Scientific
Research

Energy-Efficient Data Collection in Wireless Sensor Networks

Mohammad Hossein Anisi, Abdul Hanan Abdullah, Shukor Abd Razak

Department of Computer Systems and Communications, Universiti Teknologi Malaysia, Skudai, Malaysia
E-mail: anisi@live.utm.my

Abstract

Wireless Sensor Networks (WSNs) are usually self-organized wireless ad hoc networks comprising of a large number of resource constrained sensor nodes. One of the most important tasks of these sensor nodes is systematic collection of data and transmits gathered data to a distant base station (BS). Hence network lifetime becomes an important parameter for efficient design of data gathering schemes for sensor networks. In this paper, we benefit both cluster and tree structures for data gathering. In our proposed energy-efficient mechanism, the most appropriates hops for data forwarding will be selected and the lifetime of the whole network will be maximized. The simulation results show that by using the proposed approach, the lifetime and the throughput of the network will be increased.

1. Introduction

Wireless Sensor Networks (WSN) consists of several sensor enable nodes which are distributed in an environment and use batteries as energy resource. These tiny sensor nodes, which consist of sensing, data processing, and communicating components, result in the idea of sensor networks based on collaborative effort of a large number of nodes. Such sensor nodes could be deployed in home, military, science, and industry applications such as transportation, health care, disaster recovery, warfare, security, industrial and building automation, and even space exploration. Among a large variety of applications, phenomena monitoring is one of the key areas in wireless sensor networks and in such networks, you can query the physical quantities of the environment [1-3].

In fact, a typical wireless sensor network is composed of a large number of sensor nodes, which are randomly dispersed over the interested area, picking up the signals by all kinds of sensors and the data acquiring unit, processing and transmitting them to a node which is called sink node. The sink node requests sensory information by sending a query throughout the sensor field. This query is received at sensor nodes (or sources). When the node finds data matching the query, the data (or response) is routed back to the sink. For example, if the sensors nodes be in a tree like structure, the base

station roles as the root of the tree and each node will have a parent [4,5]. Therefore, the data items can be transmitted hop by hop from the leaf nodes to the root.

In WSNs, data compression refers to the use of compression techniques to reduce the amount of bytes required to code the different pieces of information and, thus, the traffic load which needs to be processed within the network.

As the sensor nodes are small and battery enable devices, they have limited energy which should be used precisely. Thus, the scarce sensor resources (in particular, the battery power) are easily over consumed. Thus, the key challenge in such phenomena monitoring is conserving the sensor energy, so as to maximize their lifetime. Most of the approaches tried to response to this challenge and this will be continue to gain a better solution.

In this paper, we propose an energy-efficient in-Network data aggregation approach in WSN. The proposed approach uses the advantages of both cluster based and tree based approaches. In this approach, the whole network consists of some clusters with the same size. Each node is related to a routing sub tree and each sub tree overwhelms a cluster and the root node of each sub tree is the head node of the related cluster. The energy consumption in wireless transmissions is equal to the square of distance between two nodes in communication. In the proposed approach, all the nodes transmit their

data to their neighbor instead of their cluster head. Therefore, the communication distance is reduced and the energy consumption of each node, each cluster and the average energy consumption of the whole networks is reduced and the network lifetime is increased. Also, in the proposed approach, the most appropriate parent according to some benchmarks will be selected for each node which can balance the network load.

2. Related Works

There are several approaches which use tree structure for collecting and aggregating data. The presented approach in [6], with combining Clustering and Directed Diffusion Protocol [7], could process, collect, and aggregate data of sensor nodes without any dependency to the related environment. This paper, with presenting a dynamic clustering structure, could enable the nodes to join to the nearest head cluster while sending data to the gateway node.

Most of data gathering algorithms focus network lifetime and saving energy [4,8-11].

In the TAG (Tiny Aggregation) approach [4], each epoch divides to some time slots and these time slots specify to different levels of routing tree in reversal form. In this manner, each node depends on its situation in the tree, and in its related time slot will send its data. The node synchronization of this approach for sending and receiving data could effectively reduce the average energy consumption.

In Directed Diffusion Approach [12,13] receivers and resources using some attributes for recognizing the produced or required information and the goal of this approach is finding an efficient multi way route between senders and receivers. In this approach, each task is represented as an interest and each interest is a set of attribute-value pairs.

The LEACH (Low-Energy Adaptive Clustering Hierarchy) protocol [14] uses a random approach for distributing energy consumption among the nodes. In this approach, the nodes organize themselves as local clusters and one node roles as a local base station or a cluster head. If the cluster heads can be selected base on a priority permanently and they also can be permanent in the whole life time of system, it is obvious that the bad luck nodes which are selected as the cluster heads will be died soon and the life of all the nodes in their cluster will be finished. Thus, LEACH chooses the cluster head among the nodes which have enough energy randomly. This can prevent the discharging of the battery of a special node. In addition, LEACH uses local data fusion for compressing the data which should be sent from

cluster heads to the base station.

FTEP [15] is a dynamic and distributed CH election algorithm based upon two level clustering schemes. If energy level of current CH falls below a threshold value or any CH fails to communicate with cluster members then election process is started which is based on residual energy of sensor nodes.

In EEMC (An Energy Efficient Multi Level Clustering) [16], CHs at each level are elected on the basis of probability function which takes into consideration the residual energy as well as distance factor very efficiently. In this scheme whole information is sent and received by sink node for cluster formation.

Steiner Points Grid Routing was proposed by, Chiu-Kuo Liang, *et al.* [17] In order to reduce the total energy consumption for data transmission between the source node and the sink node, a different virtual grid structure instead of virtual grid in GGR is constructed. The idea is to construct the virtual grid structure based on the square Steiner trees [18].

The paper in [19] presents a new version of LEACH protocol called VLEACH which aims to reduce energy consumption within the wireless network. In this approach, by selecting a Vice-CH, cluster nodes data will always reach the BS; no need to elect a new CH each time the CH dies. This will extend the overall network life time.

3. Data Compression with Conditional Entropy

The goal of data compression is removing the data over-head of sensors and reducing the correlations of data for achieving beneficial information for the base station. Data source of each sensor node is illustrated with a discrete random variable. The entropy of discrete random variable of X is illustrated with $H(X)$ that is equal to minimum number of bits which are required for coding X without losing any information [20]. The common entropy of two random variables of X and Y is equal to the minimum number of bits which are required for coding X and Y together. If X includes some information from Y, then the common entropy is equal to $H(X|Y)$ and $H(X|Y) \le H(X)$ which ensure a considerable reduction in the volume of coded data. In The proposed approach, the conditional entropy is used for coding the data of sensors.

For modeling the data correlations we have used the model in [21]. In fact, the data correlation of the nodes is a function of distance. Thus, with a network of N nodes $(X_1, X_2, \cdots, X_N)$ in which the data production of each node is equal to $H(X_i) = H_1\{i = 1, 2, \cdots, N\}$, we have:

$$H\left(X_j \middle| X_i\right) = \left(1 - \frac{1}{\frac{d}{c}+1}\right) H_1 \quad \{i = 1, 2, \cdots, N\}$$

In this equation c is a constant which presents the amount of data correlation and d is the distance between nodes X_j and X_i.

4. Proposed Approach

We assume that the whole network is divided in to several clusters; each cluster has a cluster-head (CH). The clustering and the selection of cluster-head (CH) can be done by using any existing protocol like LEACH, or the optimized versions of LEACH such as [19] and [22].

The proposed algorithm works in each cluster independently and performs in two phases.

4.1. Information Packet Flow

In this phase, the cluster head transmit the information packet to its neighbors. The information packets include the information below:

- **Node's location:** Each node should now it location in prior
- **Current Energy:** Remaining energy of a node
- **Hop count:** Number of hops from cluster head
- **Data Label:** Data value which is sensed by a node

When a node receives the information packet, it considers the sender as one of its possible parents and stores its information. Then, it updates the node location, current energy and data label fields of the packets with its own, increments the hop count and transmit the packet to its neighbors. This process will be done until all the nodes in the cluster receive the information packet.

4.2. Tree Construction and Data Forwarding

When the entire nodes received the information packet, each node selects it parent which should sent its data to it. This selection will be done based on the filters below:

First, among the possible parent, the one which has the least hop distance from the cluster head (Closest node to cluster head) will be selected.

If there is more than one node with the least hop distance, the nodes which have the most current energy will be selected as the parent.

If there is more than one node with the least hop distance and the most current energy, the node which has the least data correlation will be selected as the parent.

All the above conditions lead to the best parent selec-

tion. Filter 1, selects the shortest path from a node to cluster head. Filter 2, increases the network lifetime by participating most durable nodes. Filter 3, reduce the networks overhead by checking the data correlation of the nodes.

4.3. Energy Model

Our energy model is like the energy model in [23]. In this model energy consumption for transmitting K bit is equal to:

$$E_{TX}\left(K, d\right) = E_{elec} \times K + \varepsilon_{amp} \times K \times d^2$$

And the energy for receiving K bit is equal to:

$$E_{RX}\left(K\right) = E_{elec} \times K$$

In these equations, d is a constant value which relates to the distance between two nodes and the parameters below are the constant values which are defined previously and they are equal to:

$$\varepsilon_{amp} = 100 \text{ pJ/bit/m}^2 \quad E_{elec} = 50 \text{ nJ/bit}$$

5. Performance Evaluation

The proposed approach is simulated and evaluated with J-Sim (Java-Based simulator) [14]. J-SIM is simulation software selected to implement the model. It was chosen because it is component-based, a feature that enables users to modify or improve it. J-Sim uses the concept of components instead of the concept of having an object for each individual node. J-Sim uses three top level components: the target node which produces stimuli, the sensor node that reacts to the stimuli, and the sink node which is the ultimate destination. For stimuli reporting, each component is broken into parts and modeled differently within the simulator; this eases the use of different protocols in different simulation runs. In our simulation analysis, sensor nodes are randomly distributed in a 160 m × 160 m area. The radio range of each node is 40 m and the default parameters for radio communication model of J-sim are used. The cluster-head is formed by the sink. Source node randomly sends packages with constant bit rate (CBR) to the sink. Packet size is 64 bytes and package rate is 5 pkt/s.

We have compared our approach with LEACH as an innovative Energy-Efficient clustering approach and VLEACH as a modern Energy-Efficient clustering approach. As it has mentioned before, our idea is not related to clustering and the selection of cluster-head (CH) and they can be done by using any existing protocol like LEACH, or the optimized versions of LEACH. Therefore, For Clustering, we have used the

mechanism of VLEACH in our simulation which is more energy-efficient in comparison to LEACH.

According to **Figure 1**, the total residual energy of the nodes will be decreased, gradually. But Comparing to other approaches, the proposed approach, because of using the mentioned technique, can remain more energy.

Figure 2 illustrates the throughputs of the mentioned approaches. Throughput of a node is defined as the average rate of successful message delivery over a communication channel. Thus, we can observe that has the highest throughput among LEACH and VLEACH.

In our proposed approach, increasing the data correlation will increase the network lifetime. In **Figure 3**, the horizontal pivot shows the Data Correlation Parameter (c) which is explained in Section 3 and the vertical pivot demonstrates the number of epochs until the dead of the first node. c has direct relation to data correlation. Larger numbers of c lead to more data correlation which increases lifetime, too.

6. Conclusions

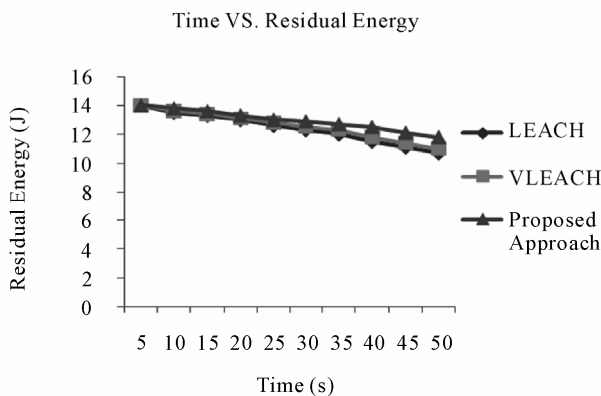In this paper, we have proposed an energy-efficient data

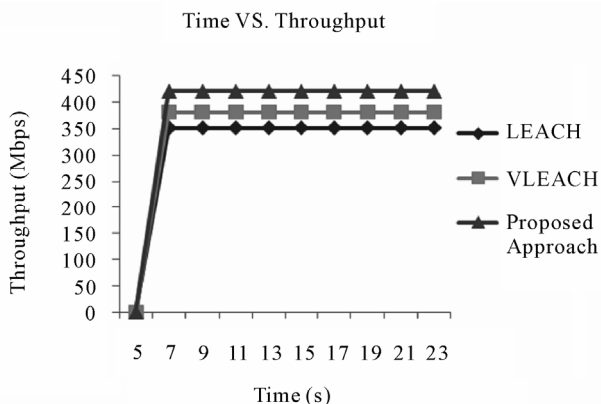**Figure 1. The reaming energy of the nodes after passing time.**

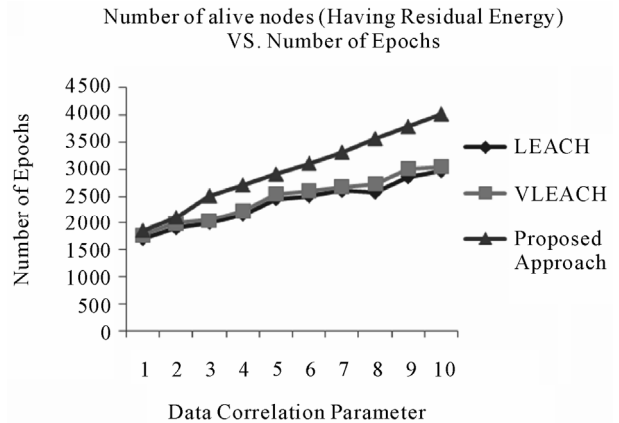**Figure 2. Different Throughput of the approaches.**

Figure 3. Increasing the network lifetime by increasing the data Correlation in the proposed approach.

collection approach in wireless sensor networks which uses an efficient strategy to forward data toward the best route. In our algorithm, there are three factors which enable the nodes to choose an appropriate parent in term of energy. These factors are distance, residual energy and data correlation. With the suggested mechanism, the remaining energy of the nodes will be increased and the life time of the whole network will be increased, too.

7. References

[1] Y. Yao and J. Gehrke, "Query Processing for Sensor Networks," *Proceedings of the First Biennial Conference on Innovative Data Systems Research*, Asilomar, 5-8 January 2003, pp. 21-32.

[2] H. Dubois-Ferrière, D. Estrin and M. Vetterli, "Packet Combining in Sensor Networks," *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems*, San Diego, 2-4 November, 2005, pp. 102-115.

[3] J. Gehrke and S. Madden, "Query Processing in Sensor Networks," *IEEE Pervasive Computing*, Vol. 3, No. 1, 2004, pp. 46-55. doi:10.1109/MPRV.2004.1269131

[4] S. Madden, M. J. Franklin, J. M. Hellerstein and W. Hong, "TAG: A Tiny AGregation Service for Ad-Hoc Sensor Networks," *Proceedings of the 5th Symposium on Operating Systems Design and Implementation*, Boston, 9-11 December 2002, pp. 131-146. doi:10.1145/1060289.1060303

[5] S. Madden, M. J. Frinklin, J. M. Hellerstein, W. Hong and D. B. Tiny, "An Acquisitional Query Processing System for Sensor Networks," *ACM Transactions on Database Systems*, New York, Vol. 30, No. 1, 2005, pp. 122-173. doi:10.1145/1061318.1061322

[6] S. Chatterjea and P. Havinga, "A Dynamic Data Aggregation Scheme for Wireless Sensor Networks," *Proceeding Program for Research on Integrated Systems and Circuits*, Veldhoven, 26-27 November 2003.

[7] Ch. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Para-

digm for Sensor Networks," *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, Boston, August, 2000, pp. 56-67.

[8] A. Chakraborty, K. Chakraborty, S. K. Mitra and M. K. Naskar, "An Energy Efficient Scheme for Data Gathering in Wireless Sensor Networks Using Particle Swarm Optimization," *Journal of Applied Computer Science & Mathematics*, Vol. 6, No. 3. 2009, pp. 9-13.

[9] Y. Wu, L. X. Yang, Y. Liu and W. Lou, "Energy-Efficient Wake-Up Scheduling for Data Collection and Aggregation," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 21, No. 2, 2010, pp. 275-287. doi:10.1109/TPDS.2009.45

[10] Z. G. Sun, Z. W. Zheng, S. H. Chen and S. J. Xu, "An Energy-Effective Clusteringalgorithm for Multilevel Energy Heterogeneous Wireless Sensor Networks," 2*nd International Conference on Advanced Computer Control* (*ICACC*), Shenyang, March 2010, Vol. 3, pp. 168-172.

[11] A. M. Brisha, "Classifying Sensors Depending on their IDs to Reduce Power Consumption in Wireless Sensor Networks," *International Journal of Online Engineering*, Vol. 6, No. 2, 2010, pp. 41-45.

[12] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A Scalableand Robust Communication Paradigm for Sensor Networks," *Proceedings of Sixth Annual International Conference on Mobile Computing and Networking* (*MobiCom* 2000), Boston, August 2000, pp. 56-67.

[13] J. Heidemann and F. Silva, C. Intanagonwiwat, R. Govindan, D. Estrin and D. Ganesan, "Building Efficient Wireless Sensor Networks with Low-Level Naming," *Proceedings of the Eighteenth ACM Symposium on Operating Systems Principles*, Vol. 35, No. 5, 2001, pp. 146-159. doi:10.1145/502034.502049

[14] C. Schurgers and M. B. Srivastava, "Energy Efficient Routing in Wireless Sensor Networks," *Proceeding of IEEE Military Communications Conference*, Vol. 1, 2001, pp. 357-361.

[15] U. B. Mahadevaswamy and M. N. Shanmukhaswamy, "An Energy Efficient Reliable Multipath Routing Protocol for Data Gathering in Wireless Sensor Networks,"

International Journal of Computer Science and Information Security, Vol. 8, No. 2, 2010, pp. 59-64.

[16] S. Hwang, G. J. Jin, C. Shin and B. Kim, "Energy-Aware Data Gathering in Wireless Sensor Networks," 6*th IEEE conference on Consumer Communications and Networking*, *CCNC*, Las Vegas, 10-13 January 2009, pp. 1-4.

[17] C. K. Liang, J. D. Lin and C. S. Li, "Steiner Points Routing Protocol for Wireless Sensor Networks," 5*th International Conference on Future Information Technology* (*Future Tech*), Busan, 21-23 May 2010, pp. 1-5.

[18] S. K. Narang, G. Shen and A. Ortega, "Unidirectional Graph-Based Wavelet Transforms for Efficient Data Gathering in Sensor Networks," *IEEE International Conference on Acoustics Speech and Signal Processing* (*ICASSP*), Dallas, 14-19 March, 2010, pp. 2902-2905.

[19] M. B. Yassein, A. Al-zou'bi, Y. Khamayseh 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. doi:10.4156/jdcta.vol3.issue2.yassein

[20] T. M. Cover and J. A. Thomas, "Elements of Information Theory," 2nd Edition, John Wiley & Sons, Chichester, 2006.

[21] S. Pattem, B. Krishnamachari and R. Govindan, "The Impact of Spatial Correlation on Routing with Compression in Wireless Sensor Networks," *Proceeding of the Third International Symposium on Information Processing in Sensor Networks*, Los Angeles, 26-27 April 2004, pp. 28-35. doi:10.1145/984622.984627

[22] X. N. Fan and Y. l. Song, "Improvement on LEACH Protocol of Wireless Sensor Network," *International Conference on Sensor Technologies and Applications* (*SENSORCOMM* 2007), Valencia, 14-20 October 2007, pp. 260-264.

[23] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, Cambridge, 4-7 January 2000, p. 10. doi:10.1109/HICSS.2000.926982