection rate. The proposed approach increases the detection rate but can be improved to other parameters also. In the proposed approach the average rule is used for classification, other parameters like “belief” function and “maximum of probability” can also be used to improve the classification accuracy. The ensemble methods provide a trade-off between false alarm rate and generalization capabilities, combination methods

reduce overall false alarm rate but in that case may also decrease the generalization capabilities. This aspect can be further investigated to deploy IDS.

In future attack detection system can be built which combines different cyber attack detection system and which may give 100% accuracy for all classes of attacks.

References

[1] G. Luger, R. Heady, A. Maccabe, and M. Servilla. "The architecture of a network-level intrusion detection system", Department of Computer Science, College of Engineering, University of New Mexico, 1990.

[2] R. P. Lippman et al., "Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation", In *Proceedings*, IEEE *DARPA Information Survivability Conference and Exposition, DISCEX'00.* vol. 2, pp. 12-26, 2000.

[3] SANS Institute, "Intrusion Detection Systems: Definition, Need and Challenges",http://www.sans.org/reading_room/whitepapers/detection/intrusion-detection-systems-definition-challenges_343, SANS Institute, 2001.

[4] T. Mahbod, et al., "A detailed analysis of the KDD CUP 99 data set", *Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009*.

[5] http://www.zentut.com/wp-content/uploads/2012/10/kdprocess.png

[6] A. K. Pujari, Data mining techniques, Universities press, Second edition, pp. 51-57, 2001.

[7] Y. Singh and A. S. Chauhan, "Neural networks in data mining", *Journal of Theoretical and Applied Information Technology,* vol. 5, no. 6, pp. 36-42, 2009.

[8] S. Peddabachigari et al. "Modeling intrusion detection system using hybrid intelligent systems", *Journal of network and computer applications,* vol. 30, no. 1, pp. 114-132, 2007.

[9] D. E. Denning, "An intrusion-detection model", *IEEE Transactions on Software Engineering,* vol. 2 , pp. 222-232, 1987.

[10] P. A. Porras and A. Valdes, "Live Traffic Analysis of TCP/IP Gateways", In *NDSS,* 1998.

[11] H. Debar, M. Dacier and A. Wespi, "Towards a taxonomy of intrusion-detection systems", *Computer Networks,* vol. 31, no. 8, pp. 805-822, 1999.

[12] H. Debar, M. Dacier and A. Wespi, "A revised taxonomy for intrusion-detection systems", In *Annales des télécommunications*, vol. 55, no. 7-8, pp. 361-378. Springer-Verlag, 2000.

[13] R. Sekar, A. Gupta, J. Frullo, T. Shanbhag, A. Tiwari, H. Yang and S. Zhou, "Specification-based Anomaly Detection: A New Approach for Detecting Network Intrusions", In P*roceedings of the 9th ACM conference on Computer and communications security*, ACM, 2002.

[14] S.J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. K. Chan, "Cost-based modeling for fraud and intrusion detection: Results from the JAM project". In IEEE *DARPA Information Survivability Conference and Exposition, 2000. DISCEX'00. Proceedings*, vol. 2, pp. 130-144. IEEE, 2000.

[15] J. McHugh, "Testing intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory", *ACM transactions on Information and system Security,* vol. 3, no. 4, pp. 262-294, 2000.

[16] A. Veda, P. Kalekar, and A. Bodhankar, "Intrusion Detection Using Data mining Techniques", *Report IIT Bombay, 2006*.

[17] A. Lazarevic, L. Ertoz, V. Kumar, A. Ozgur, and J. Srivastava, "A Comparative Study of Anomaly Detection Schemes in Network Intrusion Detection", *SDM*, 2003.

[18] K. Kendall, *A database of computer attacks for the evaluation of intrusion detection systems*, Massachusetts Inst of Tech Cambridge Dept of Electrical Engineering And Computer Science, 1999.

[19] "Nsl-kdd data set for network based intrusion detection systems." Available on: http://nsl.cs.unb.ca/NSL-KDD/, March 2009.

[20] G. I. Webb, "Multiboosting: A technique for combining Boosting and Bagging", *Machine Learning*, vol. 40, pp. 159-196, 2000.

[21] W. L. Salvatore, J. S. Kui, and W. Mok, "A Data Mining Framework for Building Intrusion Detection Models", *IEEE Symposium on Security and Privacy*, 1999.

[22] Y. Chen, K. Hwang, and W. Ku, "Collaborative detection of DDoS attacks over multiple network domains", *IEEE Transactions On Parallel And Distributed Systems*, vol. 18, no. 12, pp. 1649-1662, Dec. 2007.

[23] S. S. Kaushik and P. R. Deshmukh, "Detection of Attacks in an Intrusion Detection System", *International Journal of Computer Science and Information Technologies*, vol. 2, no. 3, pp. 982-986, 2007.

[24] H. Naguen, and Y. Choi, "Proactive Detection of DDoS Attacks Utilizing k-NN Classifier in an Anti-DDos Framework," *International Journal of Electrical and Electronics Engineering*, vol. 4, no. 4, pp. 247-252, 2010.

[25] Z. Liu, and Y. Lai, "A data mining framework for building intrusion detection models based on IPv6", *Advances in Information Security and Assurance*, Springer Berlin Heidelberg, pp. 608-618, 2009.

[26] S. Singh and S. Silakari, "Generalized Discriminant Analysis algorithm for feature reduction in Cyber Attack Detection System", *International Journal of Computer Science and Information Security*, vol. 6, no. 1, 2009.

[27] V. Balaji, and K. Varalakshmi, "Differentiating Network Attacks using C4.5 Algorithm for Multiboosting", *International Journal of emerging Technologies in Computational and Applied Sciences* (IJETCAS), pp. 231-235, 2013.

[28] T. Lappas, and K. Pelechrinis, "Data mining techniques for (network) intrusion detection systems", (white paper) *Department of Computer Science and Engineering UC Riverside, Riverside CA* 92521, 2007.

[29] W. Daunicht, "Autoassociation and novelty detection by neurornechanics", *Science*, vol. 253, no. 5025, pp. 1289-1291, 1991.

[30] D. Yeung and C. Chow, "Parzen-Window Network Intrusion Detectors", *In Proceedings 16th International Conference on Pattern Recognition*, vol. 4. IEEE, 2002.

[31] L. Didaci, G. Giacinto and F. Roli, "Ensemble Learning for Intrusion Detection in Computer Networks", *Workshop Machine Learning Methods Applications, Siena, Italy*, 2002.

[32] A. A. Freitas, "A survey of evolutionary algorithms for data mining and knowledge discovery", *Advances in evolutionary computing*, Springer Berlin Heidelberg, pp. 819-845, 2003.

[33] Z. Zhou, J. Wu and W. Tang, "Ensembling Neural Networks: Many Could Be Better Than All", *Artificial Intelligence*, vol. 137, no. 1-2, pp. 239-263, Elsevier, 2002.

[34] S. Singh, S. Agrawal, Murtaza, A. Rizvi and R. S. Thakur, "Improved Support Vector Machine for Cyber Attack Detection", *In Proceedings of the World Congress on Engineering and Computer Science 2011* WCECS 2011, October 19-21, San Francisco, USA, 2011.

[35] S. Amari and S. Wu, "Improving support vector machine classifiers by modifying kernel functions", *Neural Networks*, vol. 12, no. 6, pp. 783-789, 1999.

[36] D. Veerwal, N. Choudhary and D. Singh, "Ensemble of Soft Computing Techniques for Intrusion Detection", *Global Journal of Computer Science and Technology Network, Web & Security*, vol. 13, pp. 13, 2013.

[37] M. J. Alhaddad, A. Ahmed, S. M. Halawani, A. H. Altalhi, and A. S. Mashat, "A Study Of The Modified KDD 99 Dataset by using Classifier Ensembles", *IOSR Journal of Engineering*, vol. 2, no. 5, pp. 961-965, 2012.

[38] R. D. Kulkarni, "Using Ensemble Methods for Improving Classification of the KDD CUP '99 Data Set", *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol. 16, no. 5, pp. 57-61, 2014.

[39] A. Patel and R. Tiwari, " Bagging Ensemble Technique For Intrusion Detection System", *International Journal For Technological Research In Engineering*, vol.2, no. 4, pp. 256-259, December, 2014.

[40] G. Kumar and K. Kumar, "The Use of Artificial-Intelligence-Based Ensembles for Intrusion Detection: A Review", *Applied Computational Intelligence and Soft Computing*, pp. 1-21, 2012.

[41] G. Kumar, "Evaluation Metrics for Intrusion Detection Systems - A Study," International Journal of Computer Science and Mobile Applications, vol. 2, no. 11, pp. 11-17, 2014.

List of Publications

Published

C. Rani, S.Goel, " CSAAES: An Expert System for Cyber Security Attack Awareness", *International Conference on Computing, Communication and Automation (ICCCA),* 2015 held at Galgotia University, Greater Noida, UP, 15-16 May, 2015, pp. 242-245, IEEE Explore, DOI: 10.1109/CCAA.2015.7148381.

Communicated

C. Rani, S. Goel, "Improving Attack Detection Accuracy using Ensemble Methods" *International Journal of Innovative Research in Computer and Communication Engineering*, vol 3, no. 7, July 2015.

Video Presentation

This is link to my YouTube video where I present brief summary of my thesis topic:
https://www.youtube.com/watch?v=86321zApLAk

Plagiarism Report

A Framework for Improving Attack Detection Accuracy using Ensemble Methods

ORIGINALITY REPORT

15%	11%	13%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS