A RESOURCEFUL INTRUSION DETECTION SYSTEM FOR MANET USING ROUGH SET THEORY AND SUPPORT VECTOR MACHINE

K.Murugesan
PG Scholar

P.Kanagaraj
PG Scholar

S.Jagatheswaran
PG Scholar

S.Periyasamy
Assistant Professor

**K.S.R College of Engineering,
Tamilnadu, India.**

Abstract— Mobile impromptu Networks has more difficult vulnerabilities compared with wired networks. Mobile impromptu networking (MANET) has become a vital technology in current years attributable to the speedy proliferation of wireless devices. They're extremely at risk of attacks thanks to the open medium, dynamically ever-changing topology and lack of centralized watching purpose. It's vital to look new design and mechanisms to safeguard these networks. Intrusion detection system (IDS) tools are appropriate for securing such networks. The most tasks of IDS is to get the intrusion from collected information. a number of the options of collected information is also redundant or contribute very little to the detection method. Therefore it's essential to pick out the vital options to extend the detection rate. Most of the present intrusion detection systems detect the intrusion by victimization sizable amount of knowledge options collected from network. During this work, we have a tendency to propose anomaly primarily based intrusion observation system to detect the malicious activities by collection statistics from network. Conjointly we have a tendency to use SVM machine learning technique associate degreed Rough pure mathematics that are wont to observe the attacks in an economical approach. Rough pure mathematics preprocesses the feature information to scale back the procedure complexness. The support vector machine is trained by victimization feature set from the Rough pure mathematics for sleuthing abnormal behavior.

Keywords— Mobile ad hoc network, attacks, feature selection, ns2 simulator, support vector machine, rough set theory, intrusion detection system, Rosetta.
.

I.   INTRODUCTION

An ad-hoc network is a self-configuring network of wireless links connecting mobile nodes that forms dynamic topologies and communicate through wireless media. Due to lack of infrastructure, each node in the network can act both as a router and a host. The wireless nature of communication and the characteristics of MANET raise several security problems. The mobile nodes communicate directly with each other and without the aid of access points, and therefore have no fixed infrastructure. They form an arbitrary topology, where the routers are free to move randomly and arrange themselves as

required. Each node or mobile device is equipped with a transmitter and receiver.

They are said to be purpose-specific, autonomous and dynamic. This compares greatly with fixed wireless networks, as there is no master slave relationship that exists in a mobile ad-hoc network. Nodes rely on each other to establish the communication, thus each node acts as a router

Therefore, in a mobile ad-hoc network, a packet can travel from a source to a destination either directly, or through some set of intermediate packet forwarding nodes.[1] Routing protocols between any pair of nodes within an ad hoc network can be difficult because the nodes can move randomly and can also join or leave the network. Ad-hoc networks are highly vulnerable to security attacks and dealing this is challenging task for the developers.

The main reasons for this difficulty are; "Shared broadcast radio channel, insecure operating environment, lack of central authority, lack of association among nodes, limited availability of resources, and physical vulnerability." Generally, when considering the security of a network, we examine it under the headings, availability, confidentiality, authentication, integrity and no repudiation. Availability refers to the fact that the network must remain operational at all times. Confidentiality ensures that certain information is never disclosed to certain users. Authentication is the ability of a node to identify the node with which it is communicating. Integrity guarantees that a message is never corrupted when transferred.

Non repudiation states that the sender of the message cannot deny having sent it. An ad-hoc network needs extra security requirements caused by its lack of proper infrastructure and the dynamic relationship between the nodes in the network. Because of the lack of infrastructure, accountability is very difficult to determine as there is "no central authority which can be referenced when it comes to making trust decisions about other parties in the network." Intrusion is defined as

"any set of actions that attempts to compromise the integrity, confidentiality or availability of resources". Intrusion detection systems (IDS) are mainly used to detect and call attention of suspicious behavior.

1.1 Mobile Ad hoc Networks

A mobile ad hoc network (MANET) is a collection of mobile hosts that can communicate with each other without any pre-established infrastructure. It is a group of wireless mobile nodes in which nodes cooperate by forwarding packets to each other for allowing them to communicate beyond direct wireless transmission range. Each node in the MANET can act as router as well as host. In order to maintain connectivity in a mobile ad hoc network all participating nodes have to perform routing of network traffic.[2] Each device in a MANET is free to move independently in any direction. It can be a standalone network or it can be connected to external networks. Due to the characteristics of MANET, it is susceptible to different security vulnerabilities.

Though various encryption and authentication techniques have been developed, it is not sufficient to handle all the attacks. It is essential to have a detection mechanism for handling the security issues. The success of communication highly depends on other nodes cooperation. Therefore, MANET has the property of rapid infrastructure-less deployment and no centralized controller which makes it convenient to many environments, such as soldiers relaying information for situational awareness on the battlefield, business associates sharing information during a meeting; attendees using laptop computers to participate in an interactive conference; and emergency disaster relief personnel coordinating efforts after a fire, hurricane, or earthquake. The other possible applications include personal area and home networking, location-based services, and sensor networks.

1.2 Intrusion Detection System

Intrusion Detection System (IDS) is to detect the unwanted attempts with high detection rate and low false positive to protect the computer system. It use all data features from network statistics to detect an intrusion. The main aim of intrusion detection is to identify malicious attacks and it has two methods such as anomaly detection and misuse detection. In misuse detection, system detects the attacks by using some well known attack patterns. The limitation of this system is that new type of attack cannot be determined. In anomaly detection approach, system compares the event with normal behavior to find the attacks. It will produce an alarm when there is any mismatch between event behaviors and normal behavior.

The main task of the intrusion detection system is to discover the intrusion from the network packet data or system traffic data. One of the major problems that the intrusion detection system might face is that the packet data or system traffic data could be overwhelming. For wireless network, due to the limited capacity of wireless devices, choosing those features that can best characterize the behavior of network is very important. In wireless network, the way a node communicates with other nodes is by exchanging messages. Therefore, a node's behavior can be obtained by monitoring the network traffic.

Each node monitors its neighboring nodes' network traffic and built a profile during offline training. Then the profile is used as a threshold to detect abnormal behavior in the network. System selects all possible features as the object of the monitoring. That is suitable for a small wireless network, which has only a few nodes. But it requires a big amount of capacity for very large network.

II. RELATED WORK

In the last few years researchers have been actively exploring many mechanisms to ensure the security of control and data traffic in wireless networks. Intrusion detection system deals with huge amount of data, which contains irrelevant and redundant features causing higher resource consumption as well as poor detection rate.

John Felix and Amitabha Das have proposed [3] a uniqueness of security vulnerabilities in ad hoc networks has given rise to the need for designing novel intrusion detection algorithms. The incoming data have been trained by SVM [1] for the future selection the number of features and the training data size are reduced by the process of association and filtering, respectively. They use linear machine learning method, FDA [1], is used to check whether the chosen training data are always optimal. This approach handles only the sinking behavior [1].

D.K.Srivastsva and K.S.Patnai have proposed a data classification as a Rough SVM [2] approach Rough-SVM,which makes great use of the advantages of Support Vector Machine's greater generalization performance and Rough Set Theory [2]in effectively dealing with vagueness and uncertainty information. Classification accuracy using Rough-SVM is much better than general SVM and general RSES [2] method.

Roman W.Swiniarski a and Andrzej Skowron b emphasize the role of basic constructs of rough set approach [3] in feature selection , namely reducts and their approximations, including dynamic reducts. The sequence of data mining steps, including

application of SVD, histograms, PCA, and rough sets for feature extraction and feature selection [3] in R.W. Swiniarski, A. Skowron pattern recognition letter 24(2003)833-849 847 designing of neural network classifiers for face images and mammographic images.

AnazidaZainal, MohdAizainihave studied to investigate the effectiveness of Rough Set Theory in identifying important features in building an intrusion detection system [4]. Rough Set [4] also used to classify the data in Feature Selection using Rough Set in intrusion Detection.

This paper has presented a preprocessing part of an intrusion detection system if both accuracy and speed are to be achieved Rough Set has demonstrated its potential capability of selecting an optimum feature subset. The results obtained indicate that the feature subset proposed by Rough Set was robust and has consistent performance throughout the experiment. With the analysis of the above given papers we have used a technique that involves an effective way of feature selection process for intrusion detection in MANET.

III. OVERVIEW OF ATTACKS

Conceptually, similar to WLAN and wired networks, attacks on ad hoc networks can be classified into passive attacks and active attacks. Passive attacks refer to eavesdropping on the network traffic, and they are difficult to detect by their very nature. Malicious nodes initiate active attacks, and they can be carried out against mobile nodes, or communication protocol and infrastructure at different layers. Flooding attack spreads extra data or fake routing control packets into the network. Depending on the routing protocol, the attacker can render single-path or multi-path flooding attacks.

Black hole attacks publicize untrue routing control information. For example, in an on-demand routing protocol, attacker may advertise itself being the best path to the destination node during the path-finding process. As a result, it intercepts all data packets being sent to the destination node. Warm hole attack directs its packets from one point to another. These packets may be replayed from the far end of the wormhole. Byzantine attacks negotiate intermediate nodes conducts attacks such as black hole and packet dropping or to create routing loops.

Packet dropping attack maliciously drops data packets. The attacker may deploy different dropping patterns. This makes itself the most difficult attack to detect. Spoofing attack spoofs a legitimate user's identity or creates misleading content to trick the victim into making an inappropriate security-relevant decision. Routing protocol attack targets against routing protocols by rushing routing control packets, Poisoning routing table, injecting or replicating packets, etc.

IV. IDS ARCHITECTURE

IDS collect the network statistics from trace file as an audit data. The statistics from collection module will send to rough set theory to reduce the features for finding the attack efficiently. Support vector machine has trained by the necessary features from Rough Set Theory. It will classify the attacks and normal behavior of the network while sending the packet. Fig 1 shows the overall architecture of proposed IDS. The architecture consists of four modules namely,

1. Data collection
2. Data reduction
3. Training and
4. Classification

4.1 Data Collection

The data collection module collects the data from network. The collection module in the IDS architecture monitors the events and packet delivery time, traffic, and topology statistics and records the feature values. The unwanted events are removed from the collected data. List of important features are shown below.

Data Packets

1- NBDataSend,

2- NBDataRecv,

3- NBDataDrop,

4- NBDataFwd

RREQ packets

5- NBRREQSend,

6- NBRREQRecv,

7- NBRREQDrop,

8- NBRREQFwd

RREP packets

9- NBRREPSend,

10- NBRREPRecv,

11- NBRREPDrop,

12- NBRREPFwd

RERR packets

13- NBRERRSend,

14- NBRERRRecv,

Corresponding Author: Mr.K.Murugesan , PG Scholar, KSR College of Engineering, Tamilnadu , India   38

15- NBRERRDrop,

16- NBRERRFwd

4.2 Data Reduction

In data reduction module there are two processes namely transformation and feature selection. Features from data collection are in non readable format and it is difficult to understand. So transformation converts the statistics from network into packet data.

Then the packet data features are passed to rough set theory for selecting best features to train the SVM. The process of feature selection and data reduction will be done using Rough set theory. Features selected by the Rough Set theory are:

4.2.1 Rough Set

Rough set theory is a new statistical tool for jagged data analysis. It has an overlap with many other theories dealing with imperfect knowledge, e.g., evidence theory, fuzzy sets, Bayesian inference and others. It can be also used for feature selection, feature extraction, data reduction, decision rule generation, and pattern extraction etc., identifies partial or total dependencies in data, eliminates redundant data, gives approach to null values, missing data, dynamic data and others. Basic Concepts of Rough Sets are Information/Decision Systems, Set Approximation, educts and Core, Rough membership and Dependency of attributes.

Rough set theory is to decrease analysis data and increase executing performance. It can be used to filter features and uses support vector machine to analyze intrusion behavior modules. It has become a valuable tool in the resolution of various problems, such as: representation of uncertain or imprecise knowledge; knowledge analysis; evaluation of quality and availability of information with respect to consistency and presence a lot of date patterns; identification and evaluation of date dependency; reasoning based an uncertain and reduction of information data.

Rough set theory is an extension of conventional set theory that supports approximations in decision making. It is an approximation of a vague concept (set) by a pair of precise concepts, called lower and upper approximations, which are a classification of the domain of interest into disjoint categories. The lower approximation is a description of the domain objects which are known with certainty to belong to the subset of interest, whereas the upper approximation is a description of the objects which possibly belong to the subset.

Rough Set Theory is a mathematical tool for approximate reasoning for decision support and is particularly well suited for classification of objects. It can also be used for feature selection and feature extraction [6].The main contribution of rough set theory is the concept or redacts. A redact is a minimal subset of attributes with the same capability of objects classification as the whole set of attributes. Reduct computation of rough set corresponds to feature ranking for IDS.

The selected features from rough set theory are used to train the SVM classifier.

Table 1 : Selected features

Method	Selected Features
Rough set	NBRREQSend, NBRREQRecv, NBRREPSend, NBRREPRecv, NBRREPDrop, NBRERRDrop

The SVM classifier was trained by using SVM training algorithm.

4.3. SVM Learning

SVM learning is a process in which a set of parameters is trained to classify an unknown behavior. To introduce the concept, let us consider the following function. The method determines a linear function f(x).which classifies the normal and a malicious activity by using the sign of the function f. plus sign indicates the normal behavior and minus sign indicates abnormal behavior.

$$F(x) = q.x+d$$
$$F(x) = \{>0 \text{ normal}$$
$$\{<0 \text{ malicious}$$

In the above function q represents orientation from origin of the hyper plane, x represents the events, and d represents the distance from the origin of the hyper plane and F (x) act as a hyper plane.

Corresponding Author: Mr.K.Murugesan , PG Scholar, KSR College of Engineering, Tamilnadu , India

4.4. SVM Classification

Classification is done by a Support Vector Machine SVM classifies the normal behavior and attacks using kernel function. SVMs also have been the ability to update the training pattern dynamically whenever there is a new pattern during classification .SVM uses a feature called kernel to solve this problem. Kernel transforms linear algorithms into nonlinear ones via a map into feature spaces. There are many kernel functions; some of them are Polynomial, radial basis functions, two layer sigmoid neural nets etc. The user may provide one of these functions at the time of training classifier, which selects support vectors along the surface of this function. SVMs classify data by using these support vectors, which are members of the set of training inputs that outline a hyper plane in feature space.
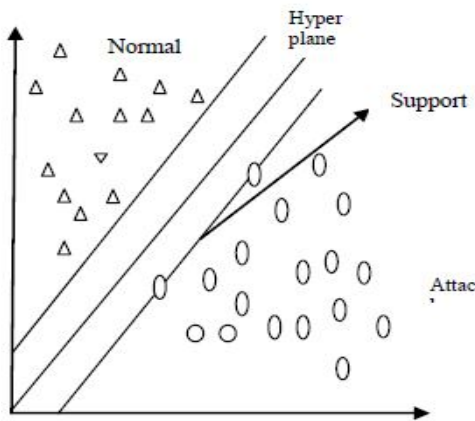

Fig 1: SVM classifier

The implementation of SVM intrusion detection system has two phases: training and testing. The main advantage of this method is speed of the SVMs, as the capability of detecting intrusions in real-time is very important. SVMs can learn a larger set of patterns and be able to scale better, because the classification complexity does not depend on the dimensionality of the feature space.

V. SYSTEM IMPLEMENTATION

The IDS uses NS-2 simulator under FEDORA environment for simulating the attacks in mobile ad hoc networks. The various parameters and its corresponding values of ns-2 simulation are given in table 2.SVM and Rough Set Theory has been implemented by LibSVM and Rosetta

Table 2: Simulation Parameters.

SNO.	PARAMETER	VALUE
1	Routing protocol	AODV
2	Simulation duration	100 seconds
3	Topology	1000 m x 500 m
4	Number of mobile nodes	20
5	Transmission range	250m
6	Traffic type	CBR/UDP
7	Data payload	512 bytes
8	Maximum speed	10 m/s

Fig 2: Architecture of IDS

Corresponding Author: Mr.K.Murugesan , PG Scholar, KSR College of Engineering, Tamilnadu , India   40
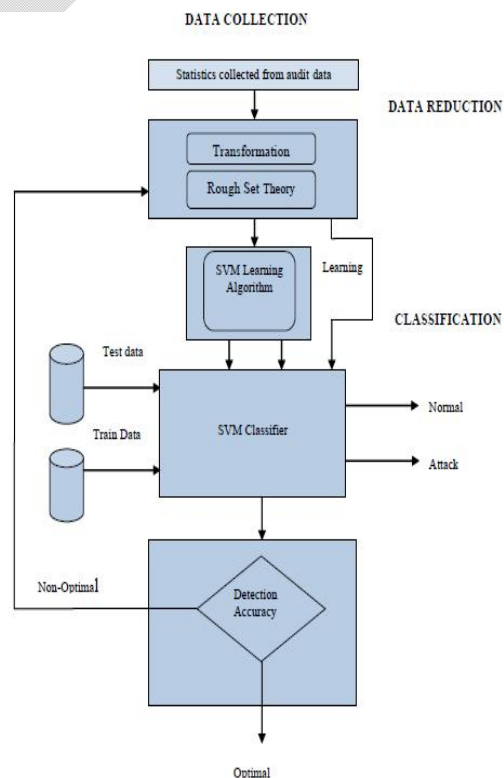
5.1 Simulated Attacks

5.1.1 Flooding attack

An attacker floods extra data or fake routing control packets into the network. Depending on the routing protocol, the attacker can render single-path or multi-path flooding attacks.

5.1.2 Routing disruption

An attacker targets against routing protocols by quickening routing control packets, Poisoning routing table, injecting or replicating packets, etc.

5.1.3 Packet Dropping Attack

In computer networking a packet drop attack is a type of denial-of-service attack in which a router that is supposed to relay packets instead discards them. This usually occurs from a router become compromising from a number of different causes.

5.2 Training Data

The attack data is compared with normal profile and labeled. The classification label consists of two classes namely, normal and abnormal. This data set has collected from trace file which is used to train the Support Vector Machine. The sample data set to train SVM , test data for SVM and output file is shown here.1 represents normal behavior and -1 represents abnormal behavior(attacks).

VI. RESULTS AND DISCUSSION

The experimental results show the performance of detection approach. Detection accuracy is used as a metrics for performance evaluation. Table summarizes the experimental results for the simulated attacks. The system performs well for flooding, because this attack technique is transparent and the consequences are obvious. Whereas packet dropping attack technique is more subtle, and the consequences are less observable.

VII. CONCLUSION

In our work, the anomaly detection with feature selection has been applied for mobile ad hoc networks to detect the intrusions. It uses network layer data to characterize the behavior of mobile nodes. Number of features and the training data size are reduced by the Rough Set Theory to reduce complexity. SVM was trained by the selected features from Rough Set Theory for detecting the intrusions effectively. Attacks were classified from normal behavior by Support Vector Machines. The system has achieved overall detection accuracy of detection with all features is 94.5%. The

observation shows that the selected features gives high detection accuracy compared with all features.

References

[1] John Felix Charles Joseph and AmitabhaDas,"Cross-Layer Detection of Sinking Behavior in Wireless Ad Hoc Networks Using SVM and FDA," IEEE transactions on dependable and secure computing, vol. 8, no. 2 march-April 2011.

[2] D. K. Srivastava and K. S. Patnaik," Data Classification: A Rough - SVM Approach," Contemporary Engineering Sciences, Vol. 3, 2010, no. 2, 77 – 86.

[3] Roman W. Swiniarski and Andrzej Skowron," Rough set methods in feature selection and recognition," Pattern Recognition Letters 24 (2003) 833–849

[4] [AnazidaZainal and MohdAizaini, "Feature Selection Using Rough Set in Intrusion Detection."IEEE TENCON 2006, 14-17th November 2006, Hongkong.

[5] ZbigniewSuraj Chair, "An Introduction to Rough Set Theory and Its Applications," ICENCO'2004, December 27-30, 2004, Cairo, Egypt.

[6] Rung-Ching Chen and Kai-Fan Cheng,"Using Rough Set and Support Vector Machine for Network Intrusion Detection, "International Journal of Network Security & Its Applications (IJNSA), Vol 1, No1, April 2009.

Corresponding Author: Mr.K.Murugesan , PG Scholar, KSR College of Engineering, Tamilnadu , India