Genetic Algorithm Enabled Prevention of Sybil Attacks for LEACH-E

R. Amuthavalli[1] & R. S. Bhuvaneswaran[2]

[1]Assistant Professor, SSN College of Engineering, Chennai, India

[2] Associate Professor, Anna University, Chennai, India

Correspondence: R.Amuthavalli, SSN College of Engineering, Chennai, India. Tel: 9-194-4388-7898. E-mail: amuthavallir@ssn.edu.in

Abstract

Wireless Sensor networks are deployed in hostile environments for critical application, especially in the military and civil domains. The sensor nodes depend uponbattery power. Sensor nodes utilize more energycompared to a normal node. This may increase delays and reduce the packet delivery ratio that cause attacks in the network. The Sybil attack is one of the dangerous attacks against sensor and ad-hoc networks, where a node illegitimately claims multiple identities. The aim of the cluster based Hierarchy routing protocol LEACH-E(Low Energy Adaptive Clustering Hierarchy-Energy) is to provide secure routing and to preserve the functionalities of the original protocol. This energy efficient protocol always elects a Cluster Head (CH) based on high energy among the cluster group.Here we propose a LEACH-E-GA for Intrusion detection (ID) in Wireless Sensor Nodes. The Genetic Algorithm is deployed into LEACH-E to provide prevention for Sybil attacks. The objective of this Genetic Algorithm (GA) is to identify its best trusted neighbors for communication using its optimization capability. LEACH-E-GA reduces an inside attack in WSN and shows reliable transmission with improved network efficiency, reduced delay and increased packet delivery ratio.

Keywords: wireless sensor networks, sybil attacks, secured rotuing protocol, genetic algorithm, leach-e, Cluster head

1. Introduction

Wireless Sensor nodes are compact, light-weighted, and battery-powered devices that can be deployed in virtually any environment. The sensor nodes are deployed in an adhoc fashion and they can communicate through a wireless medium. Here routing typically begins with neighbor discovery. Low Energy Adaptive Clustering Hierarchy (LEACH) (AlakeshBraman et al. 2014;Meena Malik et al. 2013) is the routing protocol that is employed for conventional routing. This Cluster based protocol helps improve the lifetime of wireless sensor network. The nodes are arranged into small groups called clusters to reduce energy dissipation. LEACH protocol, which is the first hierarchical cluster-based routing protocol, is divided into rounds; each round consists of two phases namely setup phase and steady phase.

In the setup phase, the nodes in the cluster decide its Cluster Head (CH) based on the signal strength and energy and the cluster head broadcasts its information to all nodes. In the steady state phase, all the nodes in the cluster group are send data to the cluster head. Data is aggregated by the cluster head and sent to the Base Station (BS). Thereby, the steady state phase much longer than the setup phase.

1.1 Sybil Attacks

Sybil node is a solitary node, which acquires the multiple characteristics, mainly the identification of the other nodes in the network. The Sybil attack is one of the primary attacks, which pave way for various other attacks to take place in the network. The Sybil nodes cannot be detected directly by checking only the ID or the node information.

In Section 1 the Leach protocol was discussed in detail. Section 2 presents the related work. Section 3 describes the problem statement and Section 4 introduces the proposed protocol **LEACH-E-GA.** Section 5 summarizes our protocol simulation results and the paper conclusion along with future enhancement is provided in Section 6.

2. Related Work

In the LEACH and LEACH-E (AlakeshBraman et al. 2014) protocols, the communication between cluster heads

and the base station requires more energy than the non-cluster nodes. This means increasing the number of clusters-heads can increase the energy consumption of the whole network and shorten the network lifetime. Therefore, it is necessary to select the optimal number of cluster heads to make the energy consumption minimum. The original LEACH-E algorithm, selects the cluster heads at random with fixed round time for the selection. It considers the remnant power of the sensor nodes in order to balance network loads and changes the round time depending on the optimal cluster size. In LEACH-C (Shuo Shi et al. 2012;Petre-CosminHuruială et al. 2010; Raed M. Bani Hani andAbdalraheem A. Ijjeh. 2013) protocol, each node transmits its information to the corresponding base station and the sink nodemakes the choice of selecting the cluster head and how to divide clusters. Then the cluster head sends this information to BS. In Hierarchy routing protocol a CH collect a data from its cluster members, aggregates all data and forward to the BS that might be located far away from it. If the CH is compromised then it will be dropped. The compromised CH will become ineffective, because the data aggregated by cluster head will never reach the base station. V-LEACH (BaniYassein. M et al. 2009) protocol, besides having a CH in the cluster, also has a vice-CH that takes the role of the CH when the CH is dropped/compromised. The vice-cluster nodes forward data directly to the BS. Messy GAs solve (Goldberg . D et al.1989) problems of coverage of local maxima by the optimal search. To choose the best CH, minimizing the energy consumption and latency is obtained by choosing the best nodes in the network. A genetic algorithm is executed on a central BS and the results are send to the nodes (Goldberg . D et al.1989). Hierarchical routing protocol (Vikram Mehta andDr.Neena Gupta. 2012) due to a battery replacement or recharging is not realistic. Choosing the routing protocol is, it must be energy-efficient to improve the network lifetime (Yang Yu et al. 2006;Manimozhi. B &, Santhi.B.2013). The optimal set of protocols is proposed to show the optimization in genetic algorithm metrics for WSNs with the QoS requirements (JiaXu,Ning et al. 2012). Cluster-based LEACH routing protocol in WSN has greater energy efficiency and the information such as node's residual energy and geometric distance send to BS, to elect CH nodes. The CH node is one hop to the BS to consume less energy than other nodes because communication of data consumes the more energy. CH nodes not only consider the residual energy of the nodes and also distance between the CH and BS also examined (Jin Fan and Parish D .J. 2007). Trust-based LEACH protocol in (Nguyen Duy Tan et al. 2012) discussed the cluster-head-assisted monitoring control. Basic classification of routing protocols in WSNs (Petre-CosminHuruială et al. 2010) has named LEACH as the most energy efficient protocol giving its advantages and disadvantages.

3. Problem Statement

Wireless sensor networks are dynamic in nature and data transmitted through various numbers of intermediate nodes. Due to the mobility and dynamic nature of the sensor nodes, the intermediate nodes may change after route discovery and route-link failure occurs. Also, any intruders can join as the intermediate node in the route. The biggest challenge for the LEACH-E protocol is that they go through topological changes in the networks and thereby their energy gets drained. To conserve energy and increase network lifetime, the node should minimize the energy dissipation and optimize communication

Changing behaviour of the node is identified by the proposed LEACH-EGenetic Algorithm using its fitness functions and thus makes the node as a trustable node. Once the node's trustiness is identified, the transmission takes place efficiently in a secured manner.

4. Proposed Algorithm: LEACH-Energy-Genetic Algorithm (LEACH-E-GA)

4.1 Leach-EProtocol

LEACH-E protocol improves the CH selection procedure. Sensor node's residual energy is the main concern, which decides whether the node become a CH or not after the first round (BaniYassein .M et al2009). Like LEACH protocol, LEACH-E is divided into rounds (Shankar .M et al. 2012). In the first round, all the nodes have the same probability of being a CH. At the end of the first round, the node, which has more residual energy, is elected as CH. LEACH–E protocol improves the cluster head selection procedure.

4.2 Leach-E-Ga (Leach-Energy-Genetic Algorithm)

This paper proposes the LEACH-EGenetic algorithm (GA) that would enhance the WSN response time, network life and minimize the delay. The Genetic algorithm proposed by (Goldberg *et al* in 1975;Wu Xinhua and Wang Sheng. 2010) improves the cluster heads selection process. Selecting the minimum number of cluster heads in the WSN is determined based on the square root of the total number of sensor nodes, to minimize the total energy consumption.

The Genetic algorithm selects an unsupervised node, which allows the network to achieve maximum coverage distance with minimum energy consumption. Genetic algorithm optimizes the behaviour of the node based on its

request and response, energy level, mobility and comparison with its record of previous transmissions. A node, whose behaviour is changed and not fit to the fitness function, is considered to be the Sybil node. The node is dropped from the network to improve the quality of the network for future communication

(Wu Xinhua & Wang Sheng. 2010) enhanced the HCR protocol using GA, which determines the clusters, CHs, Cluster-members and the schedules for transmission. In this paper the GA can be used at any place in the network like base station CH, or in administration and it provides more energy efficiency by identifying the Sybil nodes to the optimizer. In each round of routing discovery, GA is applied. The optimizer chooses the best trusted neighbor nodes using the GA fitness function. The fitness function is based on the node behavior, direct distance to destination node, and energy and trust value of the nodes in the route. LEACH-E is enhanced by GA at the base station. GA creates the energy efficient clusters for more numbers of transmissions. In terms of GA representation, nodes are called [assigned] as chromosomes. The head node, member node in a cluster can be represented as a tuple < X, Y>. A population contains a constant number of chromosomes, whereas the best chromosome can be used for next new population generation.

Genetic_Algorithm ( )

{

➢   Initialize population and Objective Function Value-[OFV]

➢   Define the Fitness function

➢   Selection

➢   Cross over

➢   Mutation

➢   Repeat the above steps until reaching the solution

}

   A population contains a group of individuals named chromosomes, which represents a finished solution for a derived problem. Each chromosome is a sequence of values of the attribute [node-energy, node-trust value, and node-distance].
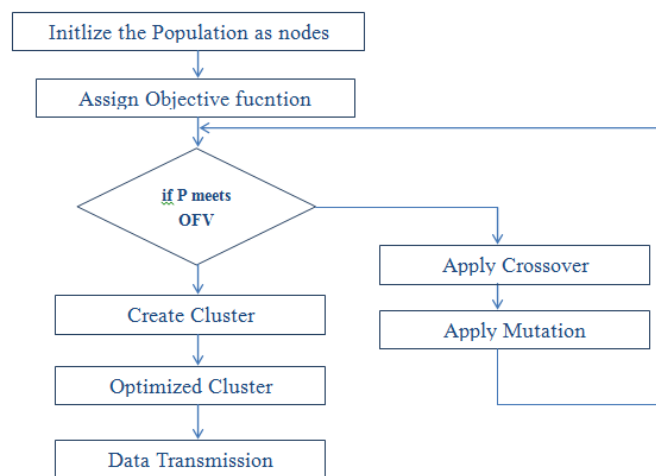

Figure 1. Flowchart for GA based WSN

4.3 Numerical Illustration

Population

All nodes in the network are considered as chromosomes. The CH and CM nodes are represented as X and Y respectively. The fitness of a chromosome is evaluated by many parameters, such as node energy, distance, node trust value. Population P consists of several chromosomes. For example, the population P is initialized as P =50 and it is generated randomly. According to the fitness the population transforms into the future generation.

Fitness Parameters

$$DD = \sum_{i=1}^{N} dist_{i,BS}$$

$$C = \sum_{i=1}^{K} dist_{i,CH} + dist_{CH,BS}$$

$$\mu = \frac{\sum_{i=1}^{HS} dist_{cluster-i}}{CH}$$

$$SD = \sqrt{\sum_{i=1}^{CH} (\mu - dist_{cluster-i})^2}$$

$$\boldsymbol{TV} = \sum_{i=1}^{N} \boldsymbol{TV(Node_i)} \geq \boldsymbol{TV_{th}}$$

$$Enry = \sum_{j=1}^{K} Enry_{Tx_{j,CH}} + K \times Enry_{Rx} + Enry_{Tx_{CH}}$$

Fitness Function

$$OBF = \sum_{i=1} (w_i \times f_i), \forall f_i \in \{C, DD, Enry, SD, TV\}$$

Where all the attributes are used as notation in the above objective constrains to represent the fitness function and it is described in the following table-1. If the node satisfied in terms of OBF, then there will be a link provides between $< X, Y>$, where X represents the node and the Y represents the cluster.

Each chromosome is evaluated according to the fitting parameters and update in each round, and it can be written as:

$$\Delta f_i = f_i - f_{i-1}$$

Δf denotes the changes in the fitness parameter values.

$$w_i = w_{i-1} + c_i \Delta f_i$$

Where, $c_i = \frac{1}{1+e^{-f_i}}$, improves the weight value compared with the previous values.

Each time the node's attribute are evaluated by the fitness function and check the arbitrary weight value. If the arbitrary weight value is the best OFV, then that chromosome is chosen as the best neighbor for transmitting data and grouped into clusters. If the node parameter has not satisfied the fitness constraints, then apply cross over on the chromosomes. Then apply mutations and compute the OFV. Repeat the same above fitness calculation until reaching the objective value. Else, replace all the bad solution based chromosomes with the newly generated chromosomes randomly for the optimization process.

Also, in this paper, it is considered to minimize latency, which means minimizing travel time of data from the end nodes to BS (Goldberget al.,1989).When each CH node sends data directly using one hop to the BS, the time is reduced, but the node consumes more energy. Minimizing energy consumption involves finding solutions where nodes communicate information on distances as short as possible and between as few nodes as possible. The results where cluster-heads are equally spread, it is found that the energy consumption is uniformly distributed in the network.

Table 1. Nomenclature

Notation	Description
i	Index and ID of the node, cluster
K	Data size
DD	Distance
C	Cluster
$Enry$	Energy
SD	Cluster Distance with Dieviationμ
TV	Trust Value
$dist_{i,BS}$	Distance between Node i and BS
$dist_{i,CH}$	Distance between Node i and CH
$dist_{CH,BS}$	Distance between CH and BS
BS	Base station
CH	Cluster Head
$cluster - i$	Ith Cluster
$Enry_{Tx_{j,CH}}$	Energy Need for Transmitting from jth node to CH
$Enry_{Rx}$	Energy Need for Receiving
$Enry_{Tx_{CH}}$	Energy Need for Transmitting to BS
w_i	Arbitrary Weight
f_i	Fitness parameter

5. Simulation and Results

The proposed schemes have been experimented in the simulation environment in NS2. The simulation parameters (Sonam Jain &SandeepSahu. 2012) are shown in Table 5.1. All the parameters such as the size of the network, what kind of propagation is going to use in the routing protocol. Which MAC layer rules are applied and the type of the antenna is used in the network model. The time duration of the entire simulation and the node deployment method with the number of nodes, number of clusters and the number of nodes in each clusters mention in the simulation settings. Finally, the node initial energy, size of the data packet used in data transmission is given in Table1.

Table 5.1. Simulation Parameters

Parameter	Level
Area	500m x 500m
Propagation Model	Two-ray ground reflection
Number of Nodes	100
MAC	802.11
Antenna	Omni directional
Simulation Time	10 s
Placement	Random
Number of Cluster	4
Nodes in Cluster	12
Node Initial Energy	90J
Equal energy (start up)	Yes
Packet size	500
Pause time	3secs

The network is deployed with 100 nodes and each of them communicates with each other. During the communication, LEACH-E-GA algorithm detects the Sybil attacks, drops those nodes and improves the network lifetime. The following graph shows the improvement of QoS parameter like packet transmission delay, Energy consumption, Packet Delivery Ratio (PDR) and Throughput compared with the normal Leach protocol.

Table 5.2. Performance matrix of protocols with change number of nodes

No.of Nodes	Normal LEACH				Genetic LEACH			
	Delay	Energy	Throughput	PDR	Delay	Energy	Throughput	PDR
20	0	0	0	0	0	0	0	0
40	10	3.62	0	0	4	12.7	18.05	0
60	70	12.74	4.03	0.1231	23	13.26	68.06	0.1231
80	88	13.05	14.06	0.4286	45	13.83	76.72	0.4286
100	114	13.25	18.85	0.5725	100	14.3	82.18	0.5725
120	160	13.47	21.61	0.6565	123	14.87	91.49	0.6565

The End-to-End delay is calculated as the average time taken by a data packet to arrive in the destination. $\sum$ (arrive-time − send-time) / $\sum$ Number of connections the lower value of the end-to-end delay means better performance of the protocol. The Comparative study of LEACH with proposed LEACH-E-GA is shown in Table 5.2. The number of nodes deployed is 20, 40, 60, 80, 100 and 120 iteratively. The performance metrics such as delay, energy, throughput and PDR computed for the existing LEACH and the Genetic LEACH is given in the Table-5.2. The overall data packet transmission is used for computing the throughput and the successful packet received at the destination is used for computing the packet delivery ratio. The energy is computed after reducing all the consumed energy from the initialized energy of the particular node. The consumed energy depends on the node participation in certain activities. From Table-5.2, it is clear that, the network performance is affected in terms of number of nodes deployed in the network.
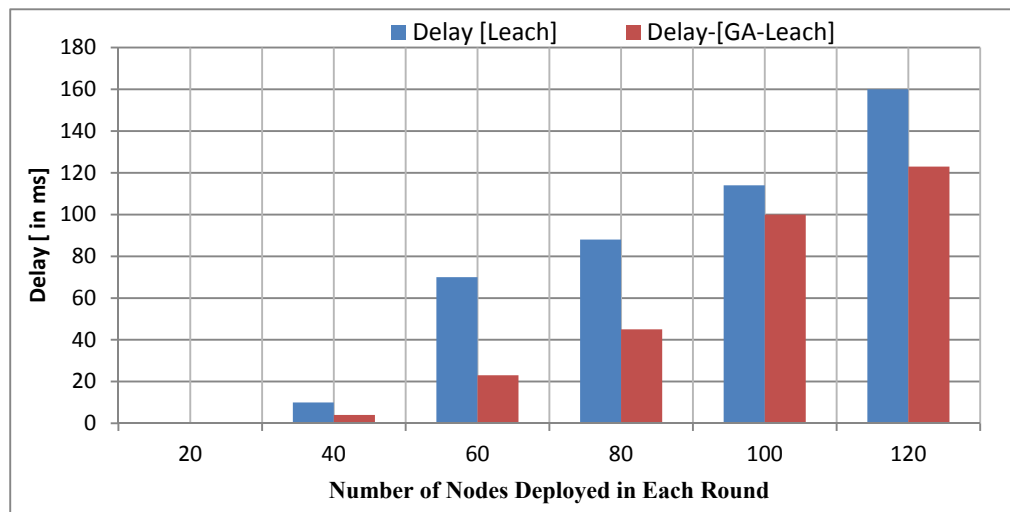

Figure 5.1 End-to-End Delay

The graph 5.1 shows that during the time increases the delay taken to transmit packes in LEACH-E-GA is minimized because of the misbehavior of the nodes eliminated at the initial stage itself.The Energy is calculated as proposed LEACH Vs Normal Leach with Genetic Leach. When the Time increases the Genetic Leach residual energy is maintained shown in Figure 5.2.In the existing LEACH, the energy efficiency is considered as the main factor and to save more energy the clustering process in applied the network. Since, the CH is changed each time of data transmission; the energy of the CH is maximized for data gathering and aggregation.

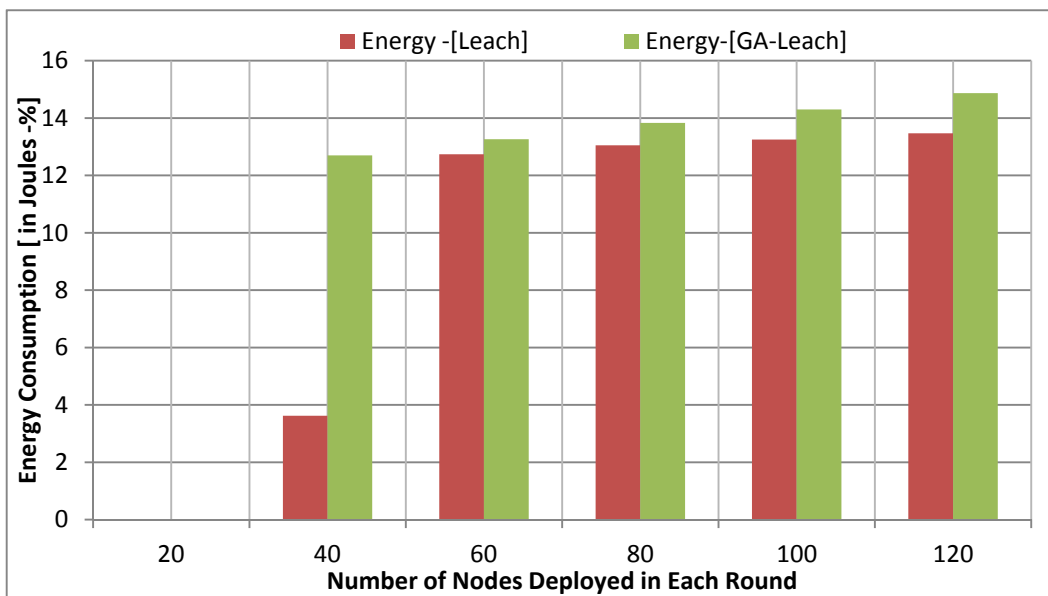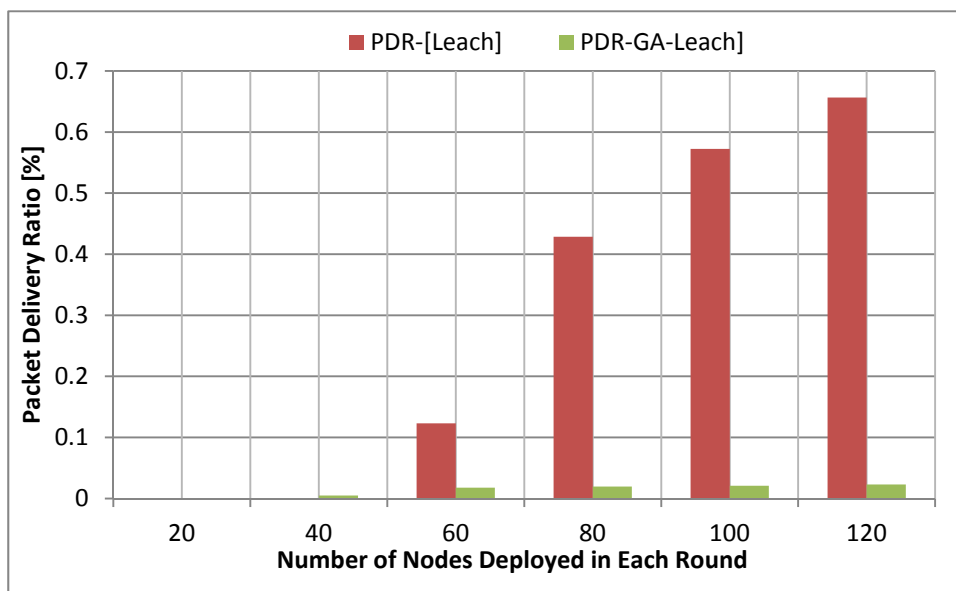Figure 5.2. Energy

Figure 5.3. Packet Delivery Ratios

In figure 5.3 packets Delivery Ratio (PDR) is the number of data packetsdelivered to the destination is calculated as the number of packets received / number of packets send. If the PDR value is more, the performance of the protocol is high. In the iteration of the time slot, the PDR value is greater than the normal Leach protocol because during data transmission the behavior of the node is controlled and misbehaving nodes are prevented from the network.
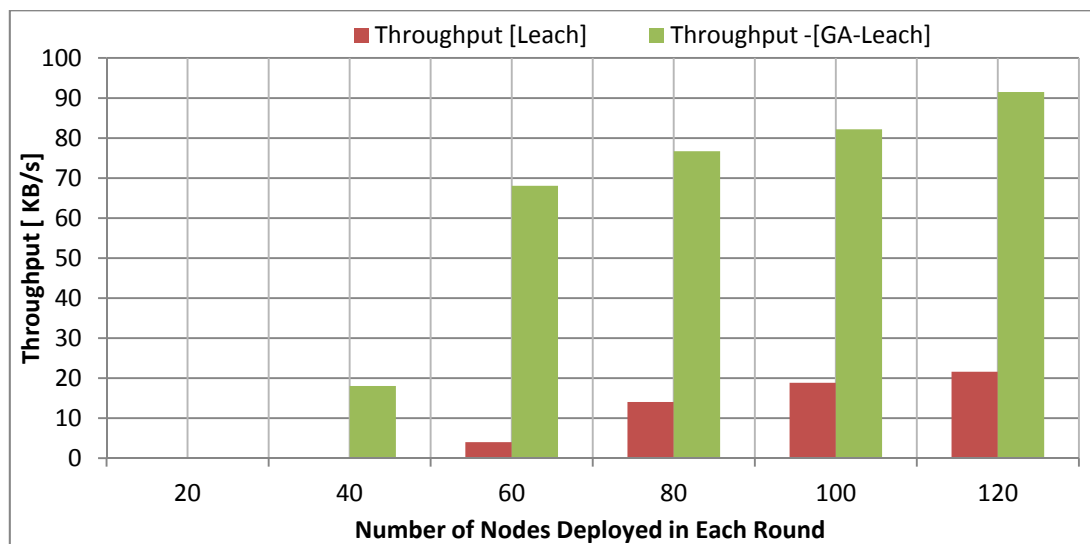
Figure 5.4. Throughput

Throughput refers to how much data can be transferred from one location to another in a given time and the total number of packets dropped during the simulation. The Sybil node is eliminated, so packet loss is less during transmission and the destination node receives the mostly all the packets delivered intend for them,sothe network performance of the protocol is increased. The performance of Genetic Leach throughput is shown in Fig. 5.4.

6. Conclusion

In this paper, it is aimed to select nodes for clustering using LEACH-E-GA in order to improve the energy efficiency with trusted nodes. Before, clustering all the nodes are optimized by the Genetic Algorithm and LEACH-E do clustering and CH election. The nodes are optimized using their attributes such as energy value, distance, trust value. From the experimental results, it is clear and concluded that the proposed LEACH-E-GA is efficient in terms of security and energy saving. This algorithm gives more effective output as proved from the graphs and tables. The packet loss is reduced in the proposed approach using a Genetic algorithm. This enables the network to continue with their further transmission without any delay and fear of attack.

6.1 Future Enhancement

In LEACH-E-GA algorithm the node behavior is controlled and network prolong lifetime is improved. With this algorithm we can extend the work to monitor the network using Intrusion Detection Protocol using Cryptography.

References

Alakesh, B., & Umapathi, G. R. (2014). A Comparative Study on Advances in LEACH Routing Protocol for Wireless Sensor Networks. *A survey International Journal of Advanced Researchin Computer and Communication Engineering, 3*(2).

Bani, Y. M. et al. (2009). Improvement on LEACH Protocol of Wireless Sensor Network(VLEACH). *International Journal of Digital Content Technology and its Applications, 3*(2).

Fan, J., & Parish, D. J. (2007). Using a Genetic Algorithm to Optimize the Performance of a wireless Sensor Network. ISBN: 1-9025-6016-7 2007 PGNet.

Goldberg, D. et al. (1989). Messy genetic algorithms: Motivation, analysis, and first results. *The Clearing house for Genetic Algorithms (TCGA),* Report 89003.

Manimozhi, B., & Santhi, B. (2013). Comparison of Different Performance Measures of Routingprotocols in WSN. *International Journal of Engineering and Technology (IJET), 5*, 208-214.

Meena, M. et al. (2013). Analysis of LEACH Protocol in Wireless Sensor Networks. *InternationalJournalof Advanced Research in Computer science and Software Engineering, 3*(2).

Nguyen, D. T. et al. (2012). An Improved LEACH Routing Protocol for Energy-Efficiency of Wireless Sensor

Networks. *Smart Computing Review, 2*(5).

Ning, J. X. et al. (2012). Improvement of LEACH protocol for WSN. *9th International Conference on Fuzzy Systems and Knowledge Discovery.*

Petre-Cosmin, H. et al. (2010). Hierarchical Routing Protocol based on Evolutionary Algorithmsfor Wireless SensorNetworks. *9th RoEduNet IEEE International Conference 2010.*

Raed, M. B. H., &Abdalraheem, A. I. (2013). A Survey on LEACH-Based Energy Aware Protocolsfor Wireless Sensor Networks. *Journal of Communications, 8*(3).

Sabarish, B. A. et al. (2012). A survey on clustering protocols in Wireless Sensor Networks.International. *Journal of Advances in Computing and Information Technology, 160-168.*

Shankar, M. et al. (2012). Performance Evaluation of LEACH Protocol in Wireless Network. *International Journal of Scientific & Engineering Research, 3*(1).

Sonam, J., & Sandeep, S. (2012). Geometric Routing Protocol based on Genetic Algorithm for DelayMinimization in MANETs. *International Journal of Computer Science and Information Technologies, 3*(3), 4122-4126.

Song, F., & Zhao, B. H. (2008). Trust-based LEACH Protocol for Wireless Sensor Networks. *Second International Conference on Future Generation Communication and Networking.*

Vikram, M., & Neena, G. (2012). Performance Analysis of QoS Parameters for Wimax *Networks. International Journal of Engineering and Innovative Technology (IJEIT), 1*(5).

Wu, X. H., & Wang, S. (2010). Performance Comparison of LEACH and LEACH-C Protocols byNS2. *Ninth International Symposium on Distributed Computing and Applications to Business, Engineeringand Science.*

Yu, Y., Viktor, K., & Prasanna, B. K. (2006). Introduction to Wireless Sensor Networks. *Information Processing and Routing in wireless sensor networks, World Scientific Publishing Co., Pte.Ltd.* Retrieved from http://www.worldscibooks.com/compsci/6288.html

Yulai Suen. A. Genetic-Algorithm Based Mobile Sensor Network Deployment Algorithm. *Final Report, The University of Texas at Austin.*