An Efficient Aggregation and Routing Algorithm Using Multi-hop Clustering in Sensor Networks⋆

Bo-Hyeong Lee, Hyung-Wook Yoon, Tae-Jin Lee, and Min Young Chung

School of Information and Communication Engineering
Sungkyunkwan University, Suwon, KOREA
{shaak,hwyoon,tjlee,mychung}@ece.skku.ac.kr

Abstract. Sensor networks consist of sensor nodes with small-size, low-cost, low-power, and multi-functions to sense, to process and to communicate. Minimizing power consumption of sensors is an important issue in sensor networks due to limited power in sensors. Clustering is an efficient way to reduce data flow in sensor networks and to maintain less routing information. In this paper, we propose a multi-hop clustering mechanism using global and local ID to reduce transmission power consumption and an efficient routing method to improve performance of data fusion and transmission. Simulation results show that our algorithm can increase life time and disperse power consumption among the nodes.

1 Introduction

Sensor networks consist of small size, low cost and low power sensor nodes with multi-functions to sense, to process and to communicate [1]. Minimizing power consumption of sensors is one of the most important issues in sensor networks due to limited power of sensor nodes. If a sensor node consumes all of its energy, the sensor network may not guarantee reliable data transmission due to possible partition of the network. Thus, algorithms to minimize nodes' energy consumption and to maximize network life time have been actively studied [1], [2], [3], [4].

Since each sensor node senses and transmits local proximity information, sensed data in neighbor nodes tend to have similar characteristics to one another. In order to reduce the amount of data for transmission, data with similar information of the nodes need to be aggregated, which is called data fusion or aggregation [2]. The mechanism with data fusion can lead to less energy consumption than the mechanism without data fusion because data fusion reduces total data flow in the overall network. In general, it is very difficult to optimize data fusion in sensor networks. So, the method to group sensor nodes into a cluster and aggregate data in the cluster has been proposed in [3].

Clustering is able to reduce total routing messages as well as sensed information data. However, the larger the size of a sensor network, the less the merit is. Hierarchical clustering can increase scalability, but this is more energy-consuming because data transmission can concentrate on specific nodes or some routes, and total length of ID in the header and routing table grows as the number of levels in the hierarchy increases.

⋆ This paper was partially supported by BK21 program. Tae-Jin Lee is the corresponding author.

In this paper, we propose a multi-hop clustering and algorithm for efficient data fusion and routing using global ID, local ID and cluster scheduling instead of using hierarchical clustering in order to reduce the number of cluster head (CH) nodes and concentrated traffic on specific nodes.

Our algorithm has the following advantages:

1) **Efficient data aggregation in a multi-hop cluster:** In each aggregation procedure, each node in a multi-hop cluster is allowed to transmit data only once by the cluster schedule, which is more efficient than the aggregation in a hierarchical cluster because some nodes in a hierarchical cluster may transmit data more than once.

2) **Distribution of energy consumption when aggregating:** In a hierarchical cluster, fixed aggregating nodes exist (e.g., CH nodes) and energy consumption concentrates on these nodes. Our algorithm extends single hop clustering and constructs multi-hop clusters to distribute energy consumption over many aggregation nodes using cost of local IDs.

3) **Load balancing by construction of multipaths** : Our algorithm constructs multipaths from CH nodes to sink node during the multi-hop clustering stage. When aggregated data are sent to the sink node, multipaths are used instead of the fixed single shortest path to extend network life time.

The remainder of the paper is organized as follows. In Section 2, we discuss some of the related works. In Section 3, we propose the mechanism with multi-hop clustering, ID construction, routing path construction, and data aggregation. In Section 4, simulation results and analysis are presented. Section 5 concludes the paper.

2   Related Works

Conventional routing protocols are not suitable to aggregate data and to maximize network life time in sensor networks due to sensor nodes' limited properties (e.g., bandwidth, transmission range, thousands of nodes in a network, etc.). So, Low-Energy Adaptive Clustering Hierarchy (LEACH), which is a clustering-based protocol for sensor networks, has been proposed [3]. In LEACH, formation of clusters among sensor nodes is conducted locally, and CH nodes compress local data of the nodes in their clusters to reduce the amount of communications. And the role of CH nodes is periodically rotated for distributing energy consumption. LEACH assumes that the base station (sink node) is located far from the other nodes and each node can transmit data to the base station directly. This assumption may not be suitable because of the limited capacity of sensor nodes (especially, small energy). In this paper, we assume that the base station is located at radio transmission distance as other normal sensor nodes. So, data transmission from a normal node to the sink node traverses node along a routing path.

Power-Efficient GAthering in Sensor Information Systems (PEGASIS) is a chain-based protocol that minimizes power consumption [5]. In PEGASIS, each node communicates only with close neighbor nodes and data are relayed to the sink node, thus reducing the amount of energy spent. However, all nodes are assumed to have global knowledge of the network, which is very difficult in real networks.

The hierarchical ad hoc routing is combined with clustering in sensor networks [10], [11]. If the number of levels the in hierarchical clustering is large, less routing information
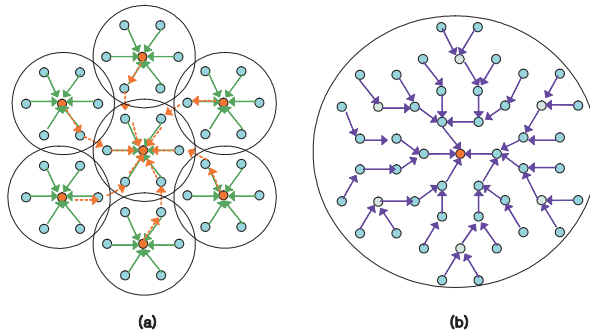
Fig. 1. (a) Hierarchical clustering and (b) multi-hop clustering.

is required [6]. However, high levels of the hierarchy is more burdensome and leads to long delay. Also, routing paths might be overlapped to transmit data as hierarchical level increases and data are concentrated in CH nodes when aggregating data (see Fig. 1). Hence, we propose and use multi-hop clustering structure, in which the distance is D hop from CH node to the nodes in a cluster, to distribute energy consumption and reduce routing overhead.

3   Proposed Multi-hop Clustering and Aggregate Routing

3.1   Multi-hop Clustering and Assignment of Node ID

Network life time, the time until the battery of any node drains-out for the first time, depends on how to determine CH nodes, i.e., clustering algorithm. Since the purpose of this paper is to realize efficient data aggregation and routing by use of local IDs for sensor nodes in a cluster, we assume that CH node is determined according to each node's current residual energy.

When the power of randomly placed sensor nodes is on, each node can become a CH node based on residual energy. If the remaining energy of a node is high, the node set short random time to become a candidate for a CH node. The node that first claims to be a CH node after random time becomes a CH node. A new CH node broadcasts CH broadcast signal with the CH's global ID, node ID, and the cluster's schedule to its neighbor nodes. The schedule of the cluster is based on the CH node's start time (see Fig. 2). Each node can know the transmission time at each level. Level i denotes the number of hops from a CH to a node in the cluster.

Neighbor nodes, which receive the CH signal, write CH ID in their information table (see Fig. 4 - (a)) and abandon being a CH node. These neighbor nodes may receive other CH broadcast signals. Then, the nodes examine their own information table, and select the CH with the smallest hop count and distance. Each node constructs a node ID based on the selected CH node's information and the node itself's (see Fig. 3). If a node determines its ID is within MAX hops from the CH, the node broadcasts CH broadcast signal again to other neighbor nodes after specified time. Nodes receiving this CH broadcast signal repeat the construction of information table with node ID.
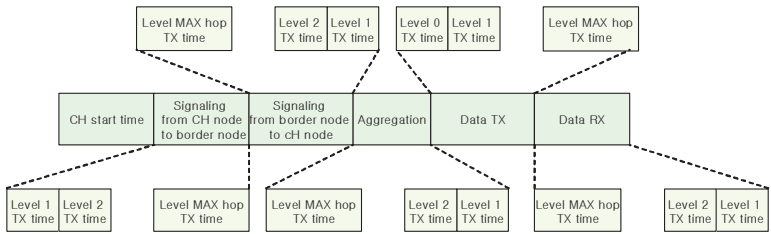
Fig. 2. Schedule in a CH broadcast signal.

CH ID	Residual energy	Hop count from CH node (level)	Shortest path cost to CH node	Link count

Fig. 3. Proposed information element (node ID).

CH ID	Neighbor node ID	Received signal power	Shortest path cost to CH node

(a) Information table

NC CH ID	Next node ID	Hop count from CH of NC	Level of NC border node	Shortest path cost from the CH of NC node

(b) NC routing table

CH source	CH destination	Previous node for reverse path	Next node for forward path	Hop #	Sequence #

(c) Inter−cluster routing table

Fig. 4. Information and intra- and inter-cluster routing table structure.

3.2   Construction of Routing Paths

Intra-cluster routing from an intra-cluster node to the CH node: When each node construct its information table with node ID, a node records the cost of the shortest path to the CH and decides the routing path to the CH node. The construction of the shortest path is similar to the Distributed Bellman-Ford (DBF) method. The difference between DBF and our proposed intra-cluster routing method is that the proposed routing mechanism can reduce the computational load of routing path (see Fig. 5).

Routing from a CH node to a neighbor cluster(NC): The border node, connected node to a neighbor cluster, plays an important role in inter-cluster routing among CH nodes. If a border node receives neighbor cluster's CH broadcast signal, the border node transmits information of the NC to its CH node. Nodes, which relay the information from the border node to the CH node, record the hop count and the other information in the NC routing table (see Fig. 4.-(b)) and construct routing paths needed when the CH node transmits data to other clusters' CH nodes. Construction of the routing path from the border node having information of NC to the CH node is similar to section 3.1 (see Fig. 6).

Inter-cluster routing from a CH to the sink node: Inter-cluster routing, which transmits aggregate data from a CH of a cluster to the sink node along the borer node of NCs, employs the same routing mechanism as in conventional ad hoc networks except replacing a node with a cluster. Routing paths determined from the previous step (see
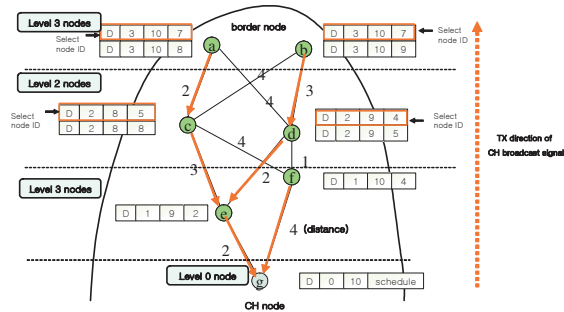
Fig. 5. Construction of intra-cluster routing path from an intra-cluster node to the CH node.
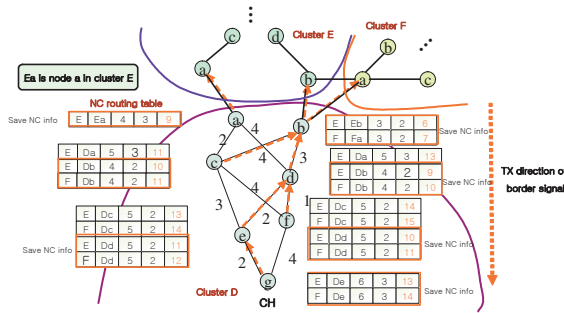

Fig. 6. Construction of routing path from a CH node to NC.

Fig. 6) are used for inter-cluster routing between CHs. Inter cluster routing table is shown in Fig. 4-(c). CH nodes manage inter-cluster routing tables which save the next cluster information to transmit aggregated data to the sink node.

3.3 Data Fusion Procedure

Nodes in a cluster periodically sense their environment and sensed transmit information to the CH node using intra-cluster routing paths. In the multi-hop clustering, the nodes at MAX hop distance from the CH node (local MAX) start to transmit in the beginning of aggregation, and the nodes at the next level transmit their own generated data and the data received from other nodes. Each node in the cluster must transmit only once to a neighbor node along the shortest path. However, data flow may be concentrated on some nodes along the shortest path, which causes unnecessary energy consumption. So, we propose the followings.

Level scheduling: When nodes start data aggregation, the nodes at MAX hop distance (level D) to the CH node transmit data. For example, when nodes in level D are able to transmit by the cluster schedule, which is informed in the clustering procedure, they transmit data to one of the nodes in level D-1. Nodes in level D-1 combine their own generated data with the received data, and transmit data again to the nodes in level D-2 when the nodes in D-1 can transmit by the cluster schedule. In this way, nodes are
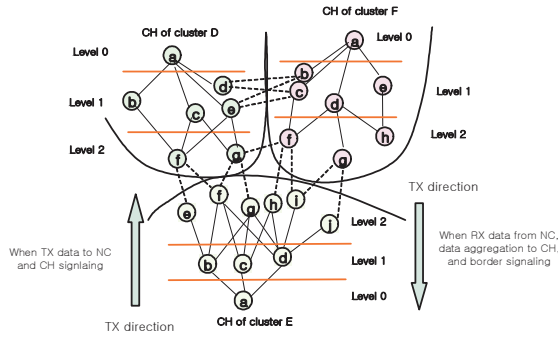
Fig. 7. Selection of receiving node from the viewpoint of cluster E.

allowed to transmit to the nodes in lower level during their level period. Finally, the CH node receives all of the nodes' data in the cluster while each node is guaranteed to transmit data only once.

Receiving node distribution: The nodes on the routing paths or receiving from many sensor nodes rapidly consumes their energy. So, it may not minimize total power consumption if we select the next hop node according to the shortest path routing. We propose two routing methods other than the shortest path. One is the random selection method and the other one is the load balancing method. In random selection method, a node transmitting data selects a random node among the nodes within the transmission distance. This realizes the random distribution of receiving nodes simply, to reduce the concentration of data to some nodes. The load balancing method selects receiving nodes based on the probability determined from cost (e.g., energy, hop count). That is a node can select a receiving node by considering the condition of both the node itself and candidate receiving nodes. For example,a node without enough energy to transmit data to a node at the long distance, increases the probability of selecting a close receiving node.

In both methods, a receiving node is selected according to the level of a transmitting node and the direction of transmission. In case of transmiting to the CH node, one of the nodes at lower level is selected. In case of transmiting to an NC, one of the nodes at higher level or in NC is selected (see Fig. 7).

4   Simulation Results

We have conducted simulations to evaluate the performance of the proposed multi-hop clustering and routing. The sensor network for the simulation consists of $50 \sim 400$ nodes distributed randomly in the square area of 100 x 100 meters. The sink node is located at the position (10,10). The proposed algorithm is used for inter-cluster routing and the AODV [7] algorithm is used for intra-cluster routing. We use the same radio model as in [8]. Signals regarding clustering and routing are assumed to be broadcast with the maximum transmission range of 20m. Data are transmitted using modified RTS-CTS-data [9] as in wireless local area network. We also assume that there are no collisions and bit errors by the cluster schedule when nodes transmit and receive data. Simulations are

Table 1. Parameters used in the simulation.

Parameter	Value
MAX transmission range	20 m
MAX hop D	3
Initial energy	0.01 J
RX energy consumption	50 nJ/bit
TX energy consumption	50nj/bit + 0.1 pJ/bit/m^2 × distance^2
Aggregation period	80 times slots
Number of simulations	1000

executed until the network life time. The parameters in the simulation are summarized in Table 1.

In Fig. 8, we observe that network life time of the proposed multi-hop clustering and routing is almost the same regardless of routing methods to balance receiving nodes. The reason is that all the data packets in a cluster must be eventually collected and transmitted through the CH node to go to the sink node. So, in order to increase the network life time, periodic reclustering may be required to switch CH nodes with those having larger residual battery power. The shortest path routing, however, shows larger variance than the other receiving node selection methods as shown in Fig 9. Basically, routing paths constructed as in Section 3 is the shortest paths. Therefore, transmitting data in a cluster tends to use the same routing paths and thus energy consumption is concentrated on the nodes along the routing paths. In case of random and load balanced routing methods, energy consumption is scattered to other nodes because transmitting nodes select the next receiving node according to the remaining energy or the probability. We also note that as the number of nodes increase, random selection surpasses the load balance since the receiving nodes tend to be distributed evenly.
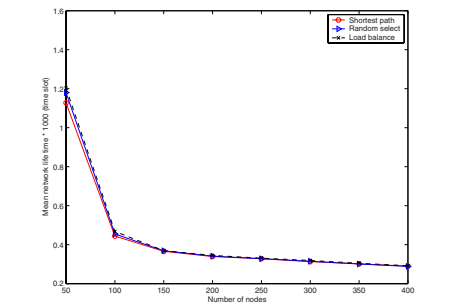

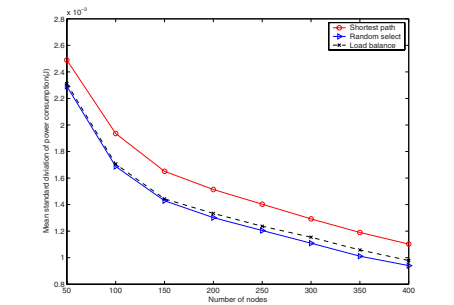**Fig. 8.** Mean network life time as different selection mechanisms of receiving nodes are used.

Fig. 9. Mean standard deviation of energy consumption in sensor nodes.

5   Conclusion

In this paper, we have proposed an efficient aggregation and routing mechanism using local ID in multi-hop cluster sensor networks. The proposed algorithm can reduce the routing computational overhead and simplify routing paths since it employs level scheduling and intra- and inter-cluster routing. In addition, each node transmits data at once in the aggregation procedure, preventing from unnecessary energy consumption which has been in the previous works. We have compared the network life time and energy consumption to evaluate the performance via simulations. The proposed load balanced routing has been shown to disperse energy consumption by selecting the next receiving node at random or by the load-balancing regime.

References

1. I. F. Akyildiz and W. Su. A Survey on Sensor Network. In *IEEE communications Magazine*, volume 40, pages 102–114, Aug. 2002.
2. W. R. Heinzelman, J. Kulik, and H. Balakrishnan. Adaptive Protocols for Information Dissemination in Wireless Sensor Networks. In *Proc. of ACM MobiCom*, pages 174–185, 1999.
3. W. R. Heinzelman, A. Chandrakasan, and H.Balakrishnan. Energy-Efficient Communication Protocol for Wireless Microsensor Networks. In *Proc. of IEEE Sys. Sci.*, pages 1–10, Jan. 2000.
4. S. Singh, M. Woo, and C. S. Raghavendra. Power-Aware Routing in Mobile Ad Hoc Networks In *Proc. of ACM MobiCom*, pages 181–190, Nov. 1998
5. S. Lindsey and C. S. Raghavendra. PEGASIS: Power-efficinet GAthering in Sensor Information System. In *Proc. of IEEE Aerospace Conference*, volume 3, pages 1125–1130, Mar. 2002.
6. J. Sucec and I. Marsic. Clustering Overhead for Hierarchical Routing in Mobile Ad-hoc Networks. In *Proc. of IEEE INFOCOM*, volume 3, pages 1698–1706, Jun. 2002.
7. C. E. Perkins and E. M. Royer. Ad-hoc On Demand Distance Vector Routing. In *Proc. of IEEE WMCSA*, pages 90–100, Feb. 1999.
8. W. Heinzelman, et. al. Energy-scalable Algorithms and Protocols for Wireless Microsensor Network. In *Proc. of International Conference on Acoustics, Speech and Signal Processing*, volume 6, pages 3722–3725, Jun. 2000.
9. P. Karn. MACA - A New Channel Access Method for Packet Radio. *ARRL/CRRL Amateur*, volume 6, pages 134–140, 2000.
10. A. Iwata, C. C.Chiang, G. Pei, M. Gerla and T. W. Chen. Scalable Routing Strategies for Ad-hoc Wireless Networks. *IEEE Journal on Selected Area in Communication*, volume 17, pages 1369–1379, Aug. 1999.
11. R. Ramanathan and M. Steenstrup. Hierarchically-organized, Multi-hop Mobile Wireless Networks for Quality-of-Service Support. *ACM/Baltzer Mobile Networks and Applications*, volume 3, pages 101–119, 1998.