Improvement of WSN Parameters Affected by Sybil Attack Using GA

Dinesh Mittal
M.Tech [ECE]
GZS Campus CET, Bathinda, India

***Abstract*— Wireless Sensor Network (WSN) consists of a set of communicating wireless mobile nodes or devices that do not have any form of fixed infrastructure or centralized authority. The security in WSN has become a significant and active topic within the research community. This is because of high demand in sharing streaming video and audio in various applications, one WSN could be setup quickly to facilitate communications in a hostile environment such as battlefield or emergency situation likes disaster rescue operation. In spite of the several attacks aimed at specific nodes in WSN that have been uncovered, some attacks involving multiple nodes still receive little attention. A reason behind this is because people make use of security mechanisms applicable to wired networks in WSN and overlook the security measures that apply to WSN. Furthermore, it may also have to do with the fact that no survey or taxonomy has been done to clarify the characteristics of different multiple node attacks.**

***Keywords*- WSN, Security, Genetic Optimization Algorithm, AODV**

I. INTRODUCTION

A network is a cluster connected with 3 or many notebook systems that are paired alongside for you to talk collectively. The WSN is an accumulation of wireless devices or hubs that impart by dispatching packets to each other or for another device/hub, without having any framework controlling information for routing. WSN hubs have boundless network and versatility to different hubs. WSN s experience the ill effects of a mixed bag of security attacks and dangers, for example, Denial of Service (DoS), flooding attack, mimic attack, wormhole attack, black hole attack, etc.

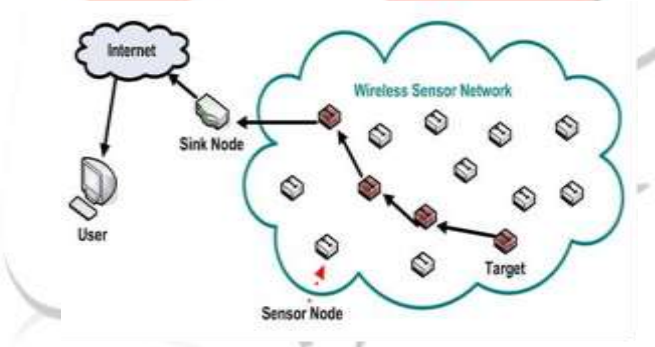

Fig. **Error! No text of specified style in document.**: Wireless Sensor Network Model

Wireless sensor networks consist of a number of sensing nodes which are distributed in a wide area. They sense an event occurring in the environment and these sensing nodes are distributed or placed according to the requirements of the application. A Wireless Sensor System includes a gang of nodes connected with typically low functionality. They work with others collectively to execute realizing tasks during granted surroundings.

Sybil Attacks

Security and antivirus programming is essential for any system. Restricted security can separate is in a Sybil attack. Sybil attack is a kind of security risk when a hub in a system guarantees various characters.

Most systems, similar to a shared system, depend on assumptions of personality, where every node speaks to one character. A Sybil attack occurs when an unreliable node is captured to claim different characters. Issues emerge when a reputation system, (for example, a record sharing reputation on a system) is deceived into believe that an attacking node has a disproportionally vast impact.

Correspondingly, an attacker with numerous personalities can utilize them to act maliciously, by either taking data or disturbing correspondence.

Initially depicted by Microsoft specialist John Douceur, a Sybil attack depends on the way that a system of PCs can't guarantee that every processing component is an unmistakable, physical PC. Various powers have tried to set up the identities of PCs on a system (or hubs) by utilizing software, for example, VeriSign, utilizing IP locations to recognize hubs, passwords and usernames, etc.

Sybil attacks have showed up in numerous situations, with wide usage for security, wellbeing and trust. For instance, a web survey can be fixed utilizing various IP locations to present countless. A few organizations have likewise utilized Sybil attack to increase better appraisals on Google Page Rank.

Sybil Attack Taxonomy

To better understand the implications of the Sybil attack and how to protect against it, taxonomy of its different forms has been discussed below:

Dimension I

Fabricated Identities

In fabricated identities, attacker tries to become the node type that is normal among various nodes, actually it produces the identity of other nodes as a legitimate node. Mainly it is found in distributed systems. E.g. if a node is known by 32-bit integer than the legitimate node also tries to become the 32-bit integer

Stolen Identities

Attackers don't fabricate new identities and steal the identity of the legitimate node. This identity theft may go undetected if the attacker destroys or temporarily disables the impersonated nodes.

Dimension II

Direct Communication:

It is one approach to perform the Sybil attack is for the Sybil hubs to communicate directly. Similarly, messages sent from Sybil hubs are really sent from one of the malicious nodes.

Indirect Communication:

In this form of the attack, no genuine hubs have the capacity to communicate with the Sybil hubs. Rather, one or a greater amount of the malicious nodes claim to have the capacity to achieve the Sybil hubs. Messages sent to a Sybil hub are routed through one of these malicious hubs, which put on a show to go on the message to a Sybil hub.

Dimension III

Simultaneous: The attacker may attempt to have his Sybil identities all partake in the system.
Non-Simultaneous: The attacker can have large number of identities by having one identity at one time.

Ad Hoc On-Demand Distance Vector (AODV) Protocol

AODV is one of reactive routing protocol means that only on demand it creates path to destination node. AODV is an on – Demand routing protocol which is confluence of DSDV as well as DSR. Route is computed on request, at the same time as it is in actual DSR by means of route detection process. That is why it is called reactive protocol. However, AODV maintains a routing table where it maintains one entry per destination unlike the DSR that maintains multiple route cache entries for each and every destination. AODV make available loop free routes even though mending link breakages but then again nothing like DSDV, it doesn't necessitate worldwide periodic routing advertisements.

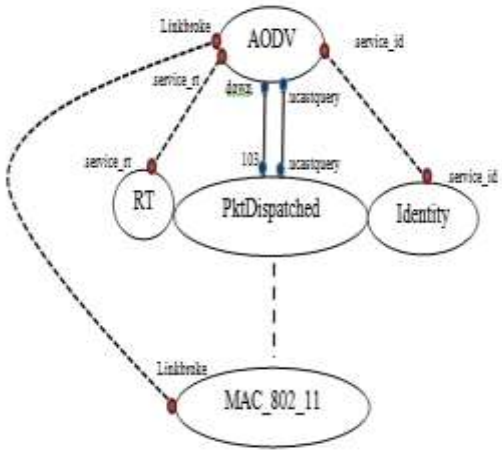

Fig. 2: Route Discovery in AODV

GA (GENETIC ALGORITHM)

According to Goldberg et al., 1989, GA is commonly used in applications where search space is huge and the precise results are not very important. The advantage of a GA is that the process is completely automatic and avoids local minima. The main components of GA are: crossover, mutation, and a fitness function. A chromosome represents a solution in GA. The crossover operations used to generate a new chromosome from a set of parents while the mutation operator adds variation. The fitness function evaluates a chromosome based on predefined criteria. A better fitness value of a chromosome increases its survival chance. A population is a collection of chromosomes. A new population is obtained using standard genetic operations such as single-point crossover, mutation, and selection operator. As a GA is relatively computation intensive, this chapter proposes executing the algorithm only at the base station. The proposed GA is used to generate balanced and energy efficient data aggregation trees for wireless sensor networks.

Step 1: At random, produce an initial population M (0).

Step 2: Compute as well as help save the actual fitness f(m) for every specific individual m in the current population M(t);

Step 3: Specify selection probabilities p (m) for every specific individual m throughout M (t) making sure that p (m) is actually proportional to f (m).

Step 4: Crank out M (t+1) by simply probabilistically choosing individuals from M (t) to produce offspring via genetic operators.

Step 5: Repeat step 2 until satisfying solution is actually attained.
However, in order for genetic algorithms to work effectively, a few criteria must be met:

- It must be relatively easy to evaluate how "good" a potential solution is relative to other potential solutions.
- It must be possible to break a potential solution into discrete parts that can vary independently. These parts become the "genes" in the genetic algorithm.
- Finally, genetic algorithms are best suited for situations where a "good" answer will suffice, even if it's not the absolute best answer.

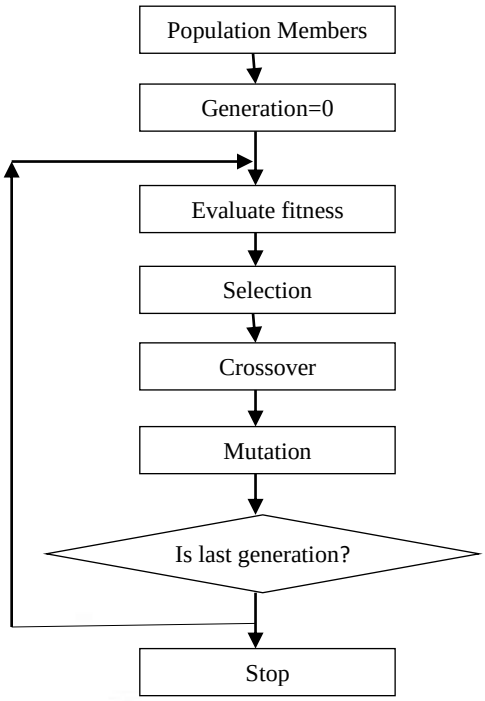

Fig. 3: Genetic Algorithm

II. RELATED WORK

Margolin et al. anticipated a recurrent fee per joining identity with the purpose of deter Sybil attackers in addition they also recommend that periodic fee is a resilient deterrent than a one-time fee. The periodic fee may possibly not be a monetary reliant payment mechanism, but then again it can also be there as a non-economic fee mechanism for example charged SMS messages, CAPTCHAs, or collaboration in the network. Barter has shown a behavior-based access in addition to admittance control structure intended for MANETs in this the nodes at first interchange their behavior profiles as well as compute separate local descriptions of normal network behavior. Scott M.Thede has discussed about Genetic Algorithm that the population is the accumulation of candidate solutions that are considered throughout the calculation. Over the generations of the calculation, new individuals are "conceived" into the population. James Newsome et al. has talked about Sybil Attack in Sensor Networks in and clarified the Sybil attack is an attack in which a single entity can control a considerable division of the system by showing different identities. Sujatha, et al. has proposed a procedure to dissect the introduction to attacks in AODV. The proposed framework is considered by taking into account Genetic Algorithm, which investigates the practices of each hub and gives insights about the attacks. The execution of MANET is investigated taking into account GAC. Anup Goyal and Chetan Kumar have suggested a systematic learning method known as Genetic Algorithm (GA), to identify illegitimate nodes. The algorithm considers the varied features in network connectivity like protocol type, network service to destination and connection status to generate a type based rules.

III. SIMULATION MODEL

Genetic Algorithm is a method of soft computing which uses the laws of selection and evolution. In computer network security; it is mainly used to find an optimal solution to a problem. The Genetic Algorithm starts by identifying a data set called population. Then these are individually encoded using bits, characters or integers and they form a chromosome. The next operation on them is an 'Evaluation Function' used to determine the genuine chromosome. During this process, two different operations namely, crossover and mutation are performed which is used to imitate the breeding and evolution. The selection of the chromosome is biased towards the fittest of the species. At last, the fit chromosome is selected once the optimization criterion is met. Figure illustrates the basic functionality of Genetic Algorithm.

Genetic Algorithm can be used to devise elementary principles for traffic in networks. These principles are used to distinguish between genuine connections against malicious ones. The malicious ones refer to the objects with illegitimacy. The rules follow the upcoming syntax and they are selected only if a particular condition is met.

The condition specified in the above syntax is usually a single or multiple network parameters. The algorithm starts initially with the population of nodes. The number of nodes in the base network is varied and hence the initial population is also variable. Then the

next step in this ID is encoding, where the members of the initial population are encoded using binary values and they are called as chromosomes.

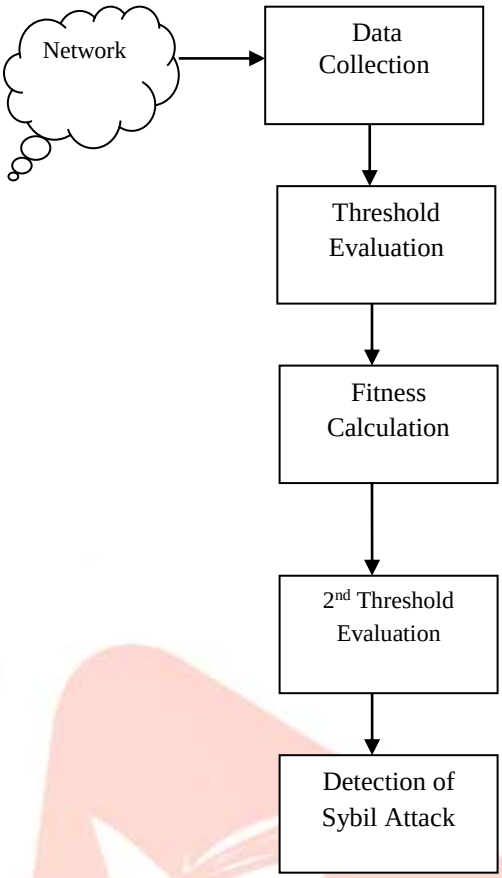

Fig. 4: Proposed Model

Each chromosome is then evaluated for an objective function by considering the various network parameters like packet drop (PD), Request Forwarding Rate (RFR), Request Receive Rate (RRR) etc. Then the threshold is determined by calculating the average of the individual network parameters. Then the fitness criterion for each every network parameter is determined based on Tournament Selection which includes Fitness remapping wherein the fit nodes are assigned '0' values and then the second threshold is determined as the weighted average of the network parameters. Then the surviving Sybil nodes are the ones with the value of all the optimal parameters to be zero. Thus these nodes are determined and plotted versus their node identification number.

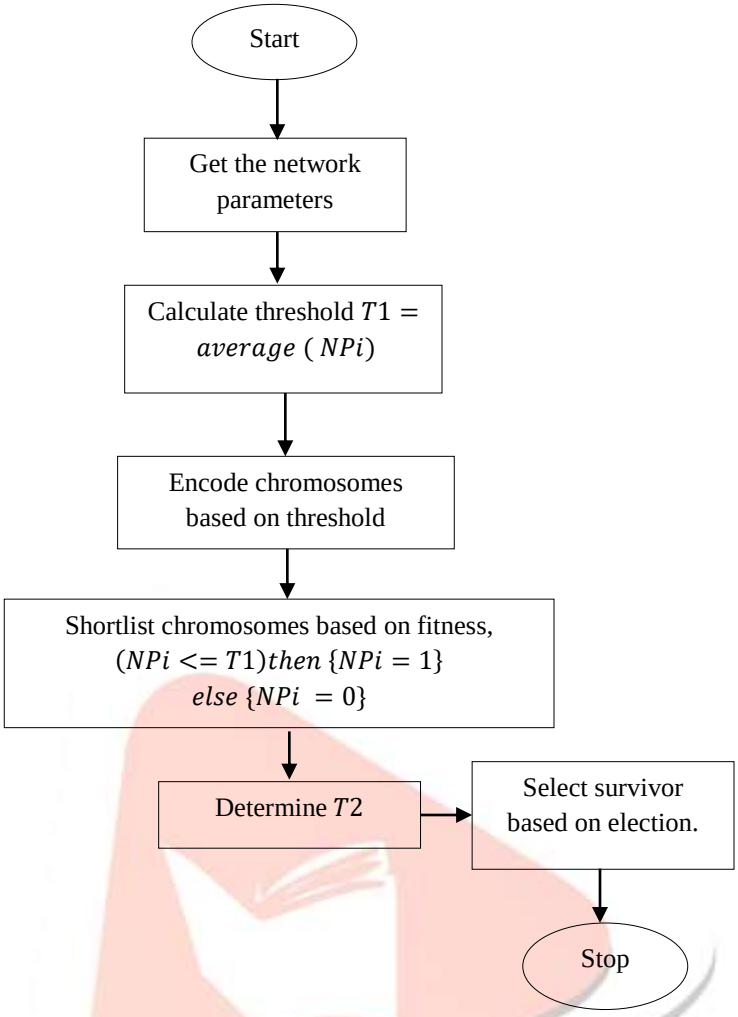

Fig. 5: Simulation Model

IV. SIMULATION RESULTS

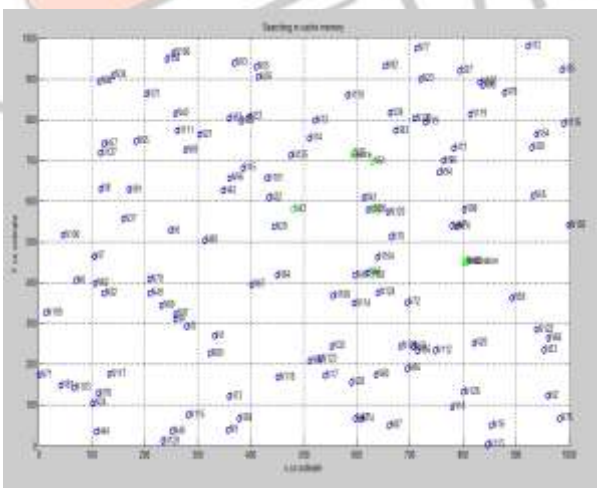

Fig. 6: Network Deployment

Figure 6 shows the simulations that were carried out by using MATLAB as the language that we use to develop the proposed framework. We used the AODV protocol to modify the network parameters that we added to the simulator and evaluate our proposed framework based on it.

In the simulation the following steps are to be followed by user:

1. Firstly ENTER the number of nodes
2 Enter the length and width of the network.
3. Enter the cluster heads and number of rounds to run for the network.
4. Then the cluster head should be plotted by the user by comparing the energy of each node with the threshold energy.

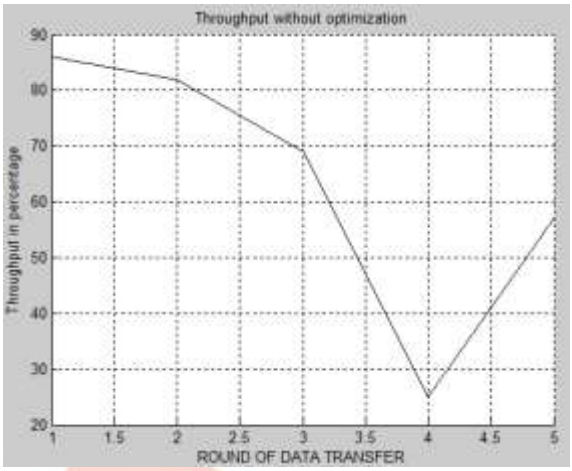

Fig. 7: Throughput without Optimization

The average throughput has been shown in above figure and it shows that the throughput value without optimization algorithms and obtained values is 59%.

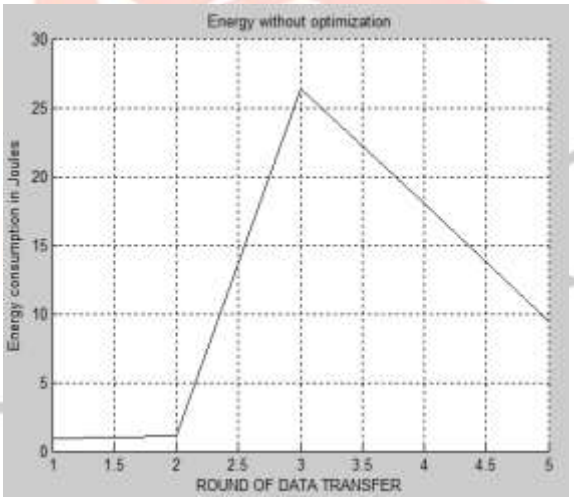

Fig. 8: Energy without Optimization

Energy consumption is the consumed energy by the nodes during transmission of data. Above graph gives the un-symmetrical view in contrast to energy consumption rate for different rounds.

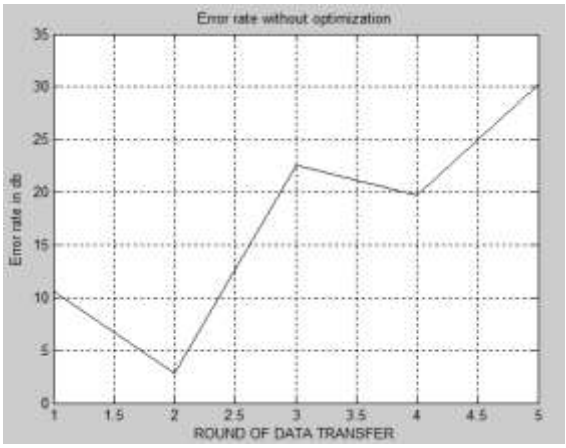

Fig. 9: Error rate without Optimization

With an offered delay of 30 for 5db. Above figure shows the Bit error rate values without optimization algorithms and obtained value is 30. Curve also demonstrates that BER has steep curve value.

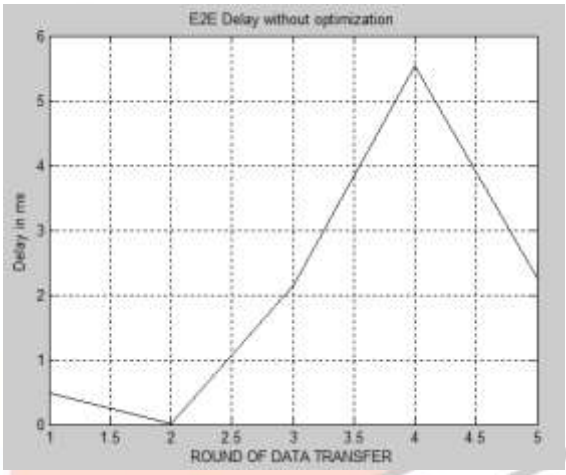

Fig. 10: End to End Delay without Optimization

Delay basically occurs when some nodes transfer data in slow speed or when there occurs a traffic in whole network. Above figure shows the delay values without optimization algorithm and obtained values is 5.1.

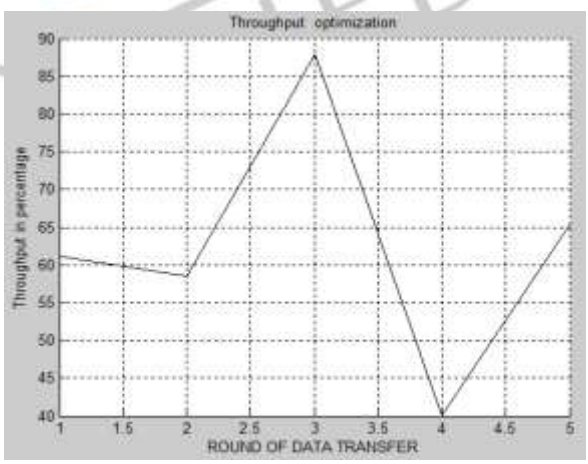

Fig. 11: Throughput with Optimization

The Sybil attacker node causes decrease in the throughput of the network. It is because number of collisions is more in system and it is optimized using GA algorithm as shown above.

Above figure shows that throughput value with GA/AODV. Total data from the source to the receiver more than the time it takes until the recipient receives the last packet. Less time translates into higher productivity.

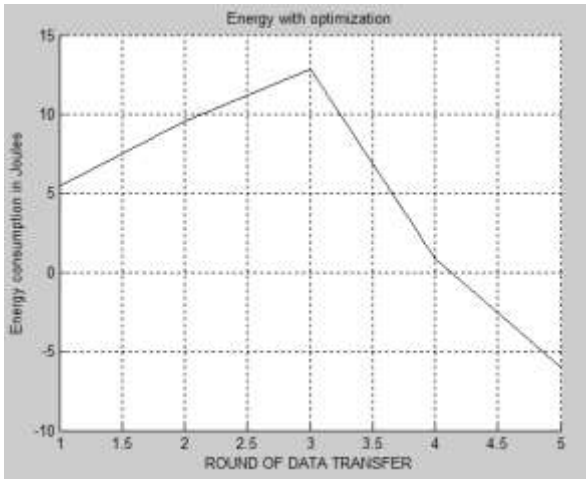

Fig. 12: Energy with Optimization

Energy Consumption is the range of the energy consumed by each node for transmission of data packets. To consume nodes efficiently and wisely is one of the important features of sensor networks. As sensor nodes are prepared with non- chargeable batteries with inadequate energy supply, a sensor network cannot works well after a fraction of nodes run out of energy. Above figure shows the energy consumption of 12 J.

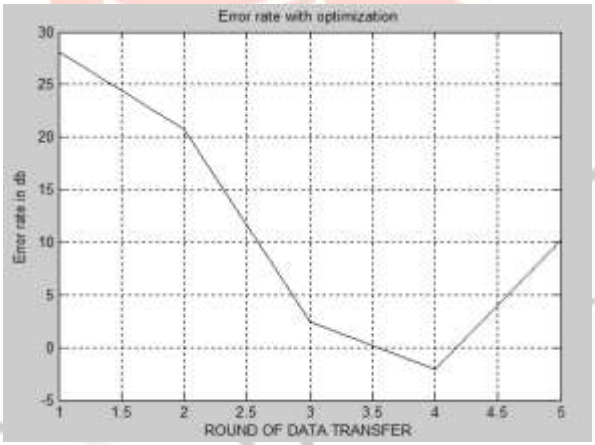

Fig. 13: Error rate with Optimization

In regard of Bit Error Rate, above figure shows that BER decreases with respect to GA i.e. 30 to 11 because the algorithm reduces the number of collisions so that it can reduces the number of packets drops caused by collisions.

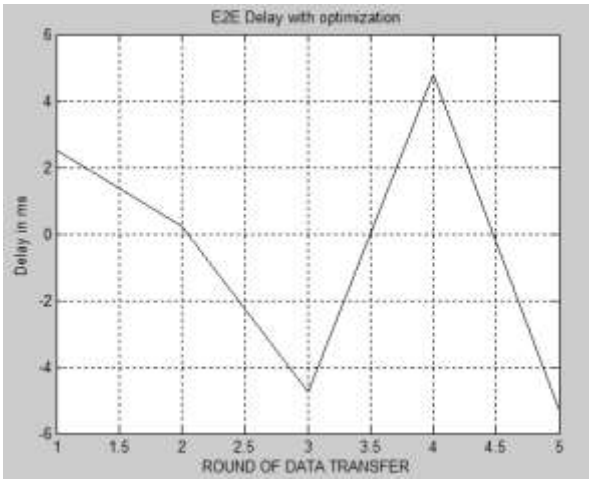

Fig. 14: E2E Delay with Optimization

The End to End Delay is a significant parameter for evaluating a protocol which must be low for good performance. Above figure shows the end to end delay with GA/AODV.

V. CONCLUSION

Peer-to-peer systems play an ever-increasingly important part of our daily lives. However, most of the peer-to-peer systems are vulnerable to Sybil attacks. In order to design more efficient and practical Sybil defenses, an implementation based on Genetic algorithm is presented.

In this thesis, the issues related to security like Sybil attack has been reviewed. Then an Intrusion Detection System (IDS) especially for Sybil attacks is implemented using Genetic Algorithm, and then tested with networks of varied node configurations. The algorithm will be tested for more number of nodes and the performance analysis will be done in terms network load, throughput of the algorithm as the node number is increased. It is concluded that Sybil attack prevention is achieved at greater rate when GA has been used.

REFERENCES

[1] C.Piro, C.Shields and B.N.Levine (2006), "Detecting the Sybil attack in mobile ad hoc networks," in Proc. Securecomm Workshops 2006, pp.1–11.

[2] E.E.Khin and T.Phyu (2014), "Impact of black hole attack on AODV routing protocol", International Journal of Information Technology, Modeling and Computing, vol. 2, pp. 66-70.

[3] James Newsome, Elaine Shi, Dawn Song and Adrian Perrig (2004), "The Sybil Attack in Sensor Networks: Analysis and Defenses", IPSN 04, April 26-27, 2004, Berkeley, California, USA.

[4] K.S.Sujatha, V.Dharmar and R.S.Bhuvaneswaran (2012), "Design of Genetic Algorithm based IDS for MANET", International Conference on Recent Trends in Information Technology (ICRTIT), IEEE, pp.28-33.

[5] Kalpana Sharma and M.K.Ghose (2010), "Wireless Sensor Networks: An Overview on its Security Threats", International Journal of Computer Applications (IJCA) special issue on MANETs, pp.42-45.

[6] P.Raghu Vamsi and Krishna Kant (2014), "Sybil Attack Detection using Sequential Hypothesis Testing in Wireless Sensor Networks", International Conference on Signal Propagation and Computer Technology (ICSPCT), IEEE, pp.698-702.

[7] Pengfui Guo, Xuezhi Wang and Yingshi Han (2010), "The Enhanced Genetic Algorithms for the Optimization Design", 3rd International Conference on Biomedical Engineering and Informatics (BMEI 2010), IEEE, pp.2990-2994.

[8] Rajeshwar Singh (2011), "Performance Evaluation of DSR and DSDV Routing Protocols for Wireless Ad Hoc Networks", International Journal for Advanced Networking and Applications 732 Vol. 02, issue. 04, pp. 732-737.

[9] S.Hazra and S.K.Setua (2012), "Sybil attack defending trusted AODV in ad-hoc network", 2nd International Conference on Computer Science and Network Technology (ICCSNT), IEEE, pp.643-647.

[10] Simranjeet Kaur, Gagangeet Singh Aujla and Sahil Vashist (2014), "Detection and Optimisation Techniques against Sybil Attack on MANET", International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), vol. 4, issue 8, pp.369-375.

[11] T.N.Manjunatha, M.D.Sushma and K.M.Shivakumar (2013), "Security Concepts and Sybil Attack Detection in Wireless Sensor Networks", International journal of Emerging Trends and Technology in Computer Science (IJETTCS), vol. 2, issue 2, pp.383-390.

[12] T.G.Dhanalakshmi, Dr.N.Bharathi and M.Monisha (2014), "Safety Concerns of Sybil Attack in WSN", International Conference on Science Engineering and Management Research (ICSEMR), IEEE, pp.1-4.

[13] Yuteng Guo, Beizhan Wang, Xinxing Zhao, Xiaobiao Xie, Lida Lin and Qingda Zhou (2010), "Feature selection based on Rough set and modified genetic algorithm for intrusion detection", 5th International Conference on Computer Science and Education (ICCSE), IEEE, pp.1441-1446.

[14] Zolidah Kasiran and Juliza Mohamad (2014), "Throughput Performance Analysis of the Wormhole and Sybil Attack in AODV", Fourth International Conference on Digital Information and Communication Technology and it's Applications (DICTAP), IEEE, pp.81-84.