Detection and Mitigation of Sybil Attack by implementing Extended Genetic Algorithm

Shilpa Goyal
Computer Science and Engineering Department
Shri Ram College of Engg & Mgmt
Palwal, India
Shilugoyal13@gmail.com

Ms. Nisha Pandey
Computer Science and Engineering Department
Shri Ram College of Engg & Mgmt
Palwal, India
n4nishalucky@gmail.com

Abstract— Today, there are several available technologies designed to build vehicular road travel easier, secure, and more enjoyable, utilizing proximity sensors, geographical positioning system, multimedia communication, etc. Although VANET is a popular application that has strengthens its roots since the last decade and made our lives much easier than ever before. But still there are various security issues in it that need to be considered. One of the major security issues relating VANET is the Sybil attack. The Sybil attack is a major threatening attack in which the attack creates several forge identities of itself in order to gain trust of the authenticated nodes to fulfill its malice presence. In this paper we will implement genetic algorithm for mitigating the Sybil attack. GA is a search technique that depends on the natural selection and genetics principles and which determines an optimal solution for even a hard issue.

Keywords- *VANET, MANET, Sybil Attack, Genetic Algorithm.*

*****

I. INTRODUCTION

Vehicular Ad Hoc Network (VANET) is a particular type of MANET in which entities that builds the network are vehicles and provides the ubiquitous computing concept for future. With VANET, vehicles can be turned into a network that will offer facilities similar like the ones utilized in homes and offices**.** Each envisioned application of VANET needs that the nodes continuously flood vital information i.e. location, speed which will increase the vehicles awareness about their whereabouts and warn drivers of unsuitable circumstances. VANET provides us with a number of applications which are categorized in three major groups namely comfort oriented applications, convenience oriented applications and safety oriented applications. The demand for communication among the vehicles has led to a number of applications which add to the comfort of the passengers. These applications have increased the road safety and comfort of the passengers.

II. GENETIC ALGORITHM

Genetic algorithms are normally a family of computational models which are motivated by the biological evolution. These algorithms encode a powerful solution to a particular problem on a simple chromosome i.e. data structure and use genetic operators to these structures so as to preserve severe information. GAs are more reliable as compared to other most search techniques because they need only information related to the quality of the solution created by every parameter set (objective function values) and not like other optimization techniques which needs derivative information, or worse yet, entire knowledge of the parameters and problem structure.

A. Genetic Algorithm has Following Features

- GA comes in the category of search algorithms that are based on computer and are random in nature. These algorithms are obtained from the natural theory of “survival of the fittest” being specified by Darwin.
- The mechanization of intelligent nature is a pre concern of this branch.
- GA is also appropriate for complicated problems.
- It generates the best of the best solutions.

B. Operators of Genetic Algorithm

GA begins with random creation of initial population and then the selection, crossover and mutation operations are conducted until best population is determined. Gas are practical and simple algorithm and easy to be implemented in power system.

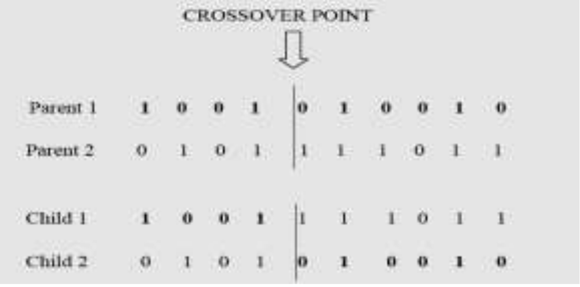
In other words, considering an initial random population generated and measured, genetic evolution happens by means of three basic genetic operators.

1) Parent selection.
2) Crossover.
3) Mutation.

The descriptions of these genetic operators are provided below:

1) Parent Selection/Selection Strategy: The selection of parents to generate successive generations plays a significant role in the GA. This permits the fitter individuals to be chosen more usually to reproduce. There is a no. of selection techniques introduced in the literature.
In this technique, n individuals are copied from the population randomly and the best of the n is introduced into population for further genetic processing. This process is repeated until the mating pool is filled.

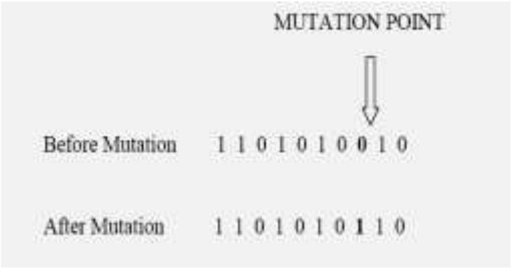
2) Crossover: Crossover is a significant operator of the GA. The primary aim of crossover is to reorganize the information of two different individuals and create a new one. It is a structured, yet randomized method of exchanging formation between strings. It encourages the exploration of new fields in search space. Cross swapping operator is used on the chosen individuals.
Here, two different cross sites of parent chromosomes are selected randomly. The cross over operation is finished by exchanging the middle substring between strings.



Parent 1	1	0	0	1	0	1	0	0	1	0
Parent 2	0	1	0	1	1	1	1	0	1	1
Child 1	1	0	0	1	1	1	1	0	1	1
Child 2	0	1	0	1	0	1	0	0	1	0

Figure 1: Crossover

3) Mutation: Mutation consists of securing the procedure of reproduction and crossover efficiently without much loss of the potentially helpful genetic material. Mutation is by itself a random walk through the string space and offers for occasional interference in the crossover operation by introducing one or more genetic elements during reproduction. This operation assures diversity in the genetic strings over large period of time and prevents stagnation in the emergence of optimal individuals. Bit wise mutation changes 1 to 0 and vice-versa. The above specified operations of selection, crossover and mutation are repeated until the best individual is detected.



Before Mutation 1 1 0 1 0 1 0 0 1 0

After Mutation 1 1 0 1 0 1 0 1 1 0

Figure 2: Mutation

III. PROPOSED ALGORITHM

In the proposed algorithm we have two phases the first one is route discovery algorithm and second one being the route reply algorithm. In the first phase i.e. route discovery the algorithm is applied when the source node does not know about the destination node then we apply genetic algorithm to find the best possible solution. We start the algorithm by generating a random population by using Dijkstra Algorithm and evaluate the fitness function f(x) for Maintaining the Integrity of the Specifications each chromosome x in the population. Now we have got our required solution then we stop else we continue repeating the four processes namely selection, crossover, mutation and accepting. In the selection process we select two chromosomes having highest value of fitness function as parent chromosomes. The offspring are obtained by crossing over the parent chromosomes. Finally apply the mutation process. The final obtained offspring is placed in the new population. The steps are repeated till we get the optimum solution. In the second phase of our algorithm the route reply is initialized. When the shortest route from the source to the destination is reached the route is replied and saved in the cache memory until the end. When the end of RREQ TTL is reached the destination adds all the received RREQ in GA and finds the two best routes from source to destination. Finally, the destination node creates a Route Reply packet, adds these two routes in it and sends the RREP packet for the source node.

Now we will discuss the algorithm step by step and also in the form of flowchart.

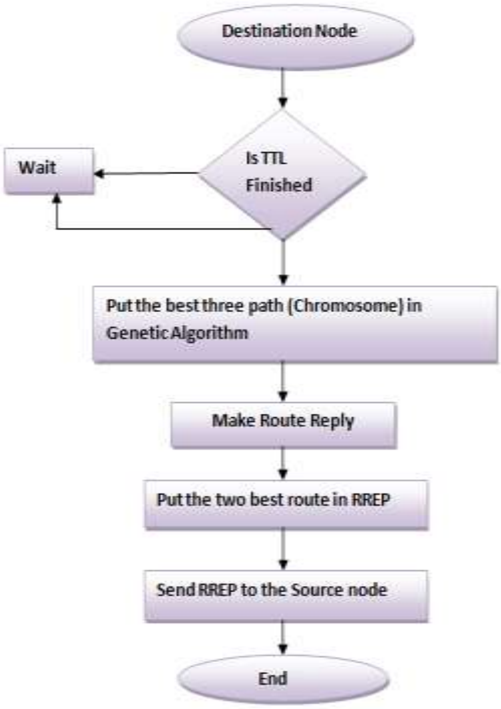

Destination Node
Is TTL Finished
Wait
Put the best three path (Chromosome) in Genetic Algorithm
Make Route Reply
Put the two best route in RREP
Send RREP to the Source node
End

Figure 3: Flowchart of Step 1 Of Genetic Algorithm

Step-1:- Initialization of Route Discovery
(i) If source node doesn't know about route to destination node
(ii) Then Apply Genetic Algorithm
1. *[Start]:* Generate random population of *n* chromosomes by using Dijkstra's Algorithm.
2. *[Fitness]:* Evaluate the fitness f(x) of each chromosome x in the population.
3. If we get our required solution then stop it, if not then repeat the following steps until the new population is complete
(i) Selection: Select two parent chromosomes from a population according to their fitness (the better fitness, the bigger chance to be selected)
(ii) Crossover: With a crossover probability cross over the parents to form a new offspring (children). If no crossover was performed, offspring is an exact copy of parents.
(iii)Mutation: With a mutation probability mutate new offspring at each locus (position in chromosome).
(iv) Accepting: Place new offspring in a new population
4. *[Replace]*: Use new generated population for a further run of algorithm
5.*[Test*]: If the end condition is satisfied, stop, and return the best solution in current population
6. *[Loop]*: Go to step **2**
(iii) end

Step-2:- Initialization of Route Reply
1. It saves received (RREQ) shortest path in its cache until the end of RREQ TTL (Time To Live).
2. When the RREQ TTL has been finished, the destination adds all received RREQ in genetic Algorithm and finds the two best routes from source to destination.
3. Finally, the destination node creates a Route Reply packet, add these two routes in it and sends the RREP packet for the source node

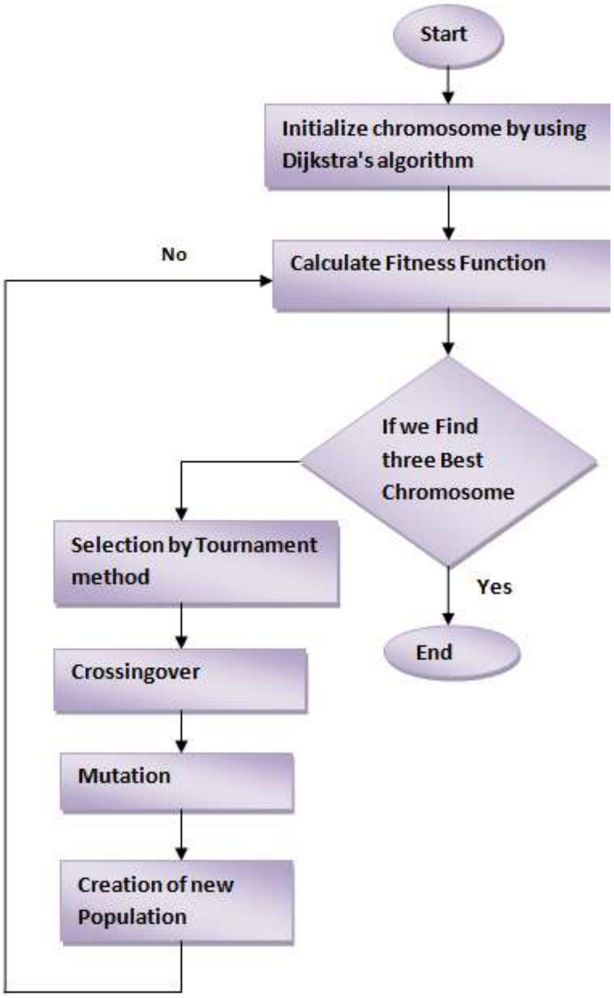

Start
Initialize chromosome by using Dijkstra's algorithm
Calculate Fitness Function
No
If we Find three Best Chromosome
Yes
End
Selection by Tournament method
Crossingover
Mutation
Creation of new Population

Figure 4: Flowchart of Step 2 of Genetic Algorithm

IV. SIMULATION ENVIRONMENT

A. *Simulation Tool Used*

The tool equipped for the simulation purpose is OPNET Modeler 14.5. OPNET is a network and application based tool used for simulation, network management and analysis. OPNET consists of various Units communication devices, almost all the protocols, architecture of various networks and technologies and thus provides simulation of their performances in the virtual environment with great ease. OPNET proves to be very helpful in various researches and it can also develop solutions which helps in the study and analysis of wireless technologies like analysis and designing of MANET protocols, WIMAX, UMTS, Wi-Fi, improving core network technology, providing power management solutions in wireless sensor networks. In our case we used OPNET for modeling of network nodes, selecting its statistics and then running its simulation to get the result for analysis.

B. *Simulation Setup*

The simulation work focuses on analyzing the performance of the network with and without the Sybil attack applied on the network. Therefore an integrated approach is used to analyze the network performance under the Sybil attack.

- Generating the network
- Adding the Sybil nodes on the network
- Applying the genetic algorithm
- Calculation fitness function
- Simulating and examining results

V. RESULTS AND ANALYSIS

There are three scenarios in our simulation work. The first scenario shows throughput without the attack. In the second scenario throughput parameter is considered to evaluate the performance of the network implementing the Sybil attack and in the last scenario the performance of the network is analyzed after implementing the genetic algorithm. Now we will present our results in three different graphs that are normal scenario(i.e. without attack), scenario with Sybil attack on it and scenario with proposed algorithm applied on it. At the last we will do comparison among the three by placing them on a single graph.

A. *Throughput Analysis for 150 nodes*

We will show our results first on 150nodes. We will present our results in three different graphs that are normal scenario (i.e. without attack), scenario with Sybil attack on it and scenario with proposed algorithm applied on it. At the last we will do comparison among the three by placing them on a single graph.

(i). The throughput evaluated below is for normal scenario for 150 nodes: **:** The first graph of normal scenario of 150 nodes of our experimentation, packets travels are shown as throughput with peak value of approx. 420551 bits per seconds and it is represented as bits per second.

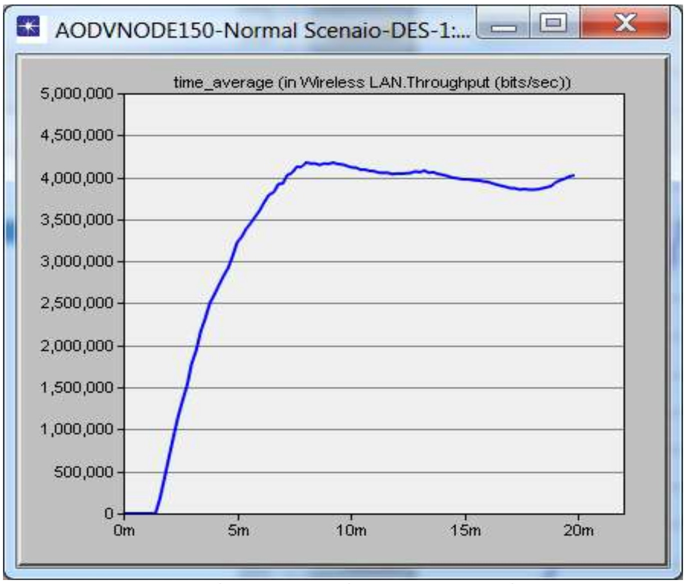


Figure 5: Graph of Throughput of Normal Scenario for 150nodes

(ii). The throughput evaluated below is for Sybil nodes for 150 nodes: the graph showing the throughput of Sybil attack for 150 nodes of our experimentation, packets travels are shown as throughput with peak value of approx. 415320 bits per seconds and it is represented as bits per second.

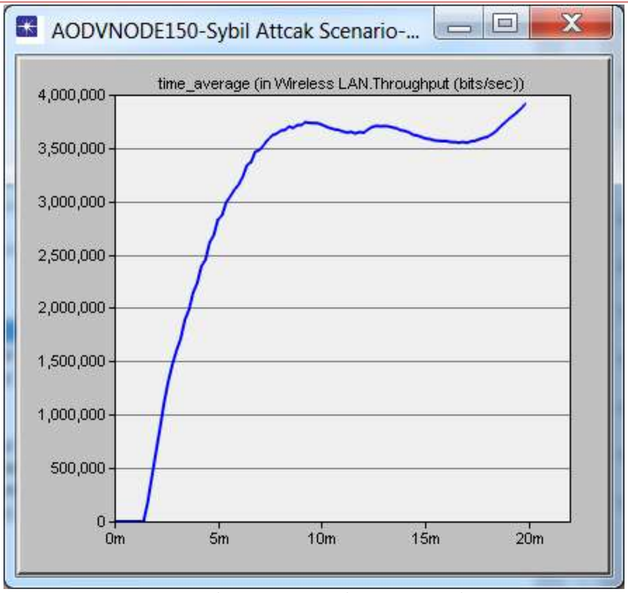


Figure 6: Graph of Throughput of Sybil Node for 150nodes

(iii). The throughput evaluated below is for mitigating the attack for 150 nodes: The graph i.e. fig 7 shows the throughput of mitigation Sybil attack for 150 nodes of our experimentation, packets travels are shown as throughput with peak value of approx. 419107 bits per seconds and it is represented as bits per second.

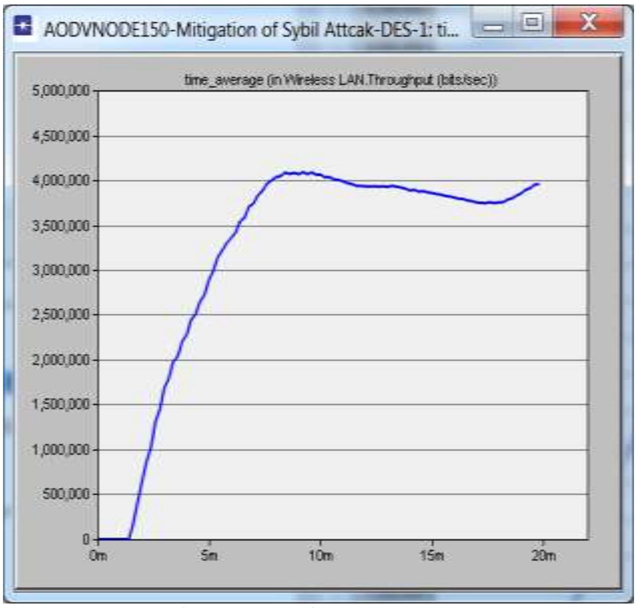


Figure 7: Graph of Throughput of Mitigation Scenario For 150nodes

(iv). Comparison of all the three above stated graphs: In first scenario of 150 nodes of our experimentation, packets travels are shown as throughput with peak value of approx. 3797692 bits per seconds and it is represented as bits per second. In second scenario which is with Sybil attack, packets drops which are represented as throughput, decreases to value of approx. 3372836 bits per second. But when we apply the proposed algorithm for mitigation of the attack the throughput increases to a value of approx. 3586683 bits per second. These observations are for the interval of 4oo seconds. Thus the results shows that the throughput has increased considerably after applying the Sybil attack for 150nodes.

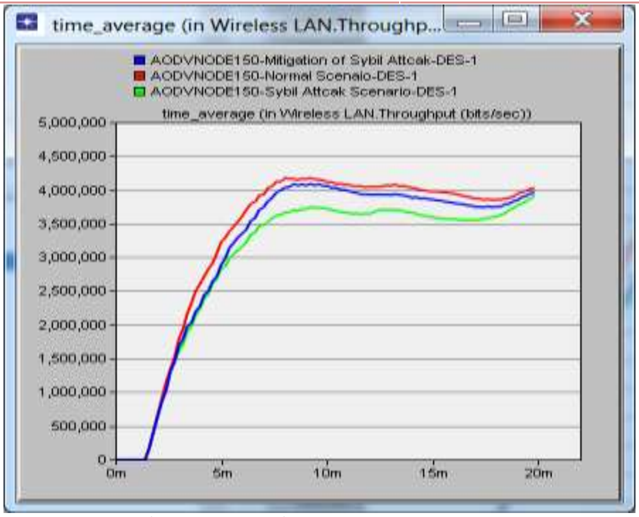


Figure 8: Graph of Comparison of All The Three Scenarios For 150nodes

B. *Throughput Analysis for 200 nodes:* Now we will show our results on 200nodes. We will present our results in three different graphs that are normal scenario (i.e. without attack), scenario with Sybil attack on it and scenario with proposed algorithm applied on it. At the last we will do comparison among the three by placing them on a single graph.

(i). The throughput evaluated below is for normal scenario for 200 nodes: The first graph of normal scenario i.e. the fig 9 of 2000 nodes of our experimentation, packets travels are shown as throughput with peak value of approx. 1895529 bits per seconds and it is represented as bits per second.

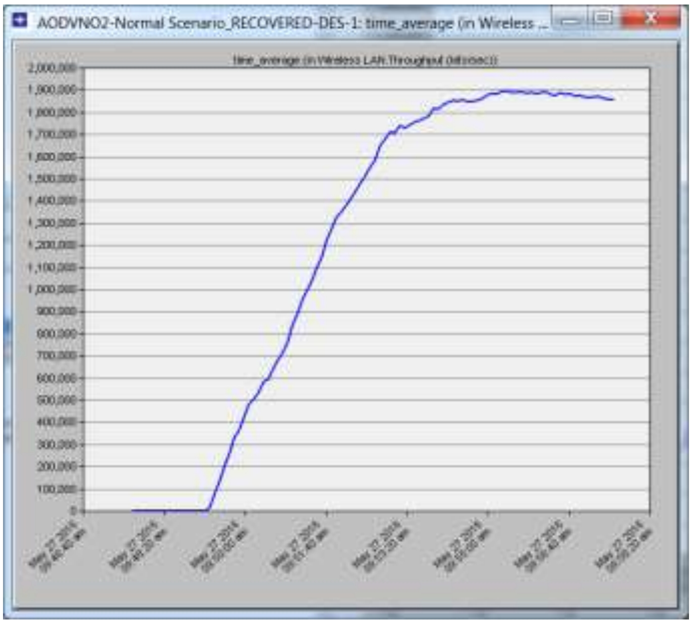


Figure 9: Graph of Throughput of Normal Scenario for 200nodes

(ii). The throughput evaluated below is for Sybil nodes for 200 nodes:.Fig 10 i.e. the graph showing the throughput of Sybil attack for 200 nodes of our experimentation, packets travels are shown as throughput with peak value of approx. 1455005 bits per seconds and it is represented as bits per second.

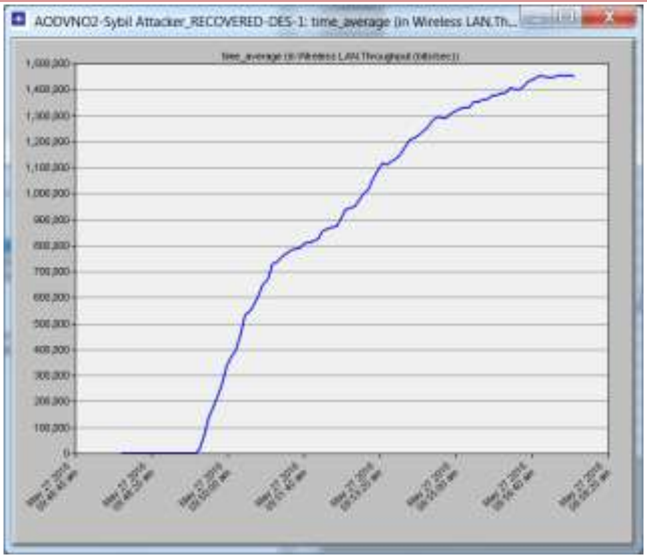


Figure 10: Graph of Throughput of Sybil Node for 200nodes

(iii). The throughput evaluated below is for mitigating the attack for 200 nodes: Fig 11 i.e. the graph showing the throughput of mitigation of Sybil attack for 200 nodes of our experimentation, packets travels are shown as throughput with peak value of approx. 1743471 bits per seconds and it is represented as bits per second.

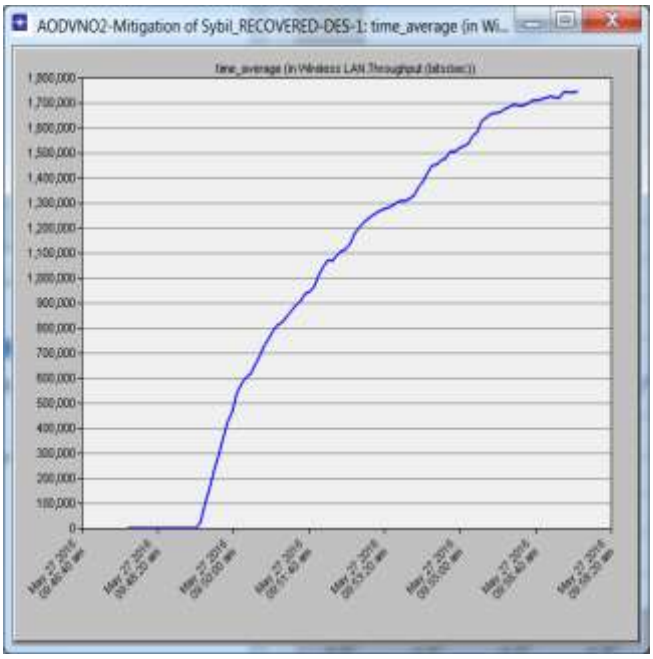


Figure 11: Graph of Throughput of Mitigation Scenario for 200nodes

(*iv*). *Comparison of all the three above stated graphs:* This drop of packets in form of throughput is due to the Sybil effect. The recovery of the throughput takes place with proposed mechanism by elimination of the Sybil attack as throughput comes to similar to the normal scenario. Thus from the graphs we conclude that the proposed algorithm proves to be an efficient approach in mitigating the Sybil attack at both the scenarios that is with 150 nodes and with 200 nodes.

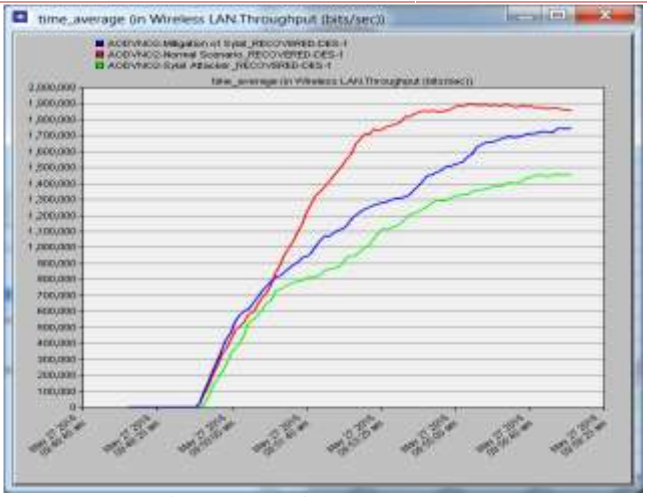


Figure 12: Graph of Comparison of All The Three Scenarios For 200nodes

VI. CONCLUSION

VANET is the most reliable wireless network for communication between the vehicles and road side support. There are various approaches that can be implemented to detect and mitigate the Sybil attack. But here in this paper we have applied genetic algorithm for detection and prevention of Sybil attack. Then we have analyzed throughput for efficient deployment of genetic algorithm for mitigating the attack. In our simulation results, we have presented results for the network with and without the attack and compared the graphs.

REFERENCES

[1] Arpita M. Bhise and Shailesh D. Kamble, “Review On Detection and Mitigation Of Sybil Attack in the Network”, International Conference on Information Security and Privacy, pp 395-401, 2016.

[2] Bhushan Vidhale and S.S. Dorle, “Performance Analysis of Routing Protocols in Realistic Environment for Vehicular Ad Hoc Networks”, IEEE, Vol. 3, Issue 11, 2011.

[3] Harsimrat Kaur and Preeti Bansal, “ Efficient Detection & Prevention of Sybil Attack in VANET”, International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 9, Sept 2015.

[4] K.Malathi and R.Manavalan, “Detection and Localization of Sybil Attack inVANET: A Review”, International Journal For Research In Applied Science And Engineering Technology, Vol. 2 Issue 9, pp 282-293, Sept 2014.

[5] Mit H. Dave, Samidha Dwivedi Sharma, “ Improved Algorithm Forintrusion Detection Using Genetic Algorithm And Snort”, International Journal of Emerging Technology and Advanced Engineering, Vol. 4, Issue 8, Aug 2014.

[6] [R. Amuthavalli and R.S. Bhuvaneswaran, “Genetic Algorithm Enabled Prevention of Sybil Attacks for LEACH-E”, Modern Applied Science, Vol. 9, Issue 9, pp 41-49, 2015.

[7] Ruhika Badhan and Bhubneshwar Sharma, “An efficient method to detect Sybil attack using genetic optimization algorithm in Manet”, International Journal of Advanced Multidisciplinary Research (IJAMR), Vol 2, Issue 10, pp 39-41, 2015.

[8] Sakshi Gupta and Taranjit Singh Aulakh, “Prevention of Sybil Attacks in VANETS Using Genetic Approach”, International Journal of Computer Science and Mobile Computing, Vol. 4, Issue 12, pg.88 – 102, Dec 2015.

[9] Tong Zhou, Romit Roy Choudhury, Peng Ning, and Krishnendu Chakrabarty, “Sybil Attacks Detection in Vehicular Ad Hoc Networks”, IEEE Journal On Selected Areas In Communications, Vol. 29, Issue 3, pp 582-594, Mar 2011.

[10] V. Geetha Devi, P.Shakeel Ahmed, P.Babu, "A Route map for Detecting Sybil Attacks in Urban Vehicular Networks", International Journal of Modern Engineering Research, Vol.3, Issue 2, pp 1157-1160, 2013.

[11] Wenshuang Liang, Zhuorong Li, Hongyang Zhang, Shenling Wang, and Rongfang Bie, "Vehicular Ad Hoc Networks: Architectures, Research Issues, Methodologies, Challenges, and Trends", International Journal of Distributed Sensor Networks, 2014.

[12] Yipin Sun, Rongxing Lu, Xiaodong Lin, Jinshu Su, Xuemin (Sherman) Shen, " A Novel and Efficient Hashchain based Certificate Management Scheme for Vehicular Communications", IEEE, pp 1-6, 2009.

[13] Shefali Khatri, Punit Sharma and Arvind Negi, "Thwarting Sybil Attack using ElGamal Algorithm", International Journal of Computer Applications, Vol 121, Issue 21, pp 44-48, Jul 2015.

[14] Sherali Zeadally, Ray Hunt, Yuh-Shyan Chen, Angela Irwin and Aamir Hassan, "Vehicular ad hoc networks (VANETS): status, results,and challenges". Springer, pp 217-241, Dec 2010.

[15] Shivani Kanwar, Sandeep Joshi and Manu Sood, "Detection of Sybil Attack in VANETs by Trust Establishment in Clusters", International Journal of Computer Engineering and Applications, Vol. 7, Issue 1, pp 51-60, Jul 2014.