Securing Vehicular Ad Hoc Networks from Sybil Attack

Priya Agrawal¹, Dushyant Singh²

M. Tech Scholar, Chandravati Group of Institutions, Bharatpur, Rajasthan, India¹

Assistant Professor, Department of CSE, Chandravati Group of Institutions, Bharatpur, Rajasthan, India²

ABSTRACT: There are various attacks that can hamper the efficiency of VANET major of them is Sybil attack. In Sybil attack the attacking node creates multiple forge identities in order to gain a generously large influence. In this paper Extended Genetic Algorithm is applied to remove Sybil attacks in VANET. In the proposed work Extended Genetic Algorithm will be applied with fitness function optimization. In this paper Genetic Algorithm is applied to mitigate Sybil attacks in VANET. In starting, the K-mean for clustering the VANET nodes have been utilized according to their distances among them. Next, the intelligent CH election utilizing soft computing methods that involves sequential hybrid Fuzzy-Genetic controller for this decision making have also been used. Proposed approach is able to locate Sybil nodes quickly without the requirement of secret information exchange and special hardware support. We evaluate our proposed approach on the realistic traffic scenario. Experiment results demonstrate that detection rate increases when optimal numbers of Sybil nodes are forged by the attacker. In this, we also compare our results with previous existing technique to show the enhancement we have done through our paper. The results are being evaluated on the basis of throughput; the whole stimulation model takes place in OPNET environment.

KEYWORDS: VANET, Security, Genetic Optimization Algorithm, Fuzzy-Genetic Controller, AODV, Clustered-AODV.

I. INTRODUCTION

Vehicular Ad-hoc Network (VANET) is a part of an ad hoc network. Vehicle in VANET interact with close vehicle or road side unit(RSU) that are fitted on vehicle or at the intersection and parking lots.[1] In VANET, two kinds of communication is done: vehicle to vehicle (V2V) and Vehicle to infrastructure(V2I). In V2V communication close vehicles transfer or obtain data utilizing short range communication facilities i.e. Wi-Fi and WAVE. Vehicle has a particular electronic gadget that permits them to obtain or relay messages. In V2I, vehicle interact with nearby rode infrastructure unit (RSU) by utilizing Wi-Fi hotspots or long range wireless communication techniques to transfer the data.[2][3]In VANET, Some applications need group interaction Services. Thus Multicast routing is the most effective technique, overcoming the broadcast and unicast routing. The quick change in the network configuration, the nodes high speed and other VANET characteristics, build the multicast routing a major challenge in vehicular scenarios. Thus, a Proficient protocol to manage a good performance in the transmission/reception of multicast packets is needed. In VANET network for transmitting data whole mobile nodes acts as router involving source and destination nodes transferring data the whole mobile nodes act as a router as well as source and destination node.

A successful technique for dealing with and maintain mobile ad hoc networks is by dividing the network into clusters. In this manner it permits better management of mobile ad hoc networks. Clustering is a technique which partitions the network into groups [6, 7]. Clustering in computer network is a division of the network into several virtual groups, depending on rules for discriminating the nodes assigned to various sub-networks. Primary objective is to obtain scalability in existence of huge networks and high mobility. Cluster-based routing is a solution to approach nodes heterogeneity, and to restrict the amount of routing information that propagates within the network. It increases the lifetime of routes, hence reducing the amount of control overhead of routing [7]. Although the clustering methods solve several issue in VANET but it still be complicating to implement and choose the optimal CH in every cluster. There are...
many routing protocols and cluster based protocols have been introduced for VANET and famous among them are AODV and DSR [8].

II. INTELLIGENT CLUSTER MANAGEMENT

The clustering methods present network scalability and to yield more effective usage of network equipments. It can be utilized for routing, resource management and location management for decreasing communications, unessential flooding RREQ packets and calculation overhead. In this section, intelligent cluster head election, cluster structure and communication techniques have been shown.

At the starting of the implementation and before the nodes movement (when time equal to zero), the setup stage requires to be started. In this state, every network node is allocated with particular information as follows:

- The current geographical position within the particular region.
- The transmission speed node (mobility).
- The battery power (0-1.4)
- The speed limit 0-10 m/sec with pause time 20 sec.
- The next geographical location after movement.
- Random way point mobility mode has been utilized by defining the current and next position arbitrarily. With presentation that every node move from its current position in straight direction to the next position at the particular speed.

After allocating the properties of the network and the nodes in setup stage, the distance matrix is computed to describe the nodes connectivity and the distances among them. This process needs the usage of the Euclidean equation depending on the position coordinate of every node with its neighbouring nodes.

To implement the routing in clustered AODV network some changes in the classical AODV protocol should be built. This is because it is based on the node’s ID no. to find the route by broadcast RREQ packets throughout the network.
Intelligent Cluster Head Election: The challenge task in the design is the intelligent head nodes election. This phase utilizes soft computing methods (Fuzzy system and Genetic algorithm) to choose the optimal head nodes that are capable to send the data through network in flexible way. There are several algorithms shown for cluster head election in network community. The introduced intelligent cluster head election design should pass via two phases: the genetic algorithm, and the fuzzy system. This consideration is made to assure that the combination of these two tools will support to obtain the advantages of both methods.

Clustered-AODV-based Routing Stage: In this phase changes in AODV protocol has been performed to be in line with intelligent clustering method in prior phase. After clustering and elects CHs in intelligent manner, two kinds of routing in the cluster-AODV network have been mentioned.

Intra-Cluster Routing phase: However AODV is a reactive routing protocol, the source and target nodes are mentioned first then the routing process will start. If the source and destination positioned in one cluster only, the AODV routing algorithm will perform inside this cluster without broadcasting the RREQ packets outside for searching the route. This is performed by selecting the shortest path between the source and target node in normal way.

Inter-Cluster Routing phase (the intelligent phase): The best route among the CHs in the network is determined which shows the optimal distance (minimum distance). This route will be taken as the backbone or the highway among the clusters to interact with each other directly through it. There are two conditions in inter-cluster routing:

The first situation when the source and destination positioned in various clusters which are at distance to each other (out of transmission range of signal). The CHs of them cannot link directly due to restricted transmission range and the packet cannot be arrived. In this case the source node utilizes classical AODV routing protocol within its cluster to arrive the CH by taking it as a local destination node. And then link to the CH of the destination via the backbone route. After arriving the destination CH return to utilize classical AODV routing protocol to arrive the final destination node in the cluster. This means that the route outside the cluster behaved proactively not reactively until arrive the CH of the target node, then the packet send to the destination node by utilizing the classical AODV protocol.

The second situation when the source and destination positioned also in different clusters but in the same signal transmission range. In this case the CHs of them will be linked directly without requirement of use the backbone path of CHs. Because the CH of destination node is reachable transmission range. This yields to decrease the resource exhaustion particularly the CHs of other clusters that not participating in the interaction.

Genetic Algorithm Stage

In this phase, the introduced method of GA when used to optimize the CH election has been explained. Table (1) represents the control parameters of GA managed in the simulation.

The GA phase includes the following steps:

1. Chromosome Encoding step: The first step in GA design is encoding the chromosomes. In the followed design, this will be the network CHs. Thus the length of chromosome is variable as the size of cluster. For example when the size of cluster is five the no. of genes in one chromosome is five and every gene show the ID of every CH in the network as illustrated in Fig (3). For the network size 100 node, the ID should not more than 100. In the simulation, various cluster size requires to be inquired from two clusters to twelve to view the effect of changing cluster size to the performance criteria. This means that the length of chromosome should change from two to twelve.

Fig 3: Chromosome representation when cluster size is five and network size is 100 nodes.

Initial Population Generation step: The initial population generation is chosen arbitrarily from the group of chromosomes on the condition of choose one node from every cluster as a candidate CH in that cluster. When the initial population generation selection is close the optimal the process will be faster to arrive optimization in short time.
Fitness Evaluation step: The fitness evaluation is the significant step in the algorithm because it decides the optimization algorithm accuracy. It is detected which one of the population is the best to support to the next generation and become strong candidate to win the optimization. The introduced algorithm utilizes an integration influencing factors that show weight metric to search dominant set. The total weight is composed by three influencing factors. These are:
i- Battery power of CHs (E)
ii- Distances between CHs (D)
iii- Mobility of CHs (M)
All these factors can be integrated together in one equation that shows the fitness equation.

\[ W_{out} = \frac{w_1 \times E}{w_2 \times D + w_3 \times M} \] (1)

Where \( w_1, w_2, w_3 \) are the weighting factors for the corresponding system control parameters and \( W_{out} \) is the fitness output. Every influencing factor can be achieved from the node information and nature. The weighting factors will be treated in FIS phase.

Crossover and Mutation step: In the first operator, one point crossover has been utilized to create new chromosomes from old ones. Next, the population generation sorted depending on the fitness evaluation from the best to the worst candidate’s chromosomes and crossover between each pairs as illustrated in Fig (4). The mutation operation is very simple but significant to prevent the locally optimum result.

![Fig 4: The one point crossover in five cluster size and 100 network size.](image)

Selection step: This step selects which chromosome will be hold in the next generation and which one will be dropped. This procedure achieved by determining the best chromosomes in the population utilizing the following equation:

\[ \text{Fitness} = \max (W_{out}) \] (2)

Where Fitness shows the vector of descending order from best to worse chromosomes after GA operation and pop shows the size of population of GA. The best fitness of the population shows the best CHs selected according to all the three parameters explained before. At last, the algorithm can be iterated many times to obtain the best results. This is due to the first attempt is not essentially provides optimal chromosomes. This iteration in the algorithm known as the iteration no. which requires a stop condition to be deployed in the introduced system.

Fuzzy Logic Controller Stage
The primary challenge in the optimization of CH election for VANET is to introduce a flexible algorithm that can be prepared to the broad variety of situations in any VANET atmosphere. The algorithm should behave professionally in atmosphere experience i.e. limited battery power, high nodes mobility and some wireless transmission restraints. Because it is usually hard to know which node will endure from its atmosphere and because the atmosphere can change unpredictably, the algorithm must be capable to adapt automatically. The Fuzzy Logic Controller (FLC) is a highly adaptive algorithm that can be utilized to obtain these objectives.

The significant issues in the fuzzy rules are to hold the output weighting factor as maximum as possible in the following cases:
1. Minimum weighting factor of average distance between elected CHs.
2. Maximum weighting factor of average battery power (energy) of elected CHs.
3. Minimum weighting factor of average mobility of elected CHs.

Fig 5: The proposed Algorithm flowchart.
III. PROPOSED SYSTEM SIMULATION AND EVALUATION

The nodes are positioned arbitrarily and move randomly in Random Way Point model. The simulation has been done by employing OPNET, a high performance language for technical computing. OPNET characterizes a family of add-on application particular solutions known as toolboxes. Very significant to most subscribers of OPNET, toolboxes permit learning and using specialized technology [11]. Table (3) represents the simulation atmosphere parameters utilized in the measurement. We take a network of nodes positioned arbitrarily in different arrangements (one source and one destination) within a 7000m X 7000 m area. The performance of intelligent CH election is measured by changing the network size (no. of nodes) ranges 20 to 100 and cluster size partitions into two groups: low clustered network (Group 1) 2 to 6 cluster and high clustered network (Group 2) 7 to 12 cluster. The introduced intelligent CH election system is checked against three parameters i.e. mobility degree, Average energy and average distance between CHs to compute the optimality and reliability of the control system.

Table 3. THE SIMULATION PARAMETERS

<table>
<thead>
<tr>
<th>Examined Protocols Cases</th>
<th>AODV without Sybil Attack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Nodes</td>
<td>150</td>
</tr>
<tr>
<td>Types of Nodes</td>
<td>Mobile</td>
</tr>
<tr>
<td>Simulation Area</td>
<td>7000*7000m</td>
</tr>
<tr>
<td>Simulation Time</td>
<td>1800 seconds</td>
</tr>
<tr>
<td>Mobility</td>
<td>Constant 10 m/s</td>
</tr>
<tr>
<td>Pause Time</td>
<td>200 seconds</td>
</tr>
<tr>
<td>Performance Parameters</td>
<td>Throughput</td>
</tr>
<tr>
<td>Mobility model used</td>
<td>Random waypoint</td>
</tr>
<tr>
<td>Data Type</td>
<td>Constant Bit Rate (CBR)</td>
</tr>
<tr>
<td>Packet Size</td>
<td>512 bytes</td>
</tr>
<tr>
<td>Traffic type</td>
<td>Http</td>
</tr>
<tr>
<td>Physical Characteristics</td>
<td>IEEE 802.11g (OFDM)</td>
</tr>
<tr>
<td>Data Rates(bps)</td>
<td>54 Mbps</td>
</tr>
<tr>
<td>Transmit Power</td>
<td>0.010</td>
</tr>
<tr>
<td>RTS Threshold</td>
<td>1024</td>
</tr>
<tr>
<td>Packet-Reception Threshold</td>
<td>-150</td>
</tr>
</tbody>
</table>

IV. RESULTS

After presenting the basic results of all simulations carried out in both scenarios, in this chapter, we analyse and discuss all these results. The performance metrics collected and presented in our results are either based on the object statistics or global statistics of the VANET model i.e. the entire network. In representing these data, we presented the average or time average values of the results in this report. We start our discussion and analysis with the two main scenarios in which the first scenario comprises of 150 mobile nodes and the latter holds 200 mobile nodes. In each scenario, we performed two simulations of a regular network operation in VANET and a VANET a Sybil attack to be precise. All simulations i.e. both scenarios were run for a time period of 30 minutes, which ranged from 0 to 1800 seconds as shown in the result graphs. After that, we analyse and compare within each scenario and also both scenarios based on their throughput and end-to-end delay. The overall simulation performance is presented in nutshell in the following table, which indicates that the elimination of Sybil attack scenario provides the better results and try to normalize the Sybil effected network to its normal state as close as possible.

Throughput:
Throughput can be defined as the ratio of the total amount of data reaches a destination from the source. The time it takes by the destination to receive the last message is called as throughput. It can express as bytes or bits per seconds (byte/sec or bit/sec). Scenario 1, represents the scenario with old algorithm normal network state, scenario 2 represents
the network that is under the Sybil attack and scenario 3 represents the mitigation of Sybil attack by using our proposed approach and implementation of the proposed method.

There are some factors that affect the throughput such as; changes in topology, availability of limited bandwidth, unreliable communication between nodes and limited energy. A high throughput is absolute choice in every network. In figure the graph represents the throughput in bits per seconds. The x-axis denotes the simulation time in minutes and the y-axis denotes throughput in bits per seconds. It can be clearly seen, that the Sybil attack decreases the overall network throughput in comparison to the normal network state. However, the entire network throughput is increased once the proposed unified mechanism is implemented. In addition to this, the state of the throughput has increased more than the no attack scenario after implementing the unified security mechanism.

**Figure: 6** Throughput of all three scenarios at 150 nodes

**Figure: 7** Throughput of all three scenarios at 200 nodes
In this research work, the network performance under Sybil attack is analysing and improving by applying integrated approach of genetic algorithm and K means clustering. This genetic based integrating approach is implemented on the selected nodes on the network and deployed in the specific area i.e 7000 x 7000 m. The findings of the research clearly states that, the implementation of such genetic based integrating approach have a significant impact on the overall network through positively. On the other hand, the implementation of such mechanisms does not only mitigate the Sybil attack effects, it also increases the overall performance above the normal state of the network. The proposed mechanism satisfactorily mitigated the effects of the Sybil attack on the network and increased the overall performance of the network. The data dropped rate decrease successfully. Our scheme is simple and efficient as compared to existing detection approaches because it does not require secret information exchange and special hardware support. As a part of future work, we would like to perform the experiments with varied transmission ranges of vehicles and with different number of Sybil identities used by attacker simultaneously. We also would like to investigate the processing time, storage capacity and communication overhead of proposed approach.

REFERENCES