CHAPTER 6

DETECTION OF BLACKHOLE ATTACK USING GENETIC APPROACH

6.1 INTRODUCTION

This chapter presents the final contribution of this research work, the scheme to detect blackhole attack using genetic approach. Genetic algorithms (GA) are adaptive heuristic search algorithms based on the evolutionary ideas of natural selection and genetics. GAs are a part of evolutionary computing, inspired by the theory on evolution – “survival of the fittest”. They represent an intelligent exploitation of a random search used to solve optimisation problems. These algorithms are implemented by converting a problem in a particular field into a model of chromosome like structure. In computer network security, it is mainly used to find an optimal solution to a problem (Wei Li 2004). In detecting intrusive activities in networks, such as blackholes, genetic algorithms can be used to evolve simple rules to distinguish the normal nodes from anomalous ones. A simple genetic approach is used in this dissertation to detect the blackhole attacks.

6.2 GENETIC ALGORITHM BASED DETECTION

The genetic algorithm starts by identifying a data set called population. These are individually encoded using bits, characters or integers and they form a chromosome. An ‘Evaluation Function’ is applied on these chromosomes to determine the genuine ones. The selection of the chromosome is biased towards the fittest of the species. The fit chromosome
is selected once the optimization criterion is met. An optimal solution is not met if the finishing is slow or the convergence is premature. This is determined by means of proper selection of the fit condition. The fit chromosomes undergo mutation and cross over operations to form the new population. Another fitness function selects the fittest surviving individuals. The basic steps of GA mentioned above are applied in the scheme to detect blackhole attack as follows.

The parameters specific to the route discovery process of the routing protocol and the behaviour of blackhole attack are derived from the simulated network data: Number of packets dropped (PD), number of RREPs transmitted to source node (SREP) and the number of RREQs forwarded to the neighbouring nodes (RFR). These parameters form the initial population required for applying genetic algorithm to the scheme.

The values of each network parameter are integers and hence these values of initial population are termed as the chromosomes. Each chromosome is then evaluated for an objective function using the first threshold value of the parameters such as PD, RFR and SREP. The first threshold is determined by calculating the average of the individual network parameters. Then the fitness criterion for every network parameter is determined based on Tournament Selection wherein the fit blackhole nodes are assigned ‘0’ and the unfit nodes are assigned a ‘1’. The selection of fit nodes is carried out by considering the fitness criterion for all the network parameters. The nodes having lesser degree of genuineness (DG) are selected to form the new population. A second threshold is determined as the weighted average of the network parameters. The fitness function based on the second threshold is applied on the new population. The nodes that have all the weighted average parameter values as zero are selected as the surviving black
hole nodes. The detection scheme using GA is shown as flow chart in Figure 6.1.

![Flow diagram of genetic detection scheme](image)

**Figure 6.1 Flow diagram of genetic detection scheme**
The process flow of the detection scheme using GAIs given below:

1. Get the network parameters PD, SREP, and RFR of the nodes in the network. (NPi)

2. Calculate the first threshold based on network parameters as

   \[ T_1(i) = \text{Average (NPi)}, \text{ where, } T_1(i) \text{ is first threshold}; \text{NPi is individual network parameter}; i = 1,2,3. \]

3. Encode the chromosomes based on threshold criterion.

4. Shortlist the chromosomes based on their fitness by calculating the optimum parameters value as:

   If (NPi<=T1(i))

   \{NPi-op =1\}

   else

   \{NPi-op =0\}

5. Determine the second threshold, T2(i) by computing the weighted average of all the network parameters of the fit chromosomes.

6. Select the survivor based on selection and recombination criteria: i.e., the nodes having value lesser than the weighted average or zero value for all the optimal parameters, become the survivor blackhole node.
6.3 RESULTS OF THE GA BASED DETECTION SCHEME

The experimental results of the detection scheme using GA to detect blackhole attack in MANETs are displayed in the following plots. The X-axis of these plots gives the node id and the total number of nodes in the simulated MANET. The Y-axis shows the output of the proposed scheme in terms of the degree of genuineness (GD), i.e., the degree to which a node is genuine. This value is computed using the various network parameters and the threshold values for each parameter. Low value (GD=0) indicates that it is a blackhole node.

![Figure 6.2 Selection of new population for 25-node MANET with 2 blackhole nodes](image)

Figure 6.2 illustrates the process of selection of new population from the chromosomes using the first threshold. The plot shows the result in a 25-node MANET with 2 blackholes and low traffic-1 scenario. It can be seen
from this figure that the nodes with id 1, 2, 23 and 24 are selected as the members for new population, as they have the least genuineness values. The surviving blackhole nodes are selected by applying another fitness function, second threshold. The result of the final selection process is given in Figure 6.3. It is understood from this figure that the nodes with id 23 and 24 have the least GD value and hence declared as blackhole nodes for this network. Two other nodes 1 and 2, which were members of the new population, did not survive after the second fitness function had operated on. This leads to the final result that the nodes 23 and 24 are blackhole nodes, which is verified with the simulation scenario, and produces accurate detection.

![Graph showing degree of genuineness vs node id](image)

**Figure 6.3** Result of the genetic detection scheme for 25-node MANET with 2 blackhole nodes
Similar inference can be made from the figures 6.4 and 6.5, which depict the selection of new population and the final result, respectively, for an ad hoc network with 50 nodes and 10 blackhole nodes in a high traffic scenario. As seen from Figure 6.4, the nodes 1,2,3,4,40,41,42,43,44,45,46, 47,48 and 49 are selected using the first threshold, to form the new population. A second fitness function selects the fittest blackhole nodes from this population, as given in Figure 6.5. It is clear from this figure that, the nodes 1,2,3 and 4 are eliminated in the final selection process to detect the blackhole nodes. This results in identifying the nodes 40 to 49 as malicious nodes, as they have the least genuininess value. This detection is found to be correct, when verified with the MANET simulations.

![Graph showing selection process](image)

**Figure 6.4** Selection of new population in 50-node MANET with 10 blackhole nodes
Figure 6.5 Result of the genetic detection scheme for 50-node MANET

6.4 PERFORMANCE ANALYSIS

The blackhole detection scheme using genetic approach has shown better performance in all aspects. Table 6.1 presents the various performance measures of this scheme.

Table 6.1 Performance analysis of GA based detection scheme

<table>
<thead>
<tr>
<th>Network size</th>
<th>Traffic intensity</th>
<th>TPR (%)</th>
<th>FPR (%)</th>
<th>Detection time (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Low</td>
<td>95.06</td>
<td>0.741</td>
<td>0.016654</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>96.3</td>
<td>0</td>
<td>0.017283</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>96.3</td>
<td>0</td>
<td>0.018301</td>
</tr>
<tr>
<td>25</td>
<td>Low</td>
<td>94.44</td>
<td>0</td>
<td>0.022421</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>100</td>
<td>0</td>
<td>0.029222</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>100</td>
<td>0</td>
<td>0.027579</td>
</tr>
<tr>
<td>50</td>
<td>Low</td>
<td>100</td>
<td>0</td>
<td>0.028132</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>100</td>
<td>0</td>
<td>0.028614</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>100</td>
<td>0</td>
<td>0.030291</td>
</tr>
</tbody>
</table>
The TPR gives the detection accuracy of the scheme and it can be inferred from the table that, the accuracy of detection increases with the size of the network. When the number of nodes in the network is less, i.e., 15, the true positive rate is comparatively lesser. The GA based scheme has produced some false alarms, especially when the network is small and has low load, due to the less information available from the network. This inference is supported by the result for 25-node network with low traffic scenario. When the information from the network is large, which is possible in large networks or network where the traffic is more, the genetic approach tends to detect the malicious activity in accurate manner. Considering the speed of detection, the GA based scheme is better and comparable with that of the FIS based scheme with subtractive clustering. The detection time increases with increase in the number of nodes and the intensity of the traffic, due to the various processes inherent to GA.

6.5 CONCLUSION

The scheme to detect blackhole attack using genetic algorithm is presented in this chapter. The various processes of genetic algorithm applied to the detection scheme are discussed and the results are presented. Results show that this scheme produces 97.74% detection accuracy with 0.023% false alarms.