

Energy Efficient Genetic Inspired Scheduling for Data Aggregation

Dr. S. Madhavi¹ and Dr. Tai-hoon Kim²

¹*Department of Computer Science and Engineering, P.V.P. Siddhartha Institute of Technology, Andhra Pradesh, India*

²*Department of Convergence Security, Sungshin Women's University 249-1, Dongseon-dong 3 - ga, Seoul, 136-742, Republic Korea*

Abstract

Energy efficient Data Aggregation Scheduling is essential in Ubiquitous sensor networks. Optimal communication distances between the nodes and the base station improves the network lifetime. Clustering improves the efficiency of the data aggregation problem. Many clustering techniques exist to find the optimal number of clusters in the network. The computation complexity of the methods to obtain optimal number of clusters is fair when genetic approaches are employed instead of the classical approaches. Since nodes in USN are dynamic, finding data aggregation schedules is difficult in a USN. Concurrent transmissions improve the throughput of the network, but the SINR perceived at the receiver should be greater than or equal to a certain threshold value for a successful transmission. Also the cluster heads and the member nodes dissipate energy. This power dissipation is more when there are more number of cluster heads. The noise and the power dissipation play a major role in decreasing the network life time. Hence in this paper we designed an efficient energy dissipation algorithm for data aggregation scheduling using the principles of Genetic algorithm and SINR. Since the battery power and bandwidth are the limited resources for the nodes in the USN our data aggregation scheduling algorithm gives equal priority for all of the following while electing a node as a cluster head. They are (a) Total power dissipation from all cluster heads in the USN, (b) The total power dissipation from the nodes at each cluster heads, (c) The SINR perceived at the cluster head should be more than a threshold value at each cluster head for the transmission to be successful, (d) The optimal schedule for the nodes in the USN so that all the nodes transmits their data finally to the base station quickly.

With the application of the genetic methods our algorithm proved to be efficient when compared with the existing algorithms in obtaining maximum network life time with minimum number of clusters.

Keywords: USN, SINR, GA, DDPA

1. Introduction

In [1] we proposed a DDPA dynamic distributed power adaptive protocol with optimal node degree and energy-efficient. In this paper we wanted to apply genetic approach for solving the data aggregation problem. A GA will typically have: (1) a representation of a chromosome, (2) an initial pool of chromosomes, (3) a fitness function, (4) a selection function and (5) a crossover operator and (6) a mutation operator [19]. Several algorithm like LEACH, SPAN, ASCENT and STEM were proposed earlier which are used for topology control. C. Zhang, F. Liu and N. Wu, proposed a Distributed Energy-efficient Unequal Clustering Routing Protocol for Wireless Sensor Networks in 2014. Hyunjo Lee, Miyoung Jang, and Jae-Woo Chang proposed a New Energy-Efficient Cluster-Based Routing Protocol Using a Representative Path in Wireless Sensor Networks in 2014. Sanjeev Kumar Gupta, Neeraj Jian, Poonam Sinha, proposed a Clustering Protocols in

Wireless Sensor Networks: A Survey in 2013. W. Liu and L. Wang proposed an improved algorithm based on LEACH protocol in 2013. M. Wu, proposed an Energy-Efficient Routing Protocols in Heterogeneous Wireless Sensor Networks in 2012. T. D. Singh studied on the analysis of Low Energy Adaptive Clustering Hierarchy (LEACH) protocol”, in 2011. Y. Ha, H. Kim, and Y. Byun, proposed Energy-efficient fire monitoring over cluster-based wireless sensor networks in 2012. R. Halke and V. A. Kulkarni, proposed a En-LEACH Routing Protocol for Wireless Sensor Network in 2012. J. S. Brunda, B. S. Manjunath, B. R. Savitha and P. Ullas proposed energy aware threshold based efficient clustering (EATEC) for Wireless Sensor Network in 2012. We compare our results with work done in previous protocols [18]. We assume the simple model as assumed in [18].

2. Proposed Efficient Energy Dissipation Ubiquitous Sensor Network Model

Consider a ubiquitous sensor network with n arbitrarily distributed ubiquitous nodes. Let a directed graph $G = (V, E)$ denote a ubiquitous network where $V = \{v_0, v_1 \dots v_{n-1}\}$ and $E = \{e_0, e_1, \dots e_{n-1}\}$. V_i denotes ubiquitous node i , and e_i denotes the edge between two ubiquitous nodes. The minimum Euclidean distance between a pair of nodes is 1. Each node can act as a cluster head and is adaptive by adjusting the transmission powers based on the communication distances.

A node is elected as a cluster head using the following constraints.

1. SINR rule. Let V_j denote a node that can transmit concurrently with node V_i as the head. For a successful transmission the SINR perceived at the receiver should be

greater than or equal to β i.e.
$$\frac{P_s / d_s^\alpha}{N_0 + \sum_{j=1}^{conc} P_r / d_j^\alpha} \geq \beta$$

Where P_s and P_r denote the transmission powers of sender and receiver nodes respectively. d_s is the distance between the sender and the receiver nodes and d_j is the distance between a concurrent transmitter and the receiver. α is the path loss ratio, which has a typical value between 2 and 4. N_0 is the ambient noise. β is the threshold for a successful transmission.

2. Optimal Energy dissipation rule[18]

To transmit a k -bit message of a distance d , the radio expends

$$E_{send}(k, d) = E_{send-elec}(k) + E_{send-amp}(k, d) = \begin{cases} kE_{elec} + E_0kd^2, & d < d_0 \\ kE_{elec} + E_{mp}kd^4, & d \geq d_0 \end{cases}$$

And to receive this message, the radio expends:

$$E_{recv}(k) = E_{recv-elec}(k) = kE_{elec}$$

Where

$$d_0 = \sqrt{\frac{E_0}{E_{mp}}} \text{ Denotes the threshold distance and the electronics energy}$$

Assume that there are N nodes randomly distributed in an $M \times M$ region. K is the number of bits in each data message, $d_{i- \text{to BS}}$ is the distance between the i^{th} CH node and the BS, and perfect data aggregation is assumed. From [18] we obtain

$$E_{Head}(i, k, d) = \begin{cases} kE_{elec}P_i + kE_{DA}(p_i + 1) + kE_{elec} + E_{fs}kd^2, & d_{i-toBS} < d_0 \\ kE_{elec}m_i + kE_{DA}(m_i + 1) + kE_{elec} + E_{mp}kd^4, & d_{i-toBS} > d_0 \end{cases}$$

Each member node only needs to transmit its sensed data to the i^{th} CH once during a frame. Thus, the energy dissipated in each member node is given by following equation

$$E_{mem-i}(l, k, d) = \begin{cases} kE_{elec} + E_{fs}kd^2, & \text{if } d_{i-toCH} < d_0 \\ kE_{elec} + E_{mp}kd^4, & \text{if } d_{i-toCH} > d_0 \end{cases}$$

The dissipation energy in communication process is the main parameter to be minimized. Minimum number of clusters results in energy savings hence GA is used to determine optimal number of CHs and their locations

3. Optimal Concurrent transmissions. Finding optimal schedules is difficult because the USN is dynamic in nature. Since two or more number of nodes are transmitting concurrently, all the nodes will be able to finally send data to their base station and frame length decreases. Two or more nodes can transmit at the same time using either CDMA/FDMA. If a node X is not within the current transmission range of Y then the node X cannot be the neighbor of a node Y and it is deselected for considering to increase its degree by node Y. The node adjusts its power in such a way that it would be able to communicate directly with more number of nodes perceiving the SINR constraint also.

4. Minimum frame length. The frame length decreases with the increase of the indegree of the cluster head node. And hence the throughput of the USN increases. Hence to achieve maximum throughput optimal schedules needs to be constructed which requires the power adjustments at the node to be adaptive

Generally time is slotted to intervals where L denotes the length of the interval. All non interfering nodes u_i are scheduled to send their aggregated data to any of their neighbors, v_i .

Construct a set of schedules from which an optimal schedule is founded and it includes

- a. When head nodes should transmit,
- b. How the head nodes gathers data from its neighbors.

The optimal schedule satisfies the following conditions.

- a. Each active node may be scheduled more than once but must be scheduled at least once.
- b. A node cannot act as a transmitter and a receiver in the same time slot, in order to avoid the primary interference [4].
- c. A Head node v_i transmits to the Base station only after all its neighbors have been scheduled [6].

3. Proposed Egsda Algorithm

A Genetic Algorithm is used to select the best population. Fitness determines the quality of the individual on the basis of the defined criteria.

EGSDA is executed in rounds. There are two phases like

1. Cluster election phase
2. Data transmission phase.

During cluster election phase the network is organized into a set of clusters by electing the cluster heads. Each member of the clusters gets a slot in the data aggregation scheduling for transmitting to its cluster head. An optimal schedule is found in the election phase.

During the data transmission phase the member nodes transmits to the cluster heads as agreed in the schedule. The network is reconstructed with the set of cluster heads acting as the member nodes.

These two phases are repeatedly executed until a stopping criterion is reached.

A GA optimizes the number of clusters and sensor connections for an arbitrary network. Each node in a network is either a cluster-head or a “member“of a cluster-head. Each regular node can only belong to one cluster-head. Each cluster-head collects data from all sensors within its cluster. If for any node the base station is nearest to any other node in the network then it by itself becomes the cluster head with no members and can directly transit data to the base station.

Chromosome representation and Population Generation: The USN nodes are represented as bits of a chromosome. A chromosome can be a binary string. Each chromosome is a sequence of 0s or 1s. The initial set of the population is a randomly generated set of individuals. Head and member nodes are represented as 1s and 0s, respectively.

Sensors are represented by a sensor set $S = \{s_1, s_2, s_3, s_4 \dots s_n\}$

The initial chromosome (X) is represented as follows. Let us assume there are 10 sensors in the network.

Then

Sensors set S	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇	S ₈	S ₉	S ₁₀
Chromosome set X	1	0	0	0	0	1	1	1	0	0

Fitness: In a GA, fitness is evaluated by the function defining the problem. The chances of survival are higher for better fitness values.

The fitness of a chromosome is determined by

- a. Aggregate energy dissipation from all the cluster heads and at an individual cluster head ,
- b. SINR perceived at the receiver node ,
- c. Optimal concurrent transmitters ,
- d. Minimum frame lengths

The dissipation energy in communication process is the main parameter to be minimized. It can be controlled by achieving minimum number of cluster heads. GA is used to determine optimal number of CHs and their locations by minimizing the following objective function F(X).

3) Selection: The selection function decides which chromosomes will participate in the evolution stage of the genetic algorithm made up by the crossover and mutation operators [18]. These new chromosomes join the existing population. The individuals (chromosomes) with better fitness values have better chances of selection. Roulette-Wheel selection, Rank selection and Tournament selection are a few selection algorithms. In two chromosomes are chosen at random from the population. First, for a predefined probability p, the more fit of these two is selected and with the probability (1-p) the other chromosome with less fitness is selected [18].

4) Crossover [18]: Crossover and mutation are two basic operators of GA. Performance of GA depend on them. There are several crossovers operators like single point, multipoint, uniform and arithmetic crossover operators form. The outcome of crossover heavily depends on the selection of chromosomes made from the population. In single-point crossover whereby a point is chosen at random and the two parent chromosomes exchange information after that point. An example of crossover is shown as follows

Chromosome 1 . . . 100000 | 001000 . . .
 Chromosome 2 . . . 000100 | 000001 . . .

Off-spring 1 . . . 100000 | 000001 . . .

Off-spring 2 . . . 000100 | 001000 . . .

5. Mutation. Mutation alters one or more gene values in a chromosome from its initial state. In mutation, the solution may change entirely from the previous solution. Hence GA can come to better solution by using mutation. Mutation occurs during evolution according to a user-definable mutation probability. This probability should be set low. If it is set too high, the search will turn into a primitive random search. There is different type of mutation operators like bit, flip bit, boundary, non uniform, uniform, Gaussian. It prevents the GA approach from premature convergence. It is used to search the solution from a whole new place instead of searching for the current better ones:

. . . 10001000 . . .

↓ Mutation

. . . 00010001 . . .

For comparison purposes like in [18] in this paper also, the proposed GA implements mutation on a per-bit basis. A chromosome selected for mutation will have a randomly selected bit changed from 0 to 1, or vice versa.

The effect of mutation on the two offspring created as a result of crossover.

Offspring 1 Offspring 2

Original 1100110110001110 1001001100000110

Mutated 1100110010001110 1001001100000110

During mutation, the eighth bit of offspring 1 is changed from 1 to 0; however, due to very low probability of mutation, there is no mutation in offspring 2. The cluster head should be at a shortest distance to every other node with optimal communication power required for transition in such as way that the transition is successful so that the SINR perceived at the cluster head is greater than or equal to a threshold value. And energy dissipation from all the cluster heads in the network should be optimal

Hence the four constraints for the fitness function is as flows

1. With **Min aggregate sum of P_{ij}**
2. The power is adjusted at a node in such a way that it will have optimal energy dissipation.

$$F(X) = w \left(\frac{E_{diss}}{E_{live}} \right) + (1 - w) \left(\frac{L}{N_{live}} \right)$$

$$E_{live} = \sum_{j=1}^{N_{live}} E_o(j) = N_{live} E_o$$

$$E_{diss} = \sum_{i=1}^L \left[(E_{CH}(i) + E_{CH-CP}(i)) + \sum_{l=1}^{N_{live}-L} (E_{mem-l}(l) + E_{mem-CP-l}(l)) \right]$$

To transmit a k-bit message a distance d, the radio expends:

$$E_{send}(k, d) = E_{send-elec}(k) + E_{send-amp}(k, d) = \begin{cases} kE_{elec} + E_0kd^2, & d < d_0 \\ kE_{elec} + E_{mp}kd^4, & d \geq d_0 \end{cases}$$

and to receive this message, the radio expends:

$$E_{recv}(k) = E_{recv-elec}(k) = kE_{elec} \quad - 2$$

Where $d_0 = \sqrt{E_0/E_{mp}}$ denotes the threshold distance and the electronics energy

$$E_{Head}(i, k, d) = \begin{cases} kE_{elec}p_i + kE_{DA}(p_i + 1) + kE_{elec} + E_{fs}kd_{i-toBS}^2, & d_{i-toBS} < d_0 \\ kE_{elec}m_i + kE_{DA}(m_i + 1) + kE_{elec} + E_{mp}kd_{i-toBS}^4, & d_{i-toBS} > d_0 \end{cases}$$

With

$$E_{mem-i}(l, k, d) = \begin{cases} kE_{elec} + E_{fs}kd_{i-toCH_i}^2, & \text{if } d_{i-toCH_i} < d_0 \\ kE_{elec} + E_{mp}kd_{i-toCH_i}^4, & \text{if } d_{i-toCH_i} > d_0 \end{cases}$$

3. **Optimal Indegree.** Minimum Aggregate distance to all its memebr nodes and optimal indegree

$$\frac{p_s/d_s^\alpha}{N0 + \sum_{j=1}^{conc} P_r/d_j^\alpha} \geq \beta$$

4. **And Optimal schedule** Optimal schedule is the one which maximizes the throughput and minimises the frame length.

A node which satisfies these criteris ais elected as a Cluster head. A set of all possible schedules are prepared with the cluster heads elected. The network is rebuilt where cluster heads acts as the nodes in the USN. The cluster heads are again elected. The member nodes transmit to the cluster heads. The process repeats until the network cannot be further clustered and finally the cluster heads forwards aggregated data to the base station. The algorithm can be stopped when there is no significant changes in the objective functions or we have met the maximum number of iterations.

The steps of the EGSDA protocol can be summarized as follows:

Construction of the data aggregations schedule consists of the following

1. Construction of a set of schedules which consists of
 - a. Election of Cluster heads
 - b. Finding which two cluster heads can transmit concurrently.
2. Find finding optimal schedule form them
3. Rebuild the network with cluster heads as the nodes.
4. If all nodes are transmitted then goto step 5 otherwise goto step 1.
5. Stop

Construction of Schedules .

Step 1: as in [18] the network is initialized by defining the parameters of the energy model, N, E₀, k, sink location and sensor field size, the parameters such as p_s, p_c, p_m and Maxgen , gen are initialized; then set rounds counter to zero.

Cluster heads are selected as follows

Step 3 Apply GA by setting randomly p_s initial () bits binary chromosome

Step 4 Calculate the objective function F(X) for all CHs chromosomes.

Step 5. Find data aggregation schedules. And find the optimal schedule from them.

Step 6: Increase the counter of generations by one (gen=gen+1).

Step 7: Select the best CHs chromosomes based on the fitness value ($1/F(X)$) using roulette wheel selection.

Step 8: From each pair of CHs parents, apply crossover operation based on the crossover rate p_c .

Step 9: Apply mutation to all genes of each child based on probability of mutation p_m .

Step 10: Calculate the objective function $F(X)$ for new CHs chromosomes .

Step 11: Select best p_s chromosomes from parents and children to be population pool for the next generation.

Step 12: if the stopping condition is met then go to step 13. Otherwise, return to step 6
step 13. stop

3. Results

In [18] the authors assumed 100 homogeneous nodes with initial energy of 0.5J, scattered randomly within a 100×100 m² sensor field. The BS was positioned at point (50, 300) m and the packets sent were 2000 bit plus 50 bit control packets. The GA parameters are set as $p_s= 50$, $p_m= 0.11$, $p_c=0.7$, weighting factor w is 0.95 and maximum number of generation (Maxgen) is 150.

They noticed that the first died after 364 rounds and all nodes died after 739 rounds in LEACH and in GAEEP protocol, first node died after 1017.7 rounds and all nodes died after 1175 rounds. Where as in our proposed protocol EGSDA we found that the first node died after 1184 rounds and all nodes died after 1592 rounds. The number of selected CHs in each round using LEACH protocol changes from 0 to 11 CHs, and in GAEEP protocol, number of CHs is approximately uniform and changes from 1 to 7 CHs. Where as in the proposed protocol EGSDA the number of cluster heads changes from 1 to 5.

The proposed protocol EGSDA prolongs the network lifetime than the LEACH protocol and GAEEP protocol. This due to that our protocol always considered four constraints with priority while selecting the CHs. Hence our proposed algorithm outperforms the above results than obtained from LEACH and GAEEP.

5. Conclusion

In this paper we proposed EGSDA which is an extension to our earlier protocol implemented in [1]. We used the principles of the genetic algorithm for implementing the EGSDA. The fitness function considers the parameters like total energy dissipation, dissipation at the cluster head, SINR perceived at the receiver and optimal schedules as the criterion to be satisfied by a node to be elected as a cluster head. The results proved that this new fitness function improved the network life time and satiability period. The number of cluster heads are also minimum in EGSDA when compared to the LEACH and GAEEP. Hence our protocol outperforms about 10% of the protocol proposed in [18].

References

- [1] S. Madhavi, "DDPA (Dynamic Distributed Power Adaptive) protocol with an Optimal node degree method", *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology* vol. 3, no. 2, December (2013).
- [2] M. Mosleh1, H. Noori and A. Nikde "Cluster Based and Energy Efficient Coverage Protocol for Wireless Sensor Networks", *International Journal of Smart Home*, vol. 8, no. 2, (2014), pp. 39-54.
- [3] T. Sharma, G.S Tomar, R. Gandhi, S. Taneja and K. Agrawal, "Optimized Genetic Algorithm (OGA) for Homogeneous WSNs", *International Journal of Future Generation Communication and Networking*, vol. 8, no. 4, (2015), pp. 131-140.
- [4] A. Azim and M.M. Islam, "A Relay Node Based Hybrid Low Energy Adaptive Clustering Hierarchy for Wireless Sensor Networks", *International Journal of Energy, Information and Communications*, vol. 3, no. 3, August (2012).
- [5] M. Abo-Zahhad, S.M. Ahmed, N. Sabor and S. Sasaki , "A New Energy-Efficient Adaptive Clustering Protocol Based on Genetic Algorithm for Improving the Lifetime and the Stable Period of Wireless

- Sensor Networks”, *International Journal of Energy, Information and Communications* Vol.5, Issue 3 (2014), pp. 47-72.
- [6] S.C.-H. Huang, P.-J. Wan, C.T. Vu, Y. Li and F. Yao, “Nearly constant approximation for data aggregation scheduling in wireless sensor networks”, *Proceedings of the 26th Conference on Computer Communications (INFOCOM’07)*, pp. 366-372.
- [7] M.B. Young, A. Al-zou'bi, Y. Khamaysch and W. Mardini, “Improvement on LEACH Protocol of Wireless Sensor Network (VLEACH)”, *International Journal of Digital Content Technology and Its Applications*, vol. 3, no. 2, (2009), pp. 132-136.
- [8] J.L. Liu and C.V. Ravishankar, “LEACH-GA: Genetic Algorithm-Based Energy-Efficient Adaptive Clustering Protocol for Wireless Sensor Networks”, *International Journal of Machine Learning and Computing*, vol. 1, no. 1, April (2011), pp. 79-85.
- [9] K.G. Vijayvargiya and V. Shrivastava, “An Amend Implementation on LEACH protocol based on Energy Hierarchy”, *International Journal of Current Engineering and Technology*, vol. 2, no. 4, December (2012), pp. 427-431.
- [10] R. Halke and V. A. Kulkarni, “En-LEACH Routing Protocol for Wireless Sensor Network”, *International Journal of Engineering Research and Applications (IJERA)*, vol. 2, no. 4, August (2012), pp. 2099-2102.
- [11] J.S. Brunda, B.S. Manjunath, B.R. Savitha and P. Ullas, “Energy Aware Threshold based Efficient Clustering (EATEC) for Wireless Sensor Network”, *International Journal of Computer Technology and Electronics Engineering (IJCTEE)*, vol. 2, no. 4, August (2012), pp. 25-30.
- [12] W. Liu and L. Wang, “An improved algorithm based on LEACH protocol”, *Journal of Applied Mechanics and Materials*, vol. 347, (2013), pp. 2725-2727.
- [13] K. Pawar, V. Pawar and T. Sharma, “Enhancement of LEACH Protocol Using Energy Heterogeneity Concept”, *International Journal of Emerging Trends and Technology in Computer Science (IJETTCS)*, vol. 2, no. 1, February (2013), pp. 49-56.
- [14] P. Bakaraniya and S. Mehta, “K-LEACH: An improved LEACH Protocol for Lifetime Improvement in WSN”, *International Journal of Engineering Trends and Technology (IJETT)*, vol. 4, no. 5, May (2013), pp. 1521-1526.
- [15] G. Smaragdakis, I. Matta and A. Bestavros, “SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks”, *Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA)*, (2004).
- [16] B.A. Attea and E. A. Khalil, “A New Evolutionary Based Routing Protocol for Clustered Heterogeneous Wireless Sensor Networks”, *Applied Soft Computing*, vol. 12, no. 7, July (2012), pp. 1950-1957.
- [17] C. Zhang, F. Liu and N. Wu, “A Distributed Energy-efficient Unequal Clustering Routing Protocol for Wireless Sensor Networks”, *International Journal of Computational Information Systems*, vol. 10, no. 6, March (2014), pp. 2369-2376.
- [18] M. Abo-Zahhad, S.M. Ahmed, N. Sabor and S. Sasaki, “A New Energy-Efficient Adaptive Clustering Protocol Based on Genetic Algorithm for Improving the Lifetime and the Stable Period of Wireless Sensor Networks”, *International Journal of Energy, Information and Communications*, vol. 5, no. 3, (2014).
- [19] P.R. Srivastava and T.H. Kim, “Application of Genetic Algorithm in Software Testing”, *International Journal of Software Engineering and Its Applications*, vol. 3, no. 4, October (2009).
- [20] H. Wang, J. Liu, J. Zhi and C. Fu, “The Improvement of Quantum Genetic Algorithm and Its Application on Function Optimization”, *Mathematical Problems in Engineering*, vol. 2013, (2013), Article ID 730749, p. 10.
- [21] Z.Y. Li and P.C. Li, “Clustering Algorithm of Quantum Self-Organization Network”, *Open Journal of Applied Sciences*, vol. 5, pp. 270-278.
- [22] V. Shariff, S.B. Satukumati, D. Bhattacharyya and T.H. Kim, “Recovery and Replacement Algorithm for a Wireless Sensor Network”, *International Journal of Control and Automation*, vol. 8, no. 2, (2015), pp. 383-392.