TWO PHASE INTRUSION DETECTION SYSTEM USING
EVOLUTIONARY ALGORITHM FOR MANETs

R. Thanuja¹ and A. Umamakeswari²

¹Department of Computer Science Engineering, School of Computing, SA STRA University, Tirumalaisamudram, Thanjavor, Tamil Nadu, India
²School of Computing, SA STRA University, Tirumalaisamudram, Thanjavor, Tamil Nadu, India

E-Mail: thanuja.r@cse.sastra.edu

ABSTRACT

Network security is one of the primary concerned now days for large organizations. MANETs are deployed in challenging real-world environments for example, disaster management, civilian emergency operation and unmanned aircraft operation. Providing Security in this type of environment is the most challenging problems as the operation environment of such network is usually is difficult to decide. This will lead to any malicious activity that will affect the operation of the networks. Various modifications are being applied to IDS regularly to detect new attacks to handle them effectively. In this paper, we have designed a effective Intrusion Detection System (IDS) to detect the attacks in a MANET that uses modified fuzzy min max neural network in a effective manner. The aim of the proposed work is to improve the detection accuracy ratio and reduce the false positive rate. IDS is designed as a two phases technique, in which the first phase use the fuzzy min max neural network algorithm to train 60% data and test 40% data of the network. In the second phase, the modified fuzzy min max neural network algorithm is applied to 60% of train data +20% test data to form a new train data set and test the remaining 20% data collected. The proposed two phase intrusion detection method was evaluated using KDD cup data set. The experimental results showed that the proposed method is better than the other methods in terms of the detection rate, accuracy and reduced false positive rate.

Keywords: intrusion detection system, MANET, neural networks, attacks, KDD data set.

1. INTRODUCTION

A mobile ad hoc network (MANET) is a wireless network with the feature of self organization and self configuration network, so that it can quickly form new networks without the help of a fixed infrastructure. Node within the transmission range of each other can communicate directly over wireless communication and those not fall within the range can use other nodes to communicate. Each node must act as both a sender and a receiver at the same time. The network topology frequently changes due to the mobility of mobile nature of the nodes as they will move to any place within the range. One of the major advantages of wireless networks is its ability to allow data communication between different parties and still maintain their mobility. The communication is limited to the distance of transmission range. The node cannot communicate with each other node when the distance between the two nodes is out of the communication range. This problem is rectified by allowing intermediate nodes to relay data communications between nodes. The MANET will be classified into two types of networks as single-hop and multihop networks. In a single-hop network, all nodes within the range communicate directly with each other node. In multihop network, nodes rely on other intermediate nodes to transmit if the destination node is out of their radio range.

MANET is capable of auto configuring and nodes can arrange themselves without the help of a fixed infrastructure, where it is used in like battle field areas. Less configuration and fast deployment make MANET ready to be used in emergency circumstances where a network structure is unavailable to install in situation like medical emergency situations.
patterns. The anomaly-based IDS is able detect abnormal activities in network. The hybrid IDS monitors’ network activity of systems according to rules defined in the dataset that describe functionality for security related issues.

Many intrusion detection systems (IDS) have been proposed in the literature for MANET, but all the methods has some drawback. This paper proposes a two phase intrusion detection system using fuzzy min max neural network. In the first phase, the fuzzy min max neural network algorithm is applied to train 60% data and test 40% data. In the second phase, the modified fuzzy min max neural network algorithm is applied to 60% of train data +20% test data to form a train data and test remaining 20% data. This two phase approach gives better detection rate.

This paper is presented as follows: The section two discusses design work of IDS in MANET, Section 3 explain the fuzzy intrusion detection and types of attacks. The Section four describes the proposed work of two phase approach. Section 5 evaluates the performance of proposed work with KDD data set and section 6 provides the experimental results discussion and further improvement in algorithm.

2. RELATED WORK

Alem et al [2] proposed an effective method based on Intrusion Detection using Anomaly Detection. It prevents both attacks by single and multiple black hole nodes. IDAD assumes every activity of a user can be monitored and abnormal activity is reported further to administrator. Black hole attack is identified by providing pre-collected set of anomaly activities, called audit data. The collected is given to the IDS system, which is able to compare every activity with audit data. If any activity of a node is out of the activity listed in the audit data, the IDS system isolates the particular node from the network. The reduction of the number of routing packets in turn minimizes network overhead and facilitates a faster communication. But the drawback is that if neighbor node gives false information then this solution gives more delay in the network.

The author [3] proposed a solution for analyzing and improving the security of AODV routing protocol against black hole attack. The packet delivery ratio is increased by 81.812% in presence of black hole attack compared to AODV and there is 13.28% rise in end-to-end delay. The disadvantage of this method it fails to detect collaborative black hole attack.

An algorithm presented in [4] to detect the black hole attack in a MANET based on the preprocessor called Pre_Process_RREP and it is simple and does not change workings of either intermediate or destination node. This solution has more network delay and cannot detect cooperative black hole Nodes.

Khattak et al [5] provide a method to get rid off black and gray hole attacks by eliminating the first and the second shortest path for data packets transmission. Secure routing procedure helps to avoid gray hole attacks to choose unwanted path from networks. Next providing security towards for data integrity and further detection of malicious node on the safe route.

Enhanced Adaptive Acknowledgment (EAACK) was proposed by Shakshuki et al [6]. It is an intrusion detection system that is capable of detecting malicious nodes in presence of false misbehavior report. EAACK relies on ACK for end to end acknowledgement scheme, Secure ACK (S-ACK) to detect misbehavior in presence of receiver collision and MRA (Misbehavior Report Authentication) to know whether the destination received the reported missing packet through a different route and digital signature for signing all acknowledge packets and verifying until they are accepted.

Monita and Ningrinla proposed intrusion detection using fuzzy logic [7]. Fuzzy Logic approach is used for three types of symptoms. First symptom, when the number of packets dropped is greater than threshold. Second symptom, when the number of packets dropped is greater than threshold and the number of packets dropped for the particular destination is greater than another threshold value which called Dest Threshold. Third symptom, when the number of packets drop is greater than threshold and the number of packets drops for the particular source is greater than DestThresold.

Fuzzy bases intrusion detection was proposed by [8]. This system efficiently detect black and gray hole attacks using fuzzy logic technique. The system also contains IPS mechanism, which gets input from fuzzy technique and provides the secure data communication over the network.

3. BACKGROUND

3.1 Fuzzy intrusion detection

Fuzzy logic is used in intrusion detection since 90’s because it is able to deal with uncertainty and complexity which is derived from human reasoning [9]. By the help of fuzzy variables or linguistic terms, intrusion detection features can be viewed easily and decision of normal and abnormal activity in the network are based on its fuzziness nature that can identify the degree of maliciousness of a node instead of yes or no conditions [7] [10]. IF-then else based fuzzy rules are used to define all situations in the network for identifying the attacks or intrusions. The fuzzy rule based system is known as fuzzy interference system (FIS) that is responsible to take decisions.

Manoj V. et al. [11] presented a scheme based on certification authority (CA) and fuzzy logic for MANETs. Some central node is authorized by service provider for assigning the keys to source node which is going to request in the network called certification authority nodes and with the help of trust agent, direct and recommended trust values are obtained periodically. Direct and recommended trust values are calculated from direct observation of one hop away neighbors with the help of algorithms. A proposed fuzzy logic based analyzer used to evaluate the value of a requested node based on the computed fuzzy table. If requested node is trusted then it would get the certification otherwise not.
Sarah and Nirkhi [12] introduced fuzzy logic based approach for forensic analysis to detect the distributed denial of service attacks (DDoS) in Manets. They suggested use of forensic analysis for intrusion detection because it is able to gather digital evidences from any system which has been compromised. It can reconstruct the compromised system and identify the location of attacker. This paper uses fuzzy Logic approach to analysis based on dynamic source routing (DSR) protocol. Three steps are followed to get the result as a forensic report: first capture the packet history file then analyzing log files algorithm and at last presenting the conclusion in terms of report.

3.2 IDS Attacks

Intrusion detection systems (IDS) are build in to monitor misuse and attacks on information systems. It will have a set of historical profiles or recorded profiles for users, matches an audit record with appropriate profile. The updating will be done on timely basis and triggers the alarm whenever it found the threat. An IDS does not usually perform any action to prevent intrusions; its main function is to alert the system administrators that there is a possible security violation; as such it is a proactive tool rather than a reactive tool. IDSs are classified into two types: host based IDS and network based IDS. HIDS monitors all the activity on a single system on network. It will check that no security violations are happened in the network. NIDS monitors activities on a whole network and analyzes traffic for any security violated rules.

Attacks will be broadly classified as:

- DOS: denial of service
- Remote to local attack: intruder access from a remote machine
- User to root attack: Intruder access to local user name
- Probing: Eavesdropping in to network and activity monitoring

4. MATERIALS AND METHODS

In this section the proposed two phase intrusion detection system using proposed algorithm is explained. Figure-1 shows the architecture of the system. Figure-2 shows the two phase algorithm.

![Diagram](image-url)
The main components of the system model are as in the following:

**Data collection:** In this process, the KDD cup data set is collected. The collected data is divided into train set (60%) and test set (40%) for further process. (Section 4 gives detailed description of the data set).

**Preprocessing:** In this process, the collected data is preprocessed. The data set contains some missing values. So the missing value record must be discarded. The data set include various types of attack like Neptune, ipsweep, saten, smurf etc. These attacks are grouped into four categories: dos, u2r, r2l and probe. Table 1 shows the attack categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Attack type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denial of service</td>
<td>Apache2, back, land, mail bomb, Neptune, pod, Process table, smurf, teardrop, udpstorm,</td>
</tr>
<tr>
<td>U2R</td>
<td>buffer_overflow, httprunnel, loadmodule, perl, ps, rootkit, sqlattack, xterm</td>
</tr>
<tr>
<td>R2L</td>
<td>ftp-write, guess_password, imap, multihop, named, phf, sendmail, snmpgetattack, snmpguess, spy, warezclient, warezmaster, worm, xlock, xsnoop</td>
</tr>
<tr>
<td>Probe</td>
<td>ipsweep, mscan, nmap, portsweep, saint, satan,</td>
</tr>
</tbody>
</table>

**Phase-I Fuzzy min max neural network:** In this step, the training data set (60%) is trained using FMM, and the test data set (40%) is classified as 4 types of attack.

Simpson (1992) proposed fuzzy min max neural network classification. The algorithm uses the hyper box fuzzy sets. A hyper box is defined as a region of dimensional pattern having length of n space and it having patterns with membership formula. The hyper box is defined by its minimum and maximum points in the network. The membership function is with respect to these hyperbox minimum-maximum points. The maximum value to which a pattern fits in the hyperbox defines function of membership. Suppose the input pattern of n-dimensions a unit cubes I^n is defined. Over the set and the membership value ranges between zero and one defines over the network. Set of patterns is defined in hyperbox has the value range from zero to one. The definition of each hyperbox fuzzy set B_j is:

\[ B_j = \{X, V_j, W_j, f(X, V_j, W_j)\} \quad \forall X \in I^n \]  

(1)

Where V and W are the points defined with lower and higher values. Figure-3 illustrates the minimum and maximum points in a 3D box.

**Algorithm-1**

*Input: KDD Cup Data Set D*

*Output: Detected Attacks*

1. Split D into Train set-Tr (60%) and Test set-Te (40%)
2. Remove missing Attribute Record
3. Group similar attack behaviors (DoS, Probe, U2R, and R2L)
4. Intrusion Detection-I

   Detected Result-I = FMMN (Tr, Te)
5. Split Detected Result –Te1 and Te2
6. Add Tr1 = Tr + Te1
7. Intrusion Detection-II

   Detected Result-II = MFMMN (Tr1, Te2)
8. Display Detected Attacks

**Figure-2.** Algorithm-1 Two phase approach.

**Figure-3.** 3D hyperbox.
The algorithm structure is made up of three layers, as in Figure-4. \( FA \) is the input layer, which equals to number of input nodes equals is same as the number of input features. \( FB \) is the hyperbox layer approach. Each \( FB \) node represents a hyperbox fuzzy set created during the learning process of networks. The connections between \( FA \) and \( FB \) nodes are the minimum and maximum points defined over the set. \( FC \) is the resultant layer of the network. It is defined by number of nodes equals to the number of output classes.

![Figure-4. Three layer FMM network.](image)

In FMM, the fuzzy min-max learning methodology is an expansion/contraction process. The training set \( D \) consists of a set of \( M \) ordered pairs \( \{X_h, d_h\} \), where, \( X_h=(x_{h1},x_{h2},...,x_{hn}) \in \Gamma \) is the input pattern and \( d_h \in \{1,2,...,m\} \) is the index of one of the \( m \) classes. The learning process is calculated by taking an ordered pair from \( D \) and it can done recursively, if necessary, to include the input. The constraint should be done recursively, and is calculated as:

\[
n\theta \geq \sum_{i=1}^{n} \left( \max(w_{ji},x_{hi}) - \min(v_{ji},x_{hi}) \right)
\]

where \( 0 \leq \theta \leq 1 \) is the size of box. If it met the condition that the condition value cannot be found, a new box is formed in the network by using new values. This will allow increase in classes as the condition are not satisfied. New classes are added with overlapping with the existing class in the network. The overlap test is performed to check if the overlap is among the same or different classes of networks. If overlapping occurs among the same classes no action will be taken place. But if overlapping occurs among different classes, the contraction process is activated and the overlap is removed.

**Phase -2 Modified fuzzy min max neural network:**

In this section the modified fuzzy min max neural network is applied to train the training data set. The original FMM is slightly modified with some formula. The Algorithm 2 is used to predict the attacks.

**Algorithm -2**

**Modified fuzzy min max neural network**

\( MFMMN \) (Train Set \( Tr \), Test Set \( Te \))

\[
A = Tr \text{ Values (} h*n \text{)}
\]

\[
h = \text{No of Rows in } Tr
\]

\[
n = \text{No of Columns in } Tr
\]

\[
B = \text{No of Boxes (initially } B=1\text{)}
\]

\[
V = \text{Minimum Value Matrix (} h*h \text{)}
\]

\[
W = \text{Maximum Value Matrix (} h*h \text{)}
\]

\[
U = \text{Binary Matrix (} B * \text{ no of class labels)}
\]

Threshold \( = 0.5 \)

1. Apply Fuzzy Concept
   Convert all \( Tr \) values into fuzzy
2. Rule Expansion
3. Rule Overlapping
4. Rule Construction
5. Extract \( B, V, W, U \)
6. For \( k=1 \) to \( h \)
   \[
   uk=\text{Matrix}(1,B);
   \]
   For \( i=1 \) to \( B \)
   \[
   sum=0;
   \]
   For \( j=1 \) to \( n \)
   \[
   sum=sum+(\max(0,1-\max(0,\text{Threshold}*(\min(1,A(k,j)-W(i,j)))))+\max(0,1-\max(0,\text{Threshold}*(\min(1,V(i,j)-A(k,j))))));
   \]
   End for
   \[
   d1=1/(2*n);
   \]
   \[
   d2=d1*sum;
   \]
   \[
   as1=d2*U(i,:);
   \]
   \[
   [a1,a2]=\text{size}(as1);
   \]
   \[
   as2=0;
   \]
   For \( j=1 \) to \( a2 \)
   \[
   as2=as2+as1(1,j);
   \]
   End For
   \[
   uk(1,i)=as2;
   \]
   End For

\[
[M1,I1]=\max(uk);
\]

\[
I2=W(I1,n);
\]

Display Predicted attack \( I2 \)

End For

In the rule expansion the FMM is used the constraints (2), MFMM is used the following formula for rule expansion,

\[
n== \text{HDC} \tag{3}
\]

Where HDC= hyperbox dimension count value. Calculate the no dimension of the hyperbox that the value exceeds the expansion coefficient.
The expansion process is calculated, when all hyperbox dimensions do not exceed the above value.

5. RESULT AND DISCUSSIONS

This section explains the performance evaluation of proposed approach. The Two phase fuzzy min max intrusion detection is implemented using Java (version 1.8), and the experiments are performed on a Intel(R) Pentium machine with a speed 2.13 GHz and 2.0 GB RAM using Windows 7 32-bit Operating System.

The proposed method is calculated using the following metric: False Positive (FP), False Negative (FN), True Positive (TP), True Negative (TN), Recall, Precision, F-Measure and Accuracy.

The KDD cup data set is used to evaluate the performance of the processed method. The data set contains 41 features and is labeled as either normal or attack type, with exactly one specific attack type. Table-2 shows the feature name of KDD cup data. Figure-5 shows the sample data.

<table>
<thead>
<tr>
<th>Table-1. Feature name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature name</td>
</tr>
<tr>
<td>Duration, Protocol type, Service, Flag, src_bytes, dst_bytes, Land, wrong_fragment, urgent, hot, num_failed_logins, logged_in, num_compromised, root_shell, su_attempted, num_root, num_file_creations, num_shells, num_access_files, num_outbound_cmds, is_host_login, is_guest_login, count, srv_count, srv_rate, srv_diff_host_rate, dst_host_serror_rate, dst_host_srv_serror_rate, dst_host_serror_rate, dst_host_srv_rerror_rate, class_label</td>
</tr>
</tbody>
</table>

| Figure-5. Sample records. |

The original data set contains both numerical and categorical attribute. For applying fuzzy concept, the data set must be converted into numerical. The categorical attribute values are converted into numerical attribute value. Figure-6 shows the sample data of categorical to numerical data conversion.

For phase-1, the data set is split into 60% for trained data and 40% for data to be tested. The 60% training set is trained using fuzzy min max neural network. The remaining 40% test set is tested based on the training data.

For phase-2, the 60% of training + 20% of test results in phase-1 are considered as a data set. The remaining 20% test is tested using proposed algorithm.
Table-3 shows the accuracy and error rate of FMM and MFMM in phase-1 and Phase-2 respectively.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Accuracy</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase-1 FMM</td>
<td>77.5%</td>
<td>22.5%</td>
</tr>
<tr>
<td>Phase-2 MFMM</td>
<td>90%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Table-4 shows the result of accuracy and error rate of FMM in both phases. Table-5 shows the result of accuracy and error rate of MFMM algorithm in both phases.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Accuracy</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase-1 FMM</td>
<td>97.5%</td>
<td>0.31%</td>
</tr>
<tr>
<td>Phase-2 MFMM</td>
<td>98.12%</td>
<td>0.18%</td>
</tr>
</tbody>
</table>
6. CONCLUSION AND FUTURE WORK

In this proposed work the two phase intrusion detection is designed using evolutionary algorithm. First the algorithm is applied on 60% training data and 40% testing data. Next, the malgorithm is applied on 60% training data + 20% of testing results in phase-1. The remaining 20% data will be used for testing purpose. The main use of this two approach is improves the accuracy of detection rate. The results showed that the proposed two phase intrusion detection method could improve the IDS in terms of detection performance and detection rate.

The MANETs are highly vulnerable to attacks and new attacks are generated regularly. The KDDCUP is highly dominated by DOS and other major attacks. In future IDS is designed using evolutionary algorithms to detect any new type of attacks and data set is further enhanced to the current environment situation.

ACKNOWLEDGEMENT

We thank the Management of our SASTRA University for their great support to carry out this research work.

REFERENCES


