Design of Soft Computing based Black Hole Detection in MANET

D. Vydeki[1], K. S. Sujatha[1] and R. S. Bhuvaneswaran[2]
[1]*Easwari Engineering College, Ramapuram, Chennai, India*
[2]*Anna Univesity, Chennai, India*

Keywords: Intrusion Detection, Black Hole, Fuzzy Logic, Genetic Algorithm.

Abstract: Mobile Ad hoc Networks (MANETs) are more vulnerable to attacks than the generic wireless network due to the lack of an underlying infrastructure and shared medium. Intrusion Detection System (IDS) provides an enhanced level of security to such networks. Application of Soft computing techniques to the detection process is proved to be more suitable as they have human-like decision-making capabilities. This paper proposes a hybrid intrusion detection system for MANETs that detects black hole attack, by combining anomaly and specification-approaches of IDS. The proposed system aims at designing two different IDS using the two fundamental soft computing mechanisms such as, Fuzzy and Genetic Algorithm (GA). The design of each IDS is tested with simulated MANETs for various traffic conditions. The performance of each system is compared based on the true and false positive rates. The experimental results show that the fuzzy based system produces 81.8% true positive rate and the GA based system results in a 100% efficient detection.

1 INTRODUCTION

A MANET is a collection of mobile platforms or nodes where each node is free to move about arbitrarily (Perkins, 2001). It is a mobile, wireless, multi-hop network that operates without the benefit of any existing infrastructure except for the nodes themselves. Such networks are assumed to be self-forming and self-healing. Routing in such networks is challenging because typical routing protocols do not operate efficiently in the presence of frequent movements, intermittent connectivity, network splits and joins. Moreover, use of wireless links make these networks very vulnerable to security attacks ranging from passive eavesdropping to active interfering.

Various cryptography algorithms and secure routing protocols attempt to prevent attacks to provide the first layer of defense. An additional layer of defense called “Intrusion detection” is often used to protect networks. There are three major approaches to IDS: misuse detection, anomaly detection and specification-based detection. This paper aims at developing a hybrid IDS that combines the latter two methods using computational intelligence techniques such as Fuzzy logic, Genetic algorithm and Neural networks. By developing a combinatorial approach, we get the benefits of both methods that reflect in the detection process.

This paper is organized as follows: Section 2 briefs about the related research work. : Section 3 provides an overview of IDS, AODV and the black hole attack. The design of proposed systems and their performance analysis is dealt in detail in Section 4. Section 5 briefs about the future work and conclusion.

2 RELATED WORK

In the recent past, plethora of research works has been carried out in IDS. Ming-Yang Su has designed IDS in which the three types of wireless nodes are defined: normal, malicious and IDS nodes (Ming, 2011). A manually defined threshold is used in their ABM (Anti-Blackhole Mechanism) function that determines whether a node is a normal, suspicious or black hole node. When a suspicious value exceeds a threshold, a nearby IDS node will broadcast a block message, informing all nodes on the network, to cooperatively isolate the malicious node.

Anup Goyal and Chetan Kumar have suggested a machine learning approach with GA, to identify

Vydeki D., S. Sujatha K. and S. Bhuvaneswaran R..
Design of Soft Computing based Black Hole Detection in MANET.
DOI: 10.5220/0004070802830288
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications and Wireless Information Networks and Systems* (WINSYS-2012), pages 283-288
ISBN: 978-989-8565-25-9

harmful/attack type of connections (Anup, 2010). The algorithm takes into consideration different features in network connections such as type of protocol, network service on the destination and status of the connection to generate a classification rule set. For this experiment, they have implemented a GA and trained it on the KDD Cup 99 data set to generate a rule set that can be applied to the IDS to identify and classify different types of attack connections.

3 IDS, ROUTING AND ATTACK

Intrusion detection technologies focus on detecting malicious activity typically from attackers that have successfully penetrated the perimeter defences. Based on the techniques used cross the security barrier, IDS can be classified into three main categories as follows:

i. ***Misuse Detection:*** In misuse detection, decisions are made on the basis of earlier knowledge of the intrusive process and what traces it might leave on the observed system (Anjum, 2007). Such a system tries to detect intrusion irrespective of any knowledge regarding the background traffic. There are several approaches in the signature detection, which differ in representation and matching algorithms employed to detect the intrusion patterns.

ii. ***Anomaly Detection:*** This technique establishes a "normal activity profile" for a system and flags observed activities that abnormally deviate from the recognized normal usage as anomalies. It must first be trained using normal data before it can be released in an operative detection mode. The main advantage of this model is that it can detect unknown attacks. On the other hand, its disadvantage is that it has high false positive alarm rate when normal user profiles, operating system, or network behavior vary widely from their normal behavior.

iii. ***Specification-based detection:*** Specification-based detection defines a set of constraints that describe the correct operation of a protocol, and monitors the execution of protocol with respect to the defined constraints. This technique may provide the capability to detect previously unknown attacks, while exhibiting a low false positive rate.

The Ad hoc on demand distance vector routing protocol (AODV) is a popular unicast routing protocol that provides quick and efficient route establishment between nodes desiring communication with minimal control overhead and minimal route acquisition delay (Sivaram, 2007). Route discovery with AODV is purely on demand and follows a route request/ route reply discovery cycle. Requests are sent using a Route Request (RREQ) message. Information enabling the creation of a route is sent back in a Route Reply message (RREP). In case where the source node receives multiple RREP messages, it will select a RREP message with the largest destination sequence number value.

AODV is efficient and scalable in terms of network performance, but it allows attackers to easily advertise falsified route information to redirect routes and to launch various kinds of attacks. In each AODV routing packet, some critical fields such as hop count, sequence numbers of source and destination, and RREQ ID, are essential to the correct protocol execution. Any misuse of these fields can cause AODV to malfunction.

One such is misuse is the black hole attack which targets on tampering the hop count and sequence number fields. A node, which poses this attack, changes the hop count value to 1 and the destination sequence number to the largest value. This makes the attacking node to be selected in the route discovery process. The black hole node then participates in the communication from source to destination. When it receives packets from source, it does not forward them to the intended destination; instead, it drops them, thus disrupting the network operation.

In this paper, the anomaly and specification based techniques are combined, and tested on networks that operate on AODV to detect black hole attacking nodes, by making use of the advantages of the two techniques to improve the detection rate.

4 SOFT COMPUTING BASED IDS

Soft computing is a multidisciplinary system defined as the fusion of fields of Fuzzy logic, neuro computing, genetic computing and probabilistic computing. Soft computing is designed to model and enable solutions for real world problems, which are difficult to do using mathematical techniques.

Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned. Fuzzy inference systems have been successfully applied in data classification, decision analysis, and expert systems. There are two types of Fuzzy Inference Systems (FIS):

1. Mamdani Type, and
2. Sugeno Type.

This paper makes use of Sugeno type FIS because it works well with optimization and adaptive techniques.

GA is a family of computational models, which uses the principles of evolution and natural selection. These algorithms convert the problem in a specific domain into a model by using a chromosome like data structure and evolve the chromosomes using selection, recombination and mutation operations. In computer security, it is mainly used to find an optimal solution to a specific problem.

Many intrusion detection systems using either signature or anomaly techniques are discussed in literature. A Specification based system is also designed for AODV using finite state machine (Seng, 2003). Some hybrid approaches combining misuse and anomaly are found in the literature. Here, it is proposed to combine the anomaly and specification based methods to improve the detection rate and to reduce the false positives. The anomaly technique detects the malicious nodes by describing the normal behaviour. Specification based mechanism crafts the rules that are related to the protocol. These two features are combined in the proposed design as follows. The simulated networks consisting of normal nodes and black hole nodes are run with various traffic conditions. The volume of traffic varies broadly to provide data in the low, medium and heavy traffic categories. This is the training mechanism of anomaly system. From the available data set, the features related to AODV protocol are extracted, which is a typical specification based approach. In addition to combining these two methods, the system is tested the using the two basic computational intelligence methods. The results for the two systems using FIS and GA are compared.

The algorithm of the proposed system is outlined as follows:

1. From the training data set, the features that are relevant to AODV are extracted. They are:

   - Number of packets dropped (PD) which dictates the behaviour of the black hole. Simply because a node drops more number of packets compared to any other node in the network, it does not lead to the decision of black hole node. This may be due to any of the reasons mentioned in assumption numbered 2.
   - Number of RREPs sent by any node (SREP) is greater for a black hole node for the simple reason that the adversary node tries to make itself available in the communication path.
   - Number of RREQs (RFR) forwarded will indicate the malicious behaviour of the black hole node as it does not forward any RREQ to the neighbouring node to avoid other nodes participating in the RDP.

2. As dictated by the specification-based approach, the threshold value for the selected parameters are defined, but using the anomaly technique, as the mean of each feature.

3. To enhance the detection process a credit allocation scheme is introduced in the system. Each node is allotted an initial credit, which is determined by the RFR. The credit is incremented or made to zero as per the following rule:

   - If the value of each parameter at each node is lesser than the respective threshold value, the credit is incremented by one.
   - Else the credit is assigned zero.

4. The total credit (TC) of each node is the sum of individual credits.

5. The parameter values of all nodes, the respective credit values are given as inputs to the FIS and GA systems.

6. The detailed algorithm used in each of the system is described in the later sections.

A novel combinatorial design approach of IDS that uses the anomaly and specification based technique is proposed. Also the design is tested using the two basic soft computing mechanisms. Generally, an IDS consists of the following fundamental components: Data collection components, Data-analysis components.

In this system, data collection is carried out using ns2, which is a network simulation tool. MANETs with 15, 25 and 50 nodes are created with AODV as routing protocol. Special data collecting (DC) nodes are placed in a strategic position so as to do the data collection process. These nodes move throughout the network and work in promiscuous mode. Hence they are capable of monitoring the behaviour of each wireless node in its vicinity. To simulate black hole attack, the hop count and the sequence number in the RREP is appropriately modified in the AODV, the changed protocol is named as ‘Black hole AODV’ (BAODV). The nodes that use this BAODV are treated as black hole nodes. Many networks with a mixture of genuine and black hole nodes are simulated. The number of black hole nodes varies from 1 to 5% of the total nodes in the network. One-to-one communications between a variable numbers

of pair of nodes are simulated with random walk mobility model. Following are the assumptions made:

1. All the nodes in the network are genuine in forwarding the RREQ packets to the neighbouring nodes in the coverage region, when they don't have a route to the destination defined in the request packet.

2. Nodes may drop the received data packet only under the following two conditions: a). The network is severely congested due to heavy traffic, or, b). The source/destination node or the intermediate node on the identified route had moved out of range of the forwarding node.

3. A genuine node may even try to establish communication with the black hole node.

4.1 FIS based IDS

The IDS that uses FIS works with Sugeno type as it does not require defuzzification process. Figure 2 shows the basic operations of the FIS. The input and credit values from the database are grouped into various clusters. Either of the following two methods is used for the purpose:

1. Subtractive Clustering
2. Fuzzy C-means Clustering (FCM)

Subtractive clustering is a fast one pass algorithm for estimating the number of clusters and the cluster centers in a set of data. In FCM, each data point belongs to a cluster to some degree that is specified by a membership grade. Both clustering mechanisms were incorporated in the proposed system and the subtractive clustering mechanism is found to be more appropriate for the given application.

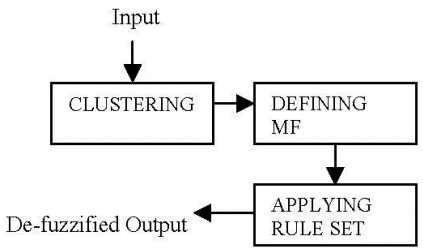

Figure 1: Sugeno FIS.

The input and output clusters are defined interms of membership functions (MFs). A MF is a means through which each point in input value is mapped to a membership value between 0 and 1. Gaussian distribution has been selected as the MFs for this system. This process is depicted as 'defining MF' in figure 1. After clustering, and mapping the input and output values to the appropriate MFs, various rule sets are applied on them. Sugeno type FIS defines its own set of rules based on the clustering. Finally defuzzified output produced by the FIS consists of values that give an indication to the performance of each node. The entire process is shown in figure 1.

The algorithm for the FIS based IDS is as given below:

- Get input: PD, SREP, RFR, TC
- Perform the clustering operation using subtractive and FCM methods.
- Define the Gaussian MF for each cluster of inputs.
- Generate a Sugeno-type 2 FIS to apply the rule set on the inputs.
- From the output of FIS, compute the standard deviation to facilitate the detection process.
- Compare the performance of each node with respect to the calculated standard deviation as described below:
  - If the FIS output of the node is lesser than the standard deviation, it is considered to be a normal node.
  - If the FIS output of a node is greater than the standard deviation, it is detected as black hole node.
- The nodes that are identified as black hole nodes are displayed with their respective node ids.

4.2 GA based IDS

The algorithm for GA based IDS is as follows:

- Assign the initial population as the network parameters derived from the data set: PD, SREP, RFR
- Calculate the initial threshold based on the network parameters as follows:

T1 = Average (NPi),

Where, T1 = First Threshold

- NPi = Network Parameter , i = 1,2,3….N.

N= Total no of nodes in the network

- Encode the chromosome based on the threshold criterion and determine the chromosomes that are above the threshold.
- Shortlist the chromosomes based on their fitness. Calculate the optimum parameter as per the following condition:

If(NPi<=T1) then { NPi-op =1}

else {NPi-op =0}, Where NPi-op is the optimal network parameter. This fitness evaluation process

results in making the respective network parameter value as '0' for fit nodes (i.e., the black hole node) and '1' for unfit nodes (i.e., the normal nodes).

- Select the survivor based on selection and recombination criteria. If all the network parameter values for any node is zero, it is selected as the survivor black hole node.
- Display the survivor node list along with its respective node id.

4.3 Simulation Results & Performance Analysis

This paper provides a combinatorial approach for soft computing based intrusion detection. The design was explained elaborately in the previous sections. This section covers the simulation of the various scenarios and the performance of each system. As mentioned earlier, ns-2 was used to simulate various MANETs that consist of different nodes. The black hole nodes and DC nodes were simulated appropriately. After the network had run with various traffic and communication scenarios, the parameters necessary for the proposed system were derived from the data log of DC nodes. The various details of the simulation are given in table 1.

Table 1: Simulation parameters.

Description	Value
Network Size	15, 25, 50
Max. Speed & Pause Time	200m/s & 2s
No. of Black hole Nodes	1,2,3 (For 15 node N/w) 1,2,5 (For 25 node N/) 3,6,10 (For 50 node N/w)
Simulation Duration	500 s
Simulation Area	1000 × 1000 sq.m
Traffic Density	Low, medium, High
Type of Communication & Application	One-to-One CBR

The various parameters such as PD, SREP, RFR and TC are given as inputs to the two systems. The performance of each system is discussed later in this section. The performance measure of any IDS can be defined in terms of the following:

- True Positive Rate (TPR)
- False Positive Rate (FPR) and
- True Negative Rate (TNR)

TPR is the measure of number of black hole nodes correctly identified. FPR is the measure of number of normal nodes identified as adversary nodes. TNR indicates the number of black hole nodes being detected as normal nodes. Our system based on the three computational intelligence techniques are analysed using the above three measures. The performance of a FIS based IDS is given in figure 2.

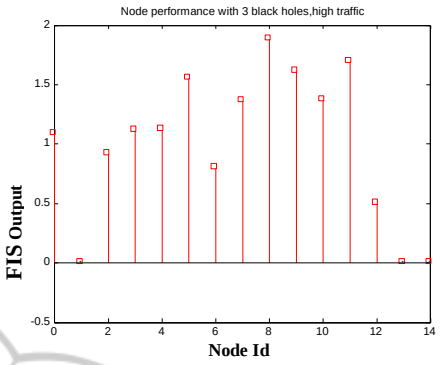

Figure 2: Detection using FIS based IDS in a MANET of 15 nodes with 3 black holes and high traffic condition.

From figure 2, it is inferred that the nodes with the least value of performance are identified as black hole nodes. Hence, the nodes having id 1, 13 and 14 are detected as black hole nodes by the FIS based IDS. As per the simulation, the nodes with id 12, 13 and 14 are simulated as black holes. While our system detects 13 and 14 correctly as black hole nodes, it has not detected node 12, which is also a black hole. It is also understood that node 1, which is a normal node is being detected as an adversary node. From the above result, it is clear that the FIS based IDS produces true positives, true negatives and also false positives. The results of detection process of various MANETs using FIS are tabulated in table 2.

From table 2, it is understood that in most of the cases, the fuzzy based system produces 100% true positive rate, which indicates its efficient detection. Exactly the efficiency is 81.8%. On the flip side, it results in false positives and false negatives. The performance of our fuzzy system for 50 nodes is plotted in figure 3.

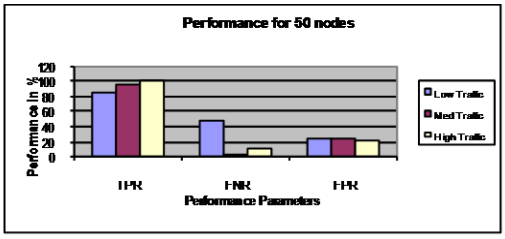

Figure 3: Analysis for 50 nodes network.

The detection of malicious nodes using GA is shown in figure 4. It can be easily deduced from figure 4 that the survivor nodes are the black hole nodes and they are displayed with their node ids. It

is understood that the nodes with id 24 and 25 were detected as adversary nodes. Earlier in the simulation for this particular case consisted of nodes 24 and 25 as black hole nodes. Moreover, GA based system does not produce any false positives or false negatives.

Table 2: Analysis of FIS based IDS.

No.of Nodes	No.of Black Hole Nodes	Traffic	TPR	FNR	FPR
15	1	Low	100	0	11.3
		Med	100	0	7
		High	100	0	7
15	2	Low	100	0	11
		Med	100	0	13.3
		High	100	0	7
15	3	Low	100	0	11
		Med	89	11	11.3
		High	89	11	7
25	1	Low	100	0	20
		Med	100	0	13
		High	100	0	13
25	2	Low	100	0	11
		Med	100	0	13
		High	100	0	13
25	5	Low	100	0	13
		Med	100	0	13
		High	100	0	13
50	3	Low	55.3	44.7	24.6
		Med	89	11	15.3
		High	100	0	22.3
50	6	Low	100	0	24.6
		Med	100	0	27
		High	100	0	27
50	10	Low	100	0	24.6
		Med	100	0	27
		High	66.7	33.3	18

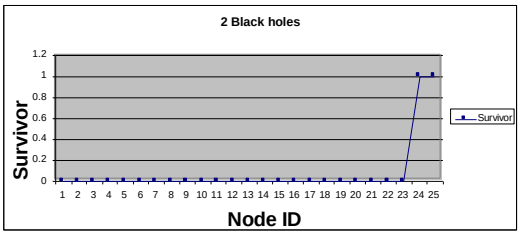

Figure 4: Detection of GA based IDS for a network of 25 nodes and 2 black holes.

5 CONCLUSION & FUTURE WORK

A soft computing based IDS that combine the specification and anomaly approaches was presented. The algorithms using computational intelligence techniques were developed to detect black hole attacks in MANETs The two fundamental soft computing techniques such as fuzzy and genetic algorithm were used. Simulation study of these systems was also exhibited. Among the techniques GA based technique has 100% detection rate compared to the 81.8% detection rate in FIS based system. In future, publishing the identity of the detected node throughout the network using the DC nodes will be carried out. Our future work includes improving the detection rate of fuzzy based system using cross layer approach. Also the possibilities of this system against other type of attacks will be explored.

REFERENCES

Charles E. Perkins, 2001, "*Ad hoc Networking*", Addison-Wesley.

Farooq Anjum, 2007, "*Security for Wireless Ad hoc Networks*", John Wiley & Sons.

Ming-Yang Su, 2011, "Prevention of Selective black hole attacks on mobile attack networks through IDS", *Computer Communications*.

Anup Goyal and Chetan Kumar, 2010, " GA-NIDS: A Genetic Algorithm based Intrusion Detection System".

C.Sivaram Murthy, B. S. Manoj, 2007, "*Ad hoc Wireless Networks, Protocols and Architecture*", Pearson Education.

Chin-Yang Tseng, 2003, Poornima Balasubramanyam, Calvin Ko, "A Specification-based Intrusion Detection System for AODV", *Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks*.