Trust based Energy Efficient Clustering using Genetic Algorithm in Wireless Sensor Networks (TEECGA)

Nivedita B. Nimbalkar
Department of Computer engg,KJCOEMR, Pune

Soumitra S. Das
Department of Computer Engg,DY Patil school of Engg, Lohgaon.Pune

Sanjeev J.Wagh, Ph.D
Department of computer Engg, KJCOMER.Pune

ABSTRACT
Wireless sensor networks are gaining lot of popularity because of its widespread applications. They consist of small sensor nodes that are low in battery and computational capability. Mostly these nodes are deployed in remote areas thus it's not easy to replace their batteries. In clustering process, clusters of the sensor nodes are formed. All the sensor nodes send the sensed data to their cluster heads and cluster heads forward the data to sink. Various techniques like fuzzy logic, neural networks, artificial intelligence and genetic algorithm etc can be used for clustering and cluster head selection in wireless sensor networks. Proposed system implements genetic algorithm based cluster head selection technique. The metrics used are residual energy, distance, number of sensor nodes, number of cluster heads and trust. Proposed system also aims at ensuring successful delivery of the data and reliability by calculating trust of all the nodes. A node with low trust value will not be selected as a cluster head. In TEECGA, multihop communication between cluster heads is used i.e. every cluster head will send the data to its nearest cluster head and finally a single cluster head will send the data to sink node which results in enhanced network lifetime. From graphical and mathematical analysis, it is proved that the proposed system is more energy efficient than classical methods of clustering and is trust based.

Keywords
Genetic Algorithm (GA), Cluster Head (CH), Clustering, Wireless Sensor Network (WSN), Sink node

1. INTRODUCTION
Wireless sensor networks are used in various applications like health care, agriculture, medical etc.[1].Achieving efficiency of energy is a major issue in wireless sensor networks. TEECGA uses genetic algorithm for selection of cluster heads and formation of clusters. As clustering minimizes communication distance, it results in saving energy [3]. Darwin's theory of evolution is the motivation behind genetic algorithm. Darwin had proposed his theory for living things. Darwin suggested that the individual who is fittest will survive in the competition of existence [2]. In genetic algorithm based cluster head selection, an individual with highest fitness is selected as cluster head selection candidate.

In GA based cluster head selection, chromosome representation is used for the network. Network will be represented as bits i.e. cluster head will be represented using '1' and member sensor node using '0'. Initially random population is generated [5]. Fitness of all the nodes is calculated. Then nodes with highest fitness are selected for application of genetic operators like crossover and mutation. One point crossover is used. A single random crossover point will be generated and bits of both the parents will be swapped at that point thus generating a new individual [4,6,7]. It's assumed that the newly generated offspring will inherit only the good characteristics' of both of its parents and thus will be a better individual. After crossover, mutation operator is applied. Mutation changes a bit from '0' to '1' and vice a versa. Mutation is applied at mutation probability [8].

After application of genetic operators, a node that has highest energy gets selected as a cluster head. All the member nodes find their nearest cluster heads and join that cluster thus forming clusters. Nature of the clusters is dynamic. They will be changed in every round.

Multihop communication between cluster heads is used. Every cluster head will find its nearest cluster head and forward its data to it. Finally a single CH will forward the data to sink. As energy of a single cluster head is consumed in forwarding the data to sink, it greatly helps in saving energy.

Successful delivery of the data is ensured using parameter 'trust'. For calculation of the trust, successful and failed type events of every node are calculated and stored in the database. Trust observed by all the neighbor nodes will be added to find the 'trust' of a particular node.

2. LITERATURE SURVEY
In paper [2], Abbas Karimi has proposed a method. Fitness function calculation depends upon difference of energy of chromosomes in current and previous round. Chromosome with least difference gets selected. Proposed method is successful in increasing lifetime of the network. Other parameter considered is number of messages received in base station. As the number increases, it prolongs networks lifetime.

In paper [6] , Sudakshina Dasgupta presents a new approach of clustering in wireless sensor networks using genetic algorithm. Here a genetic algorithm is applied by base station. Basic idea is to evaluate the fitness of chromosomes. Depending upon the fitness value, next generation can be found that means chromosomes with higher fitness can be used to produce off springs for next generation. As a result of genetic algorithm, clusters are determined and the network details are broadcasted to all the nodes and clusters will be formed accordingly.

Simulation is done in Matlab. Performance is compared with LEACH. Proposed method was found more efficient than LEACH.

In paper [7], Shiuyan Jin proposed a genetic algorithm based clustering technique.Author has explained the scaling window concept. If individuals have similar fitness value then which one to select as cluster head becomes problematic. So to select

a better individual as a cluster head, they have subtracted minimum fitness value from fitness value of each individual.

i.e. $fit(i)=fit(i)-fit_{min}$. They have used 2D environment with 100 nodes. They have tried two different sink positions (0,0) and (100,100). Value of certain parameters does not change throughout like population size, selection type, crossover rate, crossover type, mutation rate and generation size.

Moslem Alfrashteh Mehr et al [4] has proposed a method to form energy efficient clusters. Proposed system also finds optimum number of cluster heads. Various parameters used are energy, distance, number of sensor nodes and number of cluster heads. Method was more energy efficient than LEACH.

3. PROPOSED SYSTEM

Proposed system aims at forming energy efficient and trust based clusters. Initially fitness of all the nodes will be calculated. After that genetic operators like crossover and mutation are applied on the chromosomes. After application of genetic operators, nodes with highest energy will be selected as cluster heads. Successful delivery of the data is ensured that's why the node that has lesser number of successful events will not be selected as a cluster head. Multihop communication between cluster heads is used to avoid transferring the data over long distance which results in increased number of rounds which in turn increases network lifetime.

3.1 Algorithmic Strategy

1. BEGIN GA
2. Select fitness parameters and generate the initial population
3. WHILE (maximum numbers of rounds are not reached)
4. Calculate the trust by using the given formula as stated by [9],

   $Ti^{A,B}=a_iS_i^{A,B}-b_iF_i^{A,B}/c_iS_i^{A,B}+d_iF_i^{A,B}$

   //$S_i^{A,B-}$Number of successful type 'i' event that A has measured for B.

   //$F_i^{A,B-}$Number of failed type 'i' event that A has measured for B.

   //a, b, c, d - Weight/Significance of successful type event vs Weight/Significance of failed type event
5. //Trust of the node for P neighbors

   $T_{=}\sum_{i=1}^{P}T_i$
6. Fitness=RE+SE+(x*(D)+(1-X)*(N-CH))+T
7. IF (Fitness of the node is highest)
8. Crossover and mutation
9. END IF
10. IF (Energy of the node is highest)

    Then Select it as a cluster head

    END IF
11. All the nodes join their nearest CH to form clusters
12. Sensor nodes send the data to CH.
13. CH will find its nearest CH and forward data to it.
14. A single CH will forward the data to sink.
15. END WHILE
16. END GA

4. EXPERIMENTAL SETUP

Proposed protocol was implemented in Java. It can be run on Windows XP/Windows Vista or on Windows 7 operating system. Database used is MySql. Apache tomcat is used as web server.Table1 shows the simulation parameters used.

Table 1 Simulation Parameters

Number of Sensor nodes	Dynamic
Network dimension	200 * 200 m^2
Initial population	Dynamic
Initial Energy	1 Joule
Crossover type	One point
Crossover rate	0.5
Mutation rate	0.1
Sink location	(650,40)
Number of generations	Dynamic

4.1 Parameters Used for Calculation of Fitness

Fitness parameters specified below are used for calculation of fitness function. Depending upon the fitness value, node will be selected as a cluster head.

Residual Energy(RE) : Residual energy is remaining energy of the node after certain number of rounds.

Sum of Energy(SE): It is energy that is required to transmit a single message from all sensor nodes to sink-(sum of energy required to transmit a single message from member sensor nodes towards the CH+(M-1)(Energy required to receive a single message from member sensor nodes+ Energy required to transmit a single message to sink))

Here, M-Number of messages

Distance (D): Sum of the distance of all the member sensor nodes towards the sink-(Sum of the distances from all member nodes towards cluster head- distances from all cluster heads towards the Sink)

Number of nodes(N): Total number of sensor nodes in a network

Number of CH(CH): Total number of cluster heads in the network

Trust (T): A value which helps in deciding which node successfully delivers the data and which node is malicious.

4.1.1 Calculation of 'SE' and 'D'

'SE' and 'D' are as stated by [4] :

$SE=\sum_{i=1}^{M}(RES_i-(REC_i+(M-1)(CHE+CES))$

Here,

RES: Energy required to send a message from all the sensor nodes to sink.

REC: Energy required to send a message from all the sensor nodes to CH.

CHE: Total Energy required by CH to receive messages from all its cluster members

CES: Energy required to send a message from CH to sink.

M- Number of messages from cluster members.

$D=\sum^{M}_{i=1}(DRS_i-(DRC_i+DCS_i))$

Here,

DRS: Sum of the distance of all the member sensor nodes towards the BS.

DRC: Sum of the distances from all member nodes towards cluster head

DCS: Sum of distances from all cluster heads towards the BS.

5. RESULT ANALYSIS

Fig 1 shows the cluster formation in TEECGA. Initially deployment of CH and sensor nodes is random. CH's are the nodes which consume more energy as compared to regular sensor nodes that's why we are trying to minimize their number. Status of the network in the last round is shown in Fig.2

Sum of residual energy of the nodes in TEECGA was compared with that of method implemented by Moslem Afrashteh Mehr et al [4] and LEACH[10]. Sum of residual energy of the TEECGA is higher as shown in the Fig.3

Number of alive nodes in TEECGA in various rounds was compared with that of method implemented by Moslem Afrashteh Mehr et al [4] and LEACH[10]. It was found that TEECGA is more efficient in terms of number of alive nodes. It is shown in Fig 4.

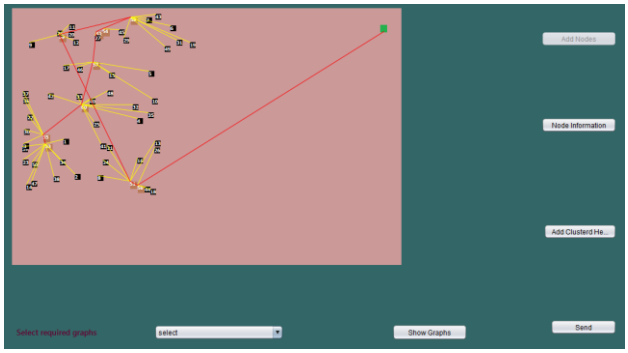

Fig.1 First Round of Cluster Formation and Data Transmission

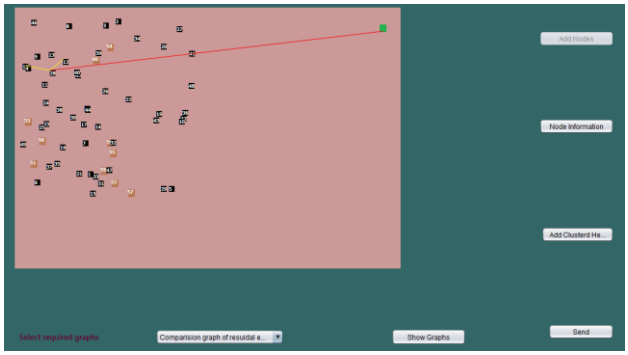

Fig.2 Last Round of Cluster Formation and Data Transmission

5.1 Lifetime of the network (using residual energy as parameter)

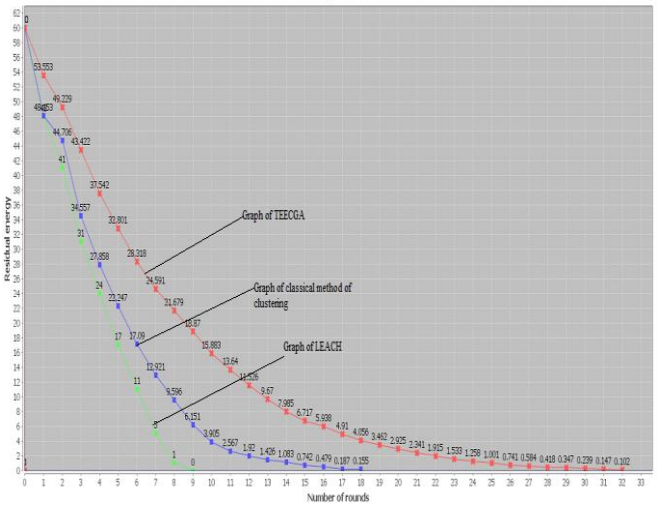

Fig.3 Number of Rounds vs Residual Energy

5.2 Lifetime of the network(using residual energy as parameter)

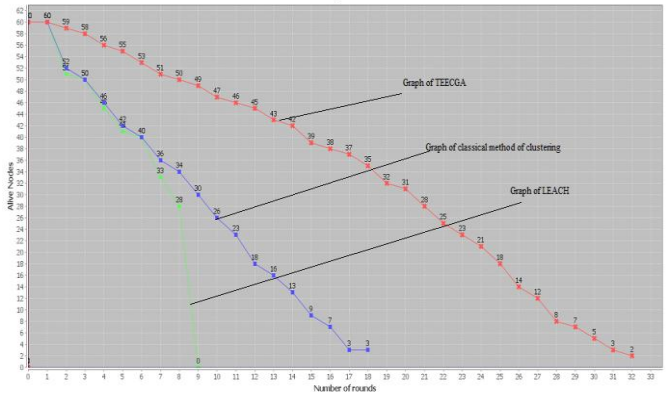

Fig.4 Number of Rounds vs Alive Nodes

6. DISCUSSION

TEECGA was designed for achieving efficiency of energy. Fittest individual gets selected as cluster head and CH's are rotated in every round. So energy of the single node is not drained thus resulting into enhanced network life time. Nodes join their nearest CH thus reducing the communication distance and saving energy. Every cluster head sends data to its nearest cluster head instead of sending it to sink thus lowering energy consumption. Trust is used as a fitness parameter which ensures successful delivery of the data.

7. CONCLUSION AND FUTURE SCOPE

TEECGA is designed to achieve efficiency of energy and trust. It is an energy efficient and trust based clustering protocol. It achieves efficiency of energy by forming energy efficient clusters using genetic algorithm. Fitness of all the nodes is calculated. Nodes with higher fitness will be selected for application of genetic operators. After that nodes with highest energy will be selected as a cluster heads. TEECGA also ensures successful delivery of the data by counting successful and failed type events of the node. The metrics used are: Residual energy, sum of energy, distance, number of sensor nodes, number of cluster heads and trust. Graphical analysis proves that the proposed system is more efficient than classical methods of clustering.

8. REFERENCES

[1] Jun Zheng, Abbas Jamalipour ,"Wireless sensor networks:a networking perspective"

[2] Abbas Karimi, S. M. Abedini, Faraneh Zarafshan, S.A.R Al- Haddad,"Cluster Head Selection Using Fuzzy Logic and Chaotic Based Genetic Algorithm in Wireless Sensor Network", J. Basic. Appl. Sci. Res., 3(4)694-703, 2013

[3] D.Srinivasa Rao, B.J.M. Ravi Kumar ,"Performance Evaluation of Genetic Based Dynamic Clustering Algorithm over LEACH Algorithm for Wireless Sensor Networks" , International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issue-4, September 2011

[4] Moslem Afrashteh Mehr, "Design and Implementation a New Energy Efficient Clustering Algorithm using Genetic Algorithm for Wireless Sensor Networks", World Academy of Science, Engineering and Technology 52 2011

[5] Sanjeev Wagh, Ramjee Prasad," Heuristic Clustering for Wireless Sensor Networks using Genetic Approach", International Journal of Wireless and Mobile Networking (IJWAMN)Vol. 1, No. 1(November 2013)

[6] Sudakshina Dasgupta, Paramartha Dutta, "An energy efficient genetic approach for clustering of wireless sensor networks"

[7] Shiyuan Jin, Ming Zhou, Annie S. Wu, "Sensor network optimization using genetic algorithm"

[8] Selim Bayrakli,Senol Zafer Erdogan , "Genetic algorithm based energy efficient clusters(GABEEC) in wireless sensor networks", ScienceDirect Computer Networks 51 (2007) 1031–1051

[9] Mobile Networks, Trust management in wireless sensor networks, Eur. Trans. Telecomms. 2010; 21:386–395. Published online 8 April 2010 in Wiley InterScience

[10] An Application-specific protocol architecture for Wireless microsensor networks," Wendi B. Heinzelman, Member, IEEE, Anantha P. Chandrakasan, Senior Member, IEEE, and Hari Balakrishnan, Member, IEEE, IEEE Transactions on Wireless Communications, VOL. 1, NO. 4, October 2002