ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 29 (2012) 2009 – 2015

Procedia Engineering

www.elsevier.com/locate/procedia

2012 International Workshop on Information and Electronics Engineering (IWIEE)

Energy Efficient and Balanced Cluster-Based Data Aggregation Algorithm for Wireless Sensor Networks

Jun Yue*, Weiming Zhang, Weidong Xiao, Daquan Tang, Jiuyang Tang

Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology Changsha-410073, China

Abstract

Data aggregation is an effectual approach for wireless sensor networks (WSNs) to save energy and prolong network lifetime. Cluster-based data aggregation algorithms are most popular because they have the advantages of high flexibility and reliability. But the problem of unbalanced energy dissipation is the inherent disadvantage in cluster-based WSNs. This paper addresses this problem in cluster-based and homogeneous WSNs in which cluster heads transmit data to base station by one-hop communication, and proposes an energy efficient and balanced cluster-based data aggregation algorithm (EEBCDA). It divides the network into rectangular grids with unequal size and makes cluster heads rotate among the nodes in each grid respectively, the grid whose cluster head consumes more energy has more sensor nodes to take part in cluster head rotation and share energy load, by this way, it is able to balance energy dissipation. Besides, it adopts some measures to save energy. The results of simulation show that EEBCDA can remarkably enhance energy efficiency, balance energy dissipation and prolong network lifetime.

Keywords: Wireless sensor networks; cluster-based; data aggregation; energy efficient; balanced energy dissipation.

1. Introduction

Recent advances in wireless communications and electronics have led to the development of WSNs, which are composed of many small-size, low-cost, low-power and multifunctional sensor nodes [1]. These nodes are densely deployed either inside or very closed to the phenomenon that is being monitored, every

* Corresponding author. Tel.: +86-013548535473.
E-mail address: yuejun.nudt@gmail.com.

doi:10.1016/j.proeng.2012.01.253

node has the ability to sense, process and transmit data to a base station (BS). Due to the limitations of size and cost, the nodes are equipped with small batteries with limited energy. Therefore, it is crucial to minimize the energy consumption to prolong network lifetime.

Data aggregation, which is the process of aggregating the data from multiple nodes to eliminate redundant transmission and provide fused data to BS, is considered as an effectual technique for WSNs to save energy [2]. The most popular data aggregation algorithms are cluster-based data aggregation algorithms, in which the nodes are grouped into clusters and each cluster consists of a cluster head (CH) and some members, each member transmits data to its CH, then, each CH aggregates the collected data and transmits the fused data to BS.

The cluster-based WSNs have an inherent problem of unbalanced energy dissipation. Some nodes drain their energy faster than others and result in earlier failure of network. Some researchers have studied this problem and proposed their algorithms which have both advantages and disadvantages. Our motivation is to propose a novel solution to this problem in the cluster-based and homogeneous WSNs, in which the CHs transmit data to BS by one-hop communication, with an objective of balancing energy consumption by an energy efficient way and prolonging network lifetime.

2. Related Work

The first proposed cluster-based algorithm for WSNs is LEACH [3]. It divides the operation into rounds, and randomly selects new CHs in each round to distribute the energy load among all nodes.

Some algorithms balance energy dissipation by forming equal clusters. HEED [4] periodically selects CHs and builds equal clusters according to a hybrid of the residual energy of nodes and a secondary parameter. The algorithm proposed in [5] divides the network into equal square zones basing on location information, the nodes in the same zone form a cluster and only new CHs are chosen in each round.

UCS [6] is the first proposed algorithm to resolve the problem of unbalanced energy dissipation by forming unequal clusters, but UCS is not practicable because the CHs are performed by some super nodes all the time and are deployed at pre-determined locations. EECS [7] and EDUC [8] are unequal cluster-based algorithms for WSNs in which the CHs transmit data to BS by one-hop communication. EECS is proposed for homogeneous WSNs, it elects some tentative CHs randomly and utilizes a cluster head competition method to choose CHs from tentative CHs, after that, each ordinary node selects CH basing on the distance from itself to CH and the distance from CH to BS to construct unequal clusters. EDUC is proposed for heterogeneous WSNs, in which each node computes its waiting time according to its residual energy firstly, when a node's waiting time is over and it has not received any cluster head competition message, it broadcasts a cluster head competition message within a radio range which is determined by the distance from itself to BS to declare that it is a CH, if a node receives cluster head competition message before its waiting time expires, it records the information of the CH and give up the competition, in the end, each ordinary node chooses the closest CH to join a cluster to form unequal clusters.

3. Problem Outline

3.1. Network model

To simplify the network model, we consider a WSN that N homogeneous nodes are uniformly dispersed within a square deployment area with border B. The left bottom vertex of deployment area locates at (O_x, O_y) in Cartesian coordinate plane. In addition, we make a few assumptions: 1) BS and all nodes are time synchronized and are stationary after deployment, BS is located at (BS_x, BS_y) which is out of deployment area; 2) communication is symmetric and every node is able to adjust its transmission power level according to the distance to the receiver; 3) BS and all nodes are location-aware.

The first two assumptions are familiar in other algorithms. The third assumption is reasonable in many applications of WSNs in which the sensed data only make sense with location information, for other applications, the locations of BS and nodes can be easily obtained by utilizing localization algorithms. The time synchronization algorithms and localization algorithms are not discussed in our work.

3.2. Energy consumption model

We use the same energy consumption model used in EECS. The free space model is used if the distance between the transmitter and receiver less than a threshold d_0, otherwise, the multipath model is used. The energy spent for transmitting an *l*-bit message over distance *d* is

$$E_{Tx}(l, d) = \begin{cases} l \times E_{elec} + l \times \varepsilon_{fs} \times d^2, & d < d_0 \\ l \times E_{elec} + l \times \varepsilon_{mp} \times d^4, & d \geq d_0 \end{cases} \quad (1)$$

where E_{elec} is the energy dissipated per bit to run the transmitter or the receiver circuit, ε_{fs} or ε_{mp}, is the energy dissipated per bit to run the transmit amplifier. To receive this message, the expended energy is

$$E_{Rx}(l) = l \times E_{elec} \quad (2)$$

The consumed energy of aggregating *m* messages with *l*-bit is

$$E_A(m, l) = m \times l \times E_{DA} \quad (3)$$

where E_{DA} is the energy dissipated per bit to aggregate message signal.

3.3. Problem statement

In the cluster-based WSNs in which the CHs transmit data to BS by one-hop communication, there are three reasons leading to unbalanced energy dissipation: 1) a CH often spends more energy than a member; 2) the amounts of received data of CHs are different; 3) the distances of transmitted data of CHs in different regions are different. Both LEACH and equal cluster-based algorithms are not able to balance the energy dissipation completely. Unequal cluster-based algorithms are considered as better solutions, they form clusters with unequal size according to the distance from each CH to BS, expect that the CH further away from BS has less members so that it is able to consume less energy to receive data and preserve more energy to transmit data. But the existing unequal cluster-based algorithms are deficient. First of all, they do not consider the distribution of CHs in CHs rotation scheme, the selected CHs are random scattered in network and the purpose of unequal clustering is affected. In addition, many unequal cluster-based algorithms make some ordinary nodes choose further CHs but not the closest CHs to form unequal clusters, so that these nodes have to spend excessive energy to transmit data.

4. EEBCDA Details

The operation of EEBCDA is also divided into rounds and every round consists of a set-up phase and a steady-state phase, especially, there is a network-division phase before the first round. The network is divided into rectangular regions firstly, called swim lanes, then, each swim lane is further partitioned into smaller rectangular regions, called grids. The node with maximal residual energy of each grid is selected as CH. The grids further away from BS are bigger and have more nodes to participate in CHs rotation. An overview of EEBCDA is depicted by Fig 1, in which the dashed lines mark the division of swim lanes, the dotted-dashed lines denote the division of grids and the dotted lines indicate the division of clusters.

4.1. Network-division phase

Without loss of generality, we assume that the BS is above the deployment area along Y-axis. At first, the deployment area is divided into S rectangular swim lanes along X-axis. All swim lanes have equal width W, and the length of each swim lane is equal to the border of deployment area. We use a sequence of integers from 1 to S as the IDs of swim lanes, and the ID of the leftmost swim lane is 1.

Then, each swim lane is partitioned into several rectangular grids along Y-axis. Each grid of each swim lane is assigned a level, we also use a sequence of integers starting from 1 as the levels of grids in each swim lane, and the level of the bottommost grid is 1. Each grid has the same width with swim lane. Both the number of grids and the length of each grid in a swim lane are related with the distance from the swim lane to BS. EEBCDA adjusts the size of each grid by setting its length. For different swim lanes, the further a swim lane is away from BS, the fewer grids it has. For same swim lane, the grid further away from BS has longer length. By this way, EEBCDA makes sure that the grid further away from BS has larger size. We define an array A with S elements, in which the k-th element is the number of grids in swim lane k. Each grid is assigned a tuple (i, j) as ID, which means that it is in swim lane i and has level j. In addition, we define S arrays to denote the lengths of grids, the v-th array H_v is the lengths of grids in swim lane v, and the w-th element h_{vw} of H_v is the length of grid (v, w). The bounds of grid (i, j) are

$$O\_x + (i-1) \times W < x \le O\_x + i \times W \tag{4}$$

$$O\_y + \sum_{k=1}^{k \le j-1} h_{ik} < y \le O\_y + \sum_{k=1}^{k \le j} h_{ik} \tag{5}$$

Finally, BS broadcasts a BS_MSG $((O\_x, O\_y), (BS\_x, BS\_y), B, W, S, A, H_1, \dots, H_S)$ message to all nodes and each node calculates the ID of the grid that it belongs to by the pseudo-code given in Fig 2, in which (x, y) is node's location, (v, w) is the ID of node's grid and a_v is the v-th element of A.

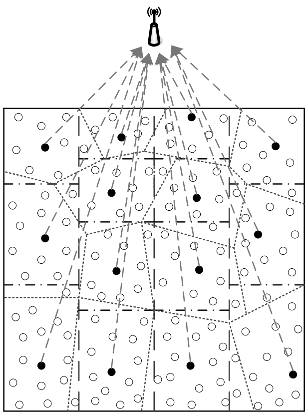

Fig. 1. An overview of EEBCDA.

```
v ← ⌈(x − O_x)/W⌉
u ← y − O_y
n ← 1
While n ≤ a_v
    If u ≤ Σ_{k=1}^{k≤n} h_vk
        w ← n
        Break
    End if
    n ← n + 1
End while
```

Fig. 2. Pseudo-code for each node to calculate its grid ID.

4.2. Set-up phase

The set-up phase performs CHs selection and cluster formation. In each round, the node with maximal residual energy in a grid is selected as the grid's CH. If there are two or more nodes with maximal residual energy, the CH is further selected according to node ID.

In the first round, the CH of a grid is selected by the cooperative work of nodes within the grid. Firstly, each node sends a *NODE_MSG* (k, (v, w), Er, (x, y)) message to other nodes in the same grid, where k is node ID, (v, w) is node's grid ID, Er is node's residual energy and (x, y) is the location of node. To save energy, the range of *NODE_MSG* only has to cover the node's grid. By this way, every node is able to collect the information of all nodes in the same grid and select the CH of its grid.

In each round after the first round, the CHs are selected by the cooperative work of the CHs of last round (LRCH). At the last time of data gathering in every round, every member transmits its residual energy information along with its data to CH. At the beginning of current round, each LRCH sorts all nodes in its cluster of last round in terms of grids, and chooses a tentative cluster head (TCH) for each grid, then, sends a *LRCH_TCH_MSG* message which contains the information of TCHs to the LRCHs of corresponding grids, the transmission distance of *LRCH_TCH_MSG* is the maximal distance between this LRCH and the vertexes of all corresponding grids. Finally, each LRCH chooses a CH from grid's TCHs.

Once the CHs have been selected, each CH broadcasts a *CH_BRDCST_MSG* message to inform other nodes. The range of *CH_BRDCST_MSG* of each CH only has to cover the CH's grid, by this way, EEBCDA not only guarantees that every ordinary node is covered by at least one message, but also saves energy. Then, each ordinary node chooses the closest CH to join a cluster. In the end, each CH sets up a TDMA schedule for its members, and transmits the schedule to the members in its cluster.

4.3. Steady-state phase

This phase of EEBCDA is similar to LEACH. Firstly, every member sends data to CH during its allocated transmission slot according to the TDMA schedule. Afterward, every CH aggregates the collected data and sends the fused data to BS. For the sake of CHs selection in next round, each member transmits its residual energy along with its data to CH at the last time of data gathering in every round.

5. EEBCDA Analysis

EEBCDA introduces a novel CHs rotation scheme which divides the network into unequal grids and chooses the node with the maximal residual energy in a grid as the grid's CH. This scheme makes CHs regularly scatter in the network. In the grid further away from BS, although the CH has more members and consumes more energy in a round, there are more nodes to take part in the grid's CH rotation. Hence, the energy dissipation of nodes in different grids is balanced on a long view. If the size of each grid is set properly, EEBCDA is able to achieve excellent performance in balancing energy dissipation.

EEBCDA adopts some measures to enhance energy efficiency. Firstly, except the first round, CHs are selected by the LRCHs not by the competition of more nodes. Secondly, the necessary information for selecting CHs is transmitted along with sensed data. Thirdly, the transmission distances and destination nodes of messages such as *NODE_MSG*, *LRCH_TCH_MSG* and *CH_BRDCST_MSG* are limited to required range. Fourthly, each ordinary node chooses the closest CH to join a cluster.

6. Simulation

We evaluate the performance of EEBCDA by simulation and use EECS for comparison. Similar to EECS, we also assume an ideal MAC layer and error-free communication links for simplicity, and take statistics of the energy dissipation whenever a node transmits or receives messages or performs data aggregation. The division of network in EEBCDA is set as S=4, W=50, A={3,4,4,3}, H_1=H_4={100,70,30},

H_2=H_3={80,60,40,20}. Like other similar algorithms, we assume that the sensed data are highly correlated in the same cluster so that each CH is able to aggregate some l-bit data messages into one l-bit data message. The other parameters of simulation are listed in Table 1.

Table 1. Parameters of simulation

Parameter	Value
Number of nodes	400
Deployment area	(0, 0)-(200, 200)m
Location of BS	(100, 250) m
Initial energy of each node	0.5 J
E_{elec}	50 nJ/bit
ε_{fs}	10 pJ/(bit • m^2)
ε_{mp}	0.0013 pJ/(bit • m^4)
d_0	87 m
E_{DA}	5 nJ/(bit • signal)
Message size	800 bit

First of all, we measure the lifetime of network. Fig 3 gives the number of living nodes over rounds. As evident from the figure, EEBCDA has a longer network lifetime than EECS. The first node of EEBCDA and EECS dies in the 591st round and 447th round, the last node dies in 698th round and 604th round, respectively, EEBCDA improves the network lifetime over EECS by 32.21% and 15.56%.

Secondly, we compare the energy dissipation of EEBCDA and EECS. We take statistics of the total residual energy of network over rounds, as shown in Fig 4. It is explicit that EEBCDA has more residual energy than EECS in every same round, which intuitively illuminates that EEBCDA is more energy efficient than EECS. The ratio of time interval between the time when the first node dies and the time when the last node dies to the full time of network is able to indicate the balanced extent of energy dissipation, and the algorithm with smaller ratio has a better performance in aspect of balancing energy dissipation. The values of this ratio of EEBCDA and EECS are 0.15 and 0.26, the result of contrast shows that EEBCDA is able to achieve more balanced energy dissipation than EECS.

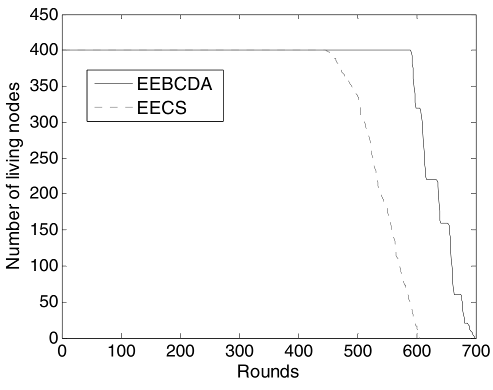

Fig. 3. The number of living nodes over rounds.

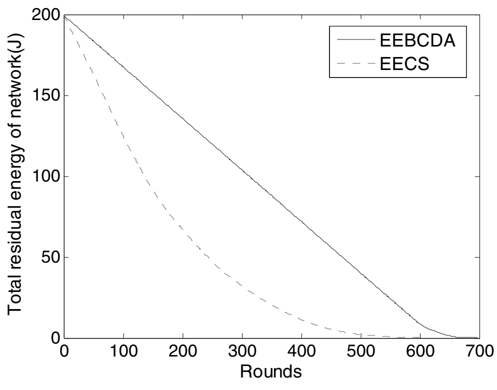

Fig. 4. The total residual energy of network over rounds.

7. Conclusion and Future Work

In this paper, we focus on the problem of unbalanced energy dissipation in cluster-based and homogeneous WSNs in which CHs transmit data to BS by one-hop communication, and propose a novel cluster-based data aggregation protocol named EEBCDA. It divides network into grids with unequal size, the grid further away from BS has bigger size and more nodes. The CHs rotation is performed in each grid. Although the CHs in the grids which are further away from BS consume more energy in each round, these grids have more nodes to participate in CHs rotation and share energy load, so that EEBCDA is able to balance energy dissipation on a long view. To improve energy efficiency, EEBCDA adopts some measures to save energy. The simulation results show that EEBCDA outperforms EECS in aspects of network lifetime, energy efficiency and balanced extent of energy dissipation.

The optimal division of network depends on the specific application, as a future work, we will formulate the parameters of EEBCDA to maximize the performance of algorithm. In addition, we will address the problem of unbalanced energy dissipation in cluster-based WSNs in which CHs transmit data to BS by multi-hop communication.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (60903225).

References

[1]Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E. A survey on sensor networks. *IEEE Commun Mag*, 2002, 40(8), p. 102–114.

[2]Rajagopalan R, Varshney PK. Data-aggregation techniques in sensor networks: a survey. *IEEE Commun Surv & Tutor*, 2006, 8(4), p. 48–63.

[3]Heinzelman WB, Chandrakasan AP, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks. *In Proc of IEEE Conf on System Sciences*, 2000, p. 3005–3014.

[4]Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans on Mobile Computing*, 2004, 3(4), p. 366–379.

[5]Al-Karaki JN, Ul-Mustafa R, Kamal AE. Data aggregation in wireless sensor networks-exact and approximate algorithms. *In Proc of Workshop on High Performance Switching and Routing*, 2004, p. 241–245.

[6]Soro S, Heinzelman WB. Prolonging the lifetime of wireless sensor networks via unequal clustering. *In Proc of IEEE Parallel and Distributed Processing Symposium*, 2005, p. 1–8.

[7]Ye M, Li C, Chen G, Wu J. An energy efficient clustering scheme in wireless sensor networks. *Ad Hoc & Sensor Wireless Networks*, 2006, 3, p. 99–119.

[8]Yu J, Qi Y, Wang G. An energy-driven unequal clustering protocol for heterogeneous wireless sensor networks, *J Control Theory Appl*, 2011, 9(1), p. 133–139.