Energy Efficient Spanning Tree for Data Aggregation in Wieless Sensor Networks

Zahra Eskandari
Department of Computer Engineering, Ferdowsi University of Mashhad
e-mail: za_es73@stu-mail.um.ac.ir

Mohammad Hossien Yaghmaee
Department of Computer Engineering, Ferdowsi University of Mashhad
Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506
e-mail: hyaghmae@ferdowsi.um.ac.ir

AmirHossien Mohajerzadeh
Department of Computer Engineering, Ferdowsi University of Mashhad
e-mail: am_mo49@stu-mail.um.ac.ir

***Abstract*—Wireless Sensor Networks (WSN) consist of some nodes that have limited processing power, memory and energy source. These constraints cause the algorithms that presented in this field focus on these constraints. Data aggregation is any process in which information is gathered and expressed in a summary form. Data aggregation has been put forward as an essential paradigm for wireless routing in sensor networks. The idea is to combine the data coming from different sources, eliminating redundancy, minimizing the number of transmissions and thus saving energy. For this purpose, sensor nodes must form aggregation tree, then forward sensed data to the root of this tree. Data is aggregated in intermediate nodes and the results are sent toward the root. In this paper we propose an energy aware algorithm for construction the aggregation tree. The proposed algorithm considers both the energy and distance parameters to construct the tree. Simulation results show that the proposed algorithm has better performance in terms of energy efficiency and number of failed nodes which increases the network lifetime.**

Keywords- Wireless Senssor Networks;Data Aggregation; Energy Aware Routing;Spanning Tree

I. INTRODUCTION

Wireless sensor networks consist of some nodes that have limited processing capability, small memory and low energy source. These nodes are equipped with the sensors that sense physical phenomenon such as light, pressure, temperature and humidity from their environment. These nodes are deployed randomly and often densely in the environment. In monitoring applications, sensor nodes sense data from the environment periodically and then transmit them to a base station which is called sink node [1]. Thereby data transmission consumes node's energy based on transmission distance. In most wireless sensor networks, the energy source of the node is limited and cannot be charged. In these networks, the energy of nodes that transmit data decrease quickly, so after some rounds of transmission, the power of nodes runs out and network can't sense data from the environment. Therefore to maximize network lifetime, data transmission must be decreased, that means we must minimize number of transmitted packets.

Energy consumption for transmission has direct correlation with distance between sender and receiver [2]. In order to decrease the energy consumption, routing in these networks is often performed in multi-hop architecture that causes energy efficiency and increase network scalability [3]. In this architecture sensor nodes sense data from their environment, then transmit them to the intermediate nodes called relay nodes. These nodes transmit data to the next nodes in the path toward the sink node. In this architecture intermediate nodes must transmit not only the received data but also their local sensed data toward the sink.

During data transmission surly collision occurs. Therefore nodes must be scheduled properly. Using scheduling mechanisms, we can solve the collision problem of data transmission in wireless sensor network, but this solution causes more delay in the network. Furthermore, transmission of this volume of data decrease relay node's energy quickly and causes energy of these nodes to be finished.

Data aggregation is a good solution for these problems. In this case, intermediate nodes aggregate data, remove redundancy in data and then transmit result to the next node in the path toward sink. As shown in figure 1, because nodes are deployed densely in the environment, the regions that the nodes sense, overlap one another and surly redundant data will be generated [4].

Data aggregation is performed during routing in wireless sensor networks. Sensor nodes transmit data to the next node in the path and intermediate nodes aggregate data and then transmit the result toward the sink. Finding the route from several nodes to the sink in a way that maximizes the shared path and redundancy removing is one of the main objectives in these routing protocols [4].

978-1-4244-2390-3/08/$25.00 ©2008 IEEE

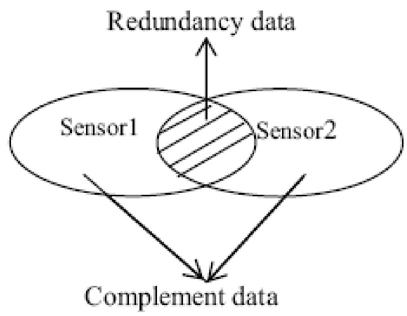

Figure 1. Sensors region overlapping

Transmission of aggregated data instead of raw data causes intermediate nodes consume less energy. Furthermore it decreases collisions and delays for retransmission. Note that data aggregation itself consumes energy [5]. Furthermore, intermediate nodes must wait to receive data from all their neighbors and then do aggregation. This waiting causes delay, but the consumed energy and delay are negligible. As described in [6] the energy consumption for data processing is less than the necessary energy for data transmission. Also by considering the collision, retransmission delay and the needed energy for retransmitting this data volume is certainly more than retransmission delay and consumed energy to transmit the aggregation result [7]. In routing with aggregation, we must construct aggregation spanning tree [8]. The spanning tree is a tree contains all network nodes and doesn't have any loop.

When an event is sensed by sensor nodes, relevant data must be forwarded to sink. The sink is the root of spanning tree and all the other nodes that sense event, construct the tree. Each intermediate node aggregates this data with data sent by its child and then transmits this aggregated data to its parent. This procedure continues until data arrives to the sink.

To decrease energy consumption, all nodes are usually initially in sleep mode [9]. When an event is detected, nodes that have sensed event, construct this aggregation spanning tree and forward data to the sink based on this tree. To evaluate the aggregation tree, in the intermediate nodes the consumed energy for sending data from the leaves to the sink must be considered. Furthermore the tree's delay which is equal to the tree's depth should be considered too. In addition to these parameters, scheduling mechanism and queuing delay can be considered as aggregation tree evaluation parameters.

In this paper, to construct aggregation tree we present a new energy efficient algorithm. To decrease number of failed nodes and to increase the network lifetime, we consider both remaining energy and distance parameters. Each node selects a node that has the most energy within neighbors as its parent. Furthermore, the distance from this parent to the root must be reasonable.

The remainder of the paper is organized as follows: in section 2, we review some existing algorithms for aggregation tree construction. The proposed algorithm is explained in section 3. In section 4 using computer simulation we evaluate the performance of the proposed algorithm under different scenarios. Finally, section 5 concludes the paper.

II. RELATED WORKS

During past few years, different aggregation algorithms have been proposed for wireless sensor networks. In [10], finding the optimal aggregation tree witch is an NP-hard problem, has been considered. Data-centric routing has been modeled and its performance has been compared with traditional end-to-end routing schemes. The authors of [10] examined the impact of source-destination placement and communication network density on the energy costs and delay associated with data aggregation. In [11] by considering the location and remaining energy of all sensors and sink, a set of trees for collecting data from all the sensors and sending them to the sink , has been constructed. Results show that the lifetime is maximized. In [12] the sink saves global network state and then by considering link cost, in centralized form, constructs minimal cost tree. In clustering algorithm given in [13] after partition network into clusters with almost 20 or 40 nodes inside each cluster, cluster's members construct aggregation tree and transmit data to cluster head. After aggregation, cluster heads transmit aggregated data to the sink in one hop or multihop manner [14]. DCTC [15], dynamically constructs an aggregation tree for mobile target tracking, dependant on target location, subset of nodes participate in tree. Espan [8] is an energy-aware spanning tree algorithm. It is a distributed protocol and facilitates the sources within an event region to perform data aggregation. In Espan, the source node which has the highest residual energy is chosen as the root. Other source nodes choose their corresponding parent node among their neighbors based on the residual energy and distance to the root. One of the most important problems of this algorithm is that nodes with least distance to root maybe selected as parent by many nodes. So these nodes consume their energy quickly and then they will be failed. In LPT [16] after selection node with most energy as root, each node selects neighbors with most energy as parent.

III. PROPOSED ALGORITHM

As mentioned earlier, to construct aggregation tree, the Espan protocol first selects a node with highest remaining energy within all network nodes as tree's root. Then each node selects closest neighbors to root as its parent. If there are multiple neighbors with equal distance, the node which has most remaining energy is selected as parent. As Espan protocol considers distance as main parameter and remaining energy as second, the network coverage is not high; because in some cases the nodes with low remaining energy are selected as parent. After local aggregation and data transmission, the remaining energy of these nodes is finished quickly. This causes the node failure and network can not coverage region completely.

LPT aims to prolong the lifetime of the sources which transmit data reports periodically. In LPT, nodes which have higher residual energy are chosen as the aggregating parents. These parents may have higher distance to root and cause more energy consumption. LPT does not consider the distance parameter in parent selection.

Unlike the Espan algorithm, the proposed algorithm considers the remaining energy in each node as the main parameter. It also considers the distance as second parameter. In the proposed algorithm, each node selects a node with most energy within its neighbors as parent. If there are some neighbors with equal energy, a neighbor with least distance will be selected.

By using this strategy, a node with low remaining energy can be alive more than that of Espan protocol. This increases the lifetime of the network and supports better coverage. To provide fairness in energy consumption, in addition to the reminding energy and distance, we consider the third parameter which is the maximum number of children. In the proposed algorithm, each node could have a predetermined maximum number of children. By using this parameter, we can provide fairness in the energy consumption between all network nodes. We determine this parameter based on nodes density and network's energy average.

In fact the proposed algorithm might select a node with higher energy but farther from root as its parent, while Espan selects a node that is closer to root and may has lower remaining energy. Furthermore, the second major difference between the proposed algorithm and Espan is that in the proposed algorithm to provide fairness each node has a predefined number of children. The proposed algorithm is a distributed algorithm which doesn't need to save global information about entire network. Each node maintains information about its neighbors in its routing table. This makes the proposed algorithm more scalable. Furthermore, in the proposed algorithm routing is done in a multihop manner.

As the transmission power is directly related to the distance, to prevent high power consumption, the proposed algorithm uses the average path's energy as a new parameter. This parameter is calculated as the sum of residual energy of each node among the path divide to the path length. When a node wants to select a new parent, a node with the highest path's energy is chosen.

Unlike the LPT algorithm, the proposed algorithm prevents of selecting a parent with high remaining energy, but far distance to the root. We use the average of path's energy to balance energy and distance parameter while the algorithms given in [8, 16] use only one of them as main parameter and the other in the next priority.

An example which helps us to understand the details of the proposed algorithm is given in figure 2. In this figure, a wireless sensor network with 8 different sensor nodes is shown. The remaining energy of nodes 1, 2, 3, 4, 5, 6, 7 and 8 are equal to 10J, 2J, 8J, 3J, 6J, 8J, 7J and 9J, respectively. Suppose that node 8 wants to select its parent. Using Espan algorithm, node 4 will be selected as parent of node 8 which has minimum distance to root, while in the proposed algorithm node 5 which has more average path's energy is selected as the parent of node 8. The average path's energy of node 8 in the proposed and Espan algorithms are equal to 7.5 and 6, respectively. In figure 2, for Espan, LPT and proposed algorithms, the spanning trees are shown. As shown in figure 2(c), LPT's tree has longer path which causes more energy consumption.

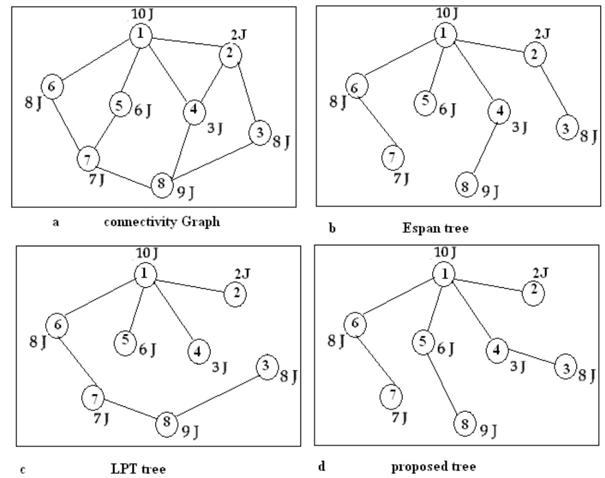

Figure 2. The spanning trees of different algorithms a) connectivity graph, b) Espan's tree, c) LPT's tree, d) proposed algorithm's tree

The details of the proposed algorithm are as below:

Define:

S_r: A message which is sent by root and contains only one field named root's ID.

S_n: A message which is sent by each node *n* and contains three fields: node's energy, node's distance_to_root, path_energy and node's ID.

Initialization:

For all nodes *n* in network set:

n.parentID=*n*

n.parent's_distance_to_root=∞

n.parent_energy=0

Calculate max_child for each node base on density and network energy

n.children_number=0

Protocol:

1. A node with highest remaining energy is selected as the root node.
2. The root node broadcasts the S_r message to all its neighbors. This message conations the root's ID.
3. Each root's neighbors *n* which receives this message, select the root node as its parent and set:
    a. n.parentID=root's ID
    b. n.diatance to root=1
    c. if (n.children_number < n.max_child)
        i. node *n* calculates path_enrgy and broadcasts the S_n message to all its neighbors

4. When each node k (except the root) receives S_n message, the following operations are done:

    If(sender's Path energy average > Parent's path energy average)

    THEN

    Node k selects the sender as its new parent

    Else

    Node k doesn't change its parent.

As mentioned above, the proposed algorithm selects a node with the most remaining energy as the root. After that other nodes which are root's neighbors and have one hop distance to the root, select this node as their parent and start to broadcast their information. Nodes that receive these messages, select their parents as described in the above algorithm. This continues until all alive nodes connect to the tree.

IV. SIMULATION RESULTS

In this section, using computer simulation, we evaluate the performance of the proposed algorithm and compare it with Espan algorithm.

A. Simulation Model

The simulation network environment is considered as a 25m*25m square area. The number of nodes (N) is varied between 50 and 500. The initial energy of each node is a random variable between 800J and 1000J. The communication range of all nodes is set to 4 meter. We assume that all nodes in network can sense occurred events and participate in aggregation process.

Suppose at each node i the variables ER_i and EA_i represent the required energy for data receiving and data aggregation, respectively. Furthermore, suppose $dist_{ij}$ and ET_{ij} are the distance and required transmission energy between node i,j, respectively. Using the energy model given in [2], the following equations are obtained:

$$ER_i = C_1 * K \tag{1}$$

$$ET_{ij} = C_2 * K * dist_{ij}^2 \tag{2}$$

$$EA_i = C_3 * K \tag{3}$$

Where C_1, C_2 and C_3 are three fixed variables which are determined based on energy and aggregation models in [17]. We assume that the transmission power is directly related to the squared distance, this hold for free space, but the exponent heavily depends on the communication medium [18].

As described in [8] if aggregation function be simple, the energy consumption for data aggregation is negligible. We assume that the input data length is equal to the output data length. In the above functions, K represents the length of transmitted and received data packet.

In this paper, a run time round is defined as the collection of one data unit from every node in the network and delivering the resulting aggregated data to the root. Network lifetime is defined as the longevity of network in terms of number of rounds performed before the failure of 5% of the total nodes of the tree [18].

B. Performance Evaluation

In this subsection, using simulation results, we evaluate the performance of the proposed algorithm. We produce some events in the simulation network. When a node detects an event, forwards sensed data to the root node using aggregation tree.

All nodes that sense an event and forward their data to the root, consumes energy for only data transmission (ET_{ij}). Intermediate nodes consume energy for both data receiving and data transmission (ER_i and ET_{ij}, respectively). Furthermore, they consume energy for data aggregation (EA_i).

At the first simulation trial, we set the value of N to 400. After each simulation round, we measure the number of alive nodes. Figure 3, shows the results of the first trial. As the proposed algorithm, selects the nodes with high remaining energy, so the node with low energy remains longer time in the network, therefore the number of alive nodes are more than that of Espan and LPT algorithms.

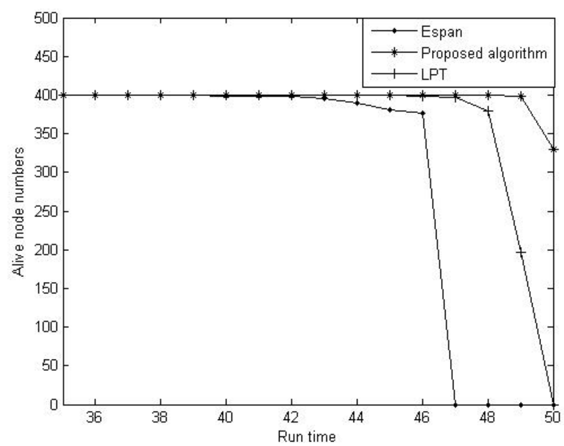

Figure 3. Number of alive nodes at N=400

As mentioned before, energy efficiency is a main objective of routing and aggregation algorithms in wireless sensor networks. In figure 4, the average hop count is plotted versus number of nodes. These values represent average depth of tree, so, the deeper tree means data transmit via longer path, and this causes more delay and also more energy consumption. As shown in figure 4, unlike the proposed algorithm the LPT transmits data via longer path which causes that energy consumption of the proposed algorithm be less than that of LPT algorithm.

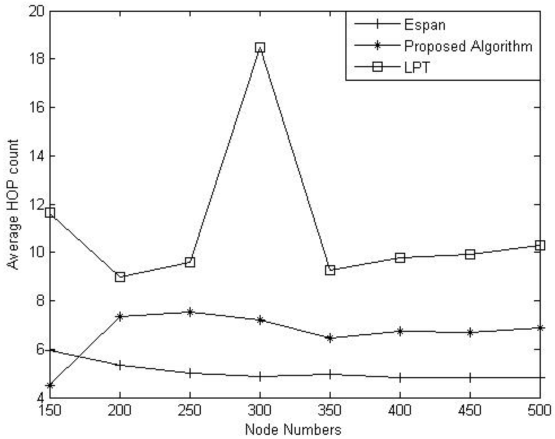

Figure 4. Average hop count to root

In figure 5, for three algorithms the average lifetime is plotted versus number of nodes. These results were obtained after 20 different simulation trials. As the proposed algorithm consumes less energy and has more alive nodes, so its lifetime is higher than Espan algorithm and LPT algorithm.

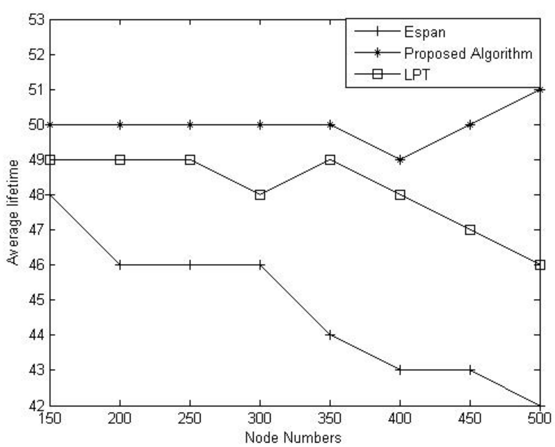

Figure 5. Average lifetime comparison

V. CONCLUSION

One of the most important constraints in wireless sensor networks is the energy consumption. To decrease the power consumption; we can decrease the number of transmitted packets. Data aggregation is any process in which information is gathered and expressed in a summary form. Data aggregation has been put forward as an essential paradigm for wireless routing in sensor networks. The idea is to combine the data coming from different sources, eliminating redundancy, minimizing the number of transmissions and thus saving energy. Data aggregation is an effective mechanism to save energy in wireless sensor networks. In this paper we propose an energy efficient algorithm to construct aggregation tree. The proposed algorithm considers both energy and distance to construct aggregation tree. Simulation results show that the proposed algorithm has better performance than existing algorithms. Simulation results confirm that the proposed algorithm decreases number of failed nodes and provides higher network lifetime.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci,"Wireless Sensor Networks: A Survey, Computer Networks", Computer Networks Journal, 2002

[2] Stephanie Lindsey, Cauligi S. Raghavendra, "PEGASIS: Power-Efficient GAthering in Sensor Information Systems", IEEE 2002

[3] Jin zhu, symeon papavassiliou, stella kafetzoglou,jie yang, "An Efficient QoS-Constrained Data Aggregation and Processing Approach in Distributed Wireless Sensor Networks", IEEE 2006

[4] Bing Liang and Qun Liu, "A Data Fusion Approach for Power Saving in Wireless Sensor Networks", IEEE 2006

[5] Hong Luo, Jun Luo, Yonghe Liu, and Sajal K. Das, "Adaptive Data Fusion for Energy Efficient Routing in Wireless Sensor Networks", IEEE 2006

[6] emanuele toscano, orazio Mirabella, iucia lo bello, An energy-efficient real-time communication framework for wireless sensor networks,

[7] yi hu, nuo yu, xiaohua jia, energy efficient real time data aggregation in wireless sensor network, IWCMC 2006

[8] Marc Lee and Vincent W.S. Wong," An Energy-Aware Spanning Tree Algorithm for data aggregation in wireless sensor networks", IEEE 2005

[9] Fei Hu,Xiaojun Cao,Carter May," Optimized Scheduling for Data Aggregation in Wireless Sensor Networks", ITCC 2005

[10] B. Krishnamachari, D. Estrin, and S. Wicker. "The Impact of Data Aggregation in Wireless Sensor Networks", *International Workshop on Distributed Event-Based Systems,* 2002.

[11] K. Kalpakis, K. Dasgupta, and P. Namjoshi, "Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks", The International Journal of Computer and Telecommunications Networking, 2003.

[12] S. Upadhyayula, V. Annamalai, and S. K. S. Gupta, "A lowlatency and energy-efficient algorithm for convergecast," EEE GLOBECOM 2003.

[13] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient,distributed clustering approach for ad hoc sensor networks,"IEEE 2004

[14] Yuanzhu Peter Chen, Arthur L. Liestman and Jiangchuan Liu," A Hierarchical Energy-Efficient Framework for Data Aggregation in Wireless Sensor Networks ",IEEE 2006

[15] W. Zhang and G. Cao, "DCTC: Dynamic convoy tree-based collaboration for target tracking in sensor networks," IEEE 2004

[16] Marc Lee and Vincent W.S. Wong," LPT for Data Aggregation in Wireless Sensor Networks", IEEE GLOBECOM 2005

[17] Xiaowei zhang, xiaohai yu, xiangqun chen, "Inter-Query Data Aggregation in Wireless Sensor Networks", IEEE 2oo5

[18] S. Upadhyayula, S.K.S. Gupta, Spanning tree based algorithms for low latency and energy efficient data aggregation enhanced convergecast (DAC) in wireless sensor networks, 2006 Elsevier