SELECTION MECHANISM OF ENERGY-EFFICIENT DATA AGGREGATION NODE IN WIRELESS SENSOR NETWORKS

Yongbin Yim, Euisin Lee, Seungmin Oh and Sang-Ha Kim
The Department of Computer Engineering, Chungnam National University, 220 Gung-dong, Yuseong-gu, Daejeon, Republic of Korea

Keywords: Aggregation, Data Aggregation Node, Energy Consumption, Redundant Sensed Data, Wireless Sensor Networks.

Abstract: In-network data aggregation is one of the most important issues for achieving energy-efficiency in wireless sensor networks since sensor nodes in the surrounding region of an event may generate redundant sensed data. The redundant sensed data should be aggregated before being delivered to the sink to reduce energy consumption. Which node should be selected as a Data Aggregation Node (DAN) for achieving the best energy efficiency is a difficult issue. To address this issue, this letter proposes a scheme to select a DAN for achieving energy-efficiency in an event region. The proposed scheme uses an analytical model to select the sensor node that has the lowest total energy consumption for gathering data from sensor nodes and for forwarding aggregated data to a sink, as a DAN. Analysis and simulation results show that the proposed scheme is superior to other schemes.

1 INTRODUCTION

Wireless sensor networks consist of a great number of sensor nodes which are deployed in an interest region for event monitoring (Akyildiz et al., 2002). The sensor nodes are normally powered by batteries with limited energy resource (Akyildiz et al., 2002)(Luo and Liu, 2007). Thus, the primary challenge for this energy-constrained sensor networks is to design efficient schemes to reduce the energy consumption.

Generally, when an event happens, many sensor nodes in a region around it generate redundant sensed data (Luo and Liu, 2007). Then, as shown in Fig 1(a), the total energy consumption will be significant if each sensor node in the event region directly disseminates its sensed data to a sink. A general solution is that a Data Aggregation Node (DAN) in the event region gathers the sensed data from the other sensor nodes and forwards aggregated data to the sink (Akyildiz et al., 2002). However, it is a difficult issue that which node should be selected as the DAN so that the energy consumption can be minimized.

Two schemes (Zhang and Cao, 2004)(Petrovic et al., 2003) have been proposed to address this issue. As shown in Fig. 1(b), the scheme proposed in (Zhang and Cao, 2004) selects the sensor node at the center of the event region as a DAN to minimize the

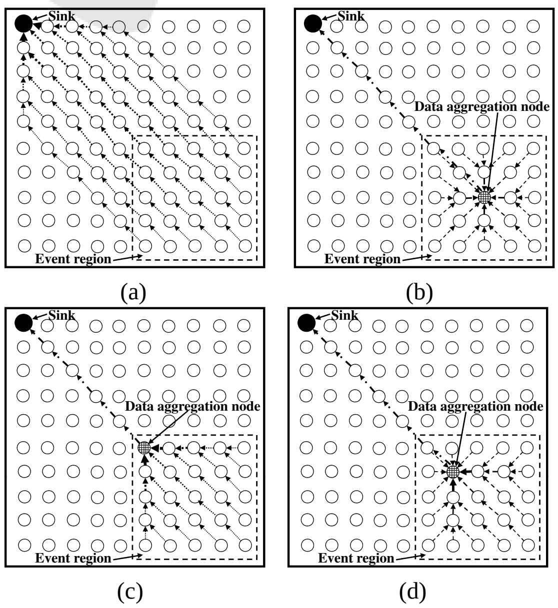

Figure 1: Data dissemination schemes from an event region to a sink: (a) the directed scheme, (b) the center scheme, and (c) the nearest scheme, and (d) the proposed scheme.

energy consumption for gathering sensed data from sensor nodes. As shown in Fig. 1(c), the scheme proposed in (Petrovic et al., 2003) selects the sensor no-

Yim Y., Lee E., Oh S. and Kim S..
SELECTION MECHANISM OF ENERGY-EFFICIENT DATA AGGREGATION NODE IN WIRELESS SENSOR NETWORKS.
DOI: 10.5220/0003907503910395
In *Proceedings of the 2nd International Conference on Pervasive Embedded Computing and Communication Systems* (PECCS-2012), pages 391-395
ISBN: 978-989-8565-00-6

de nearest to the sink as a DAN to minimize the energy consumption for forwarding the aggregated data to the sink. However, the center scheme hardly reduces the energy consumption for gathering sensed data from sensor nodes in the event region if the event region is small, and however consumes much energy for forwarding the aggregated data to the sink if the event region is far from the sink. In contrast, the nearest scheme hardly reduces the energy consumption for forwarding the aggregated data to the sink if the event region is close to the sink, and however consumes much energy for gathering sensed data from sensor nodes in the event region if the event region is big.

Hence, to address this issue, as shown in Fig. 1(d), this letter proposes a scheme which selects a sensor node in the event region as the DAN so that the total energy consumption for gathering sensed data and for forwarding aggregated data can be minimized. This letter presents an analytical model for calculating the energy consumption for gathering sensed data and for forwarding aggregated data. Analysis and simulation results show that the proposed scheme outperforms both the center scheme and the nearest scheme in terms of the energy consumption.

2 THE PROPOSED SCHEME

2.1 Network Model

We describe a network model to implement our work. A sink and sensor nodes are deployed in a sensor network. Each node is aware of its own location information through GPS device or localization techniques (Karp and Kung, 2000). All sensor nodes can know the location of the sink by programming the location to the sensor nodes or flooding the location by the sink. If an event happens, sensor nodes in its surrounding region detect it and generate data with their won location information because many applications in wireless sensor networks require the location of source data, for example, target tracking and habitat monitoring. Then, the sensor nodes disseminate their data to the sink by geographic routing (Karp and Kung, 2000). After receiving data with the location information from the sensor nodes, the sink calculates the location information of the event region and selects one among the sensor nodes in the event region to function as a Data Aggregation Node (DAN), through the location information of the sensor nodes. The sink sends a DAN_Selection message with the location information of the event region to the DAN by geographic routing. The DAN floods a DAN_Announcement message with its location information in the event region through the well known geocasting protocols (Stojmenovic, 2004). Through the messages, the other sensor nodes in the event region get to be aware of location information of the DAN and disseminate their data to the DAN. The DAN gathers data from sensor nodes in the event region and forwards aggregated data to the sink.

2.2 Data Aggregation Node (DAN) Selection

We develop an analytical model to select a DAN in an event region. In the analytical model, the total energy consumption function E_t for data dissemination from sensor nodes in the event region to the sink consists of the energy consumption function E_g that the DAN gathers the sensed data from the sensor nodes, and the energy consumption function E_f that the DAN forwards the aggregated data to the sink.

The energy consumption model proposed in (Heinzelman et al., 2002) is exploited by our analytical model which defines the communication cost ($E_C(k,r)$) as the energy consumption of transmitting ($E_T(k,r)$) and receiving ($E_R(k,r)$) a k-bit packet with a distance r:

$$E_C(k,r) = E_T(k,r) + E_R(k,r). \quad (1)$$

$E_T(k,r)$ and $E_R(k,r)$ are defined as

$$E_T(k,r) = E_{elec} \cdot k + \varepsilon_{amp} \cdot k \cdot r^2 \quad (2)$$

$$E_R(k,r) = E_{elec} \cdot k, \quad (3)$$

where the transmitter of sensor nodes dissipates E_{elec} = 50 nJ/bit to run the transmitter or receiver circuitry and ε_{amp}= 100 pJ/bit/m^2 for transmit amplifier. Since every sensor node uses same transmission power, the r is its transmission range.

Consider a set of sensor nodes S = $\{n_1, n_2, ... ,n_N\}$ in an event region. If the sink has topology information of all sensor nodes, it can optimally select a sensor node n_o in the set S as the DAN whose total energy consumption cost is minimal. The total energy consumption cost function E_{t_o} of the optimal scheme is defined as

$$E_{t\_o}(o) = E_{g\_o}(o) + E_{f\_o}(o). \quad (4)$$

E_{g_o} and E_{f_o} are defined as follows.

$$E_{g\_o}(o) = \sum_{i=1, i\neq o}^{N} real\_hops(i,o) \cdot E_C(sen\_size, r) \quad (5)$$

$$E_{f\_o}(o) = real\_hops(o,s) \cdot E_C(aggre\_size, r) \quad (6)$$

Here, the $real\_hops(i,o)$ and $real\_hops(o,s)$ are the number of real hop counts between two nodes o and i, and between the node o and the sink s, respectively. The *sen_size* is a packet size of sensed data generated in a sensor node. The *aggre_size* is a packet size of aggregated data that the DAN aggregates the gathered data from the other source nodes in the set S. The *aggre_size* depends on data aggregation ratio (Luo and Liu, 2007) which is the ratio of the amount of outgoing data to that of incoming data at a DAN. This ratio at a DAN may vary widely according to applications in sensor networks.

Although the optimal scheme can minimize the total energy consumption, it is not practical since it assumes that the sink can know real hop counts between any two nodes from knowledge of the network topology. Therefore, we propose a scheme to derive heuristically the total energy cost function through only location information in geographic routing when the sink knows only location information of sensor nodes in the event region. We assume that sensor nodes are densely and uniformly deployed in a sensor field and all sensor nodes have the same transmission range. Accordingly, the total energy consumed by transmitting a data packet along a multi-hop path in geographic routing is proportional to the Euclidean distance between a source node and a destination node. This assumption is justified by the fact that the Euclidean distance between two nodes in a dense and uniform wireless sensor network is approximately proportional to the hop count between the same nodes (Niculescu and Nath, 2003). We note that such an energy model is also adapted by several existing energy-efficient communication protocols in wireless sensor networks (Kim et al., 2003).

About the set S = $\{n_1,\ n_2,\ ...\ \ ,n_N\}$, the proposed scheme defines the total energy consumption cost function E_{t_p}(j) of each node n_j in the set S as

$$E_{t\_p}(j) = E_{g\_p}(j) + E_{f\_p}(j). \tag{7}$$

E_{g_p} and E_{f_p} are defined as follows.

$$E_{g\_p}(j) = \sum_{i=1, i\neq j}^{N} geo\_hops(i,j) \cdot E_C(sen\_size, r) \tag{8}$$

$$E_{f\_p}(j) = geo\_hops(j,s) \cdot E_C(aggre\_size, r) \tag{9}$$

Here, the $geo\_hops(i,j)$ is the number of the expected hop counts between locations of two nodes i and j, and the $geo\_hops(i,j)$ is the number of the expected hop counts between locations of sensor node i and the sink s. The $geo\_hops(i,j)$ is calculated as

$$geo\_hops(i,j) = \left(\frac{d(i,j)}{Single\_Hop\_Pro_{ave}}\right) + 1, \tag{10}$$

where the $d(i,j)$ is defined as the Euclidean distance between the nodes i and j, and the $Single\_Hop\_Pro_{ave}$ means an average single-hop progress and is defined as the expected value of the difference between the before-hop distance (between the sender node and the destination node) and the after-hop distance (between the next-hop node and the destination node) (Chen et al., 2007). We use a value calculated by the equation (14) in (Chen et al., 2007) as the $Single\_Hop\_Pro_{ave}$ where ρ is the average number of neighbors within the transmission range r of the sender and is given by $\rho = \pi r^2 \lambda$ where λ is the expected number of nodes within a unit area.

Hence, the sink determines a sensor node n_j in the set S as the DAN, whose $E_{t\_p}(j)$ is minimal.

3 ANALYSIS

Based on our analytical model, we analyze the four schemes in Fig. 1. The total energy consumption of the direct scheme in Fig. 1(a) is defined as

$$E_{t\_d} = \sum_{i=1}^{n} geo\_hops(i,g) \cdot E_C(sen\_size, r). \tag{11}$$

The total energy consumption of the center scheme in Fig. 1(b) is defined as

$$E_{t\_c} = E_{g\_c}(c) + E_{f\_c}(c). \tag{12}$$

E_{g_c} and E_{f_c} are defined as follows.

$$E_{g\_c}(c) = \sum_{i=1, i\neq c}^{n} geo\_hops(i,c) \cdot E_C(sen\_size, r) \tag{13}$$

$$E_{f\_c}(c) = geo\_hops(c,g) \cdot E_C(aggre\_size, r) \tag{14}$$

The energy consumption of the nearest scheme in Fig. 1(c) is defined as

$$E_{t\_n} = E_{g\_n}(n) + E_{f\_n}(n). \tag{15}$$

E_{g_n} and E_{f_n} are defined as follows.

$$E_{g\_n}(n) = \sum_{i=1, i\neq n}^{N} geo\_hops(i,n) \cdot E_C(sen\_size, r) \tag{16}$$

$$E_{f\_n}(n) = geo\_hops(n,g) \cdot E_C(aggre\_size, r) \tag{17}$$

The energy consumptions of the proposed scheme in Fig. 1(d) and the optimal scheme are the equation (7) and (4), respectively.

4 PERFORMANCE EVALUATION

We compare the performance of the proposed scheme (PS) with that of the optimal scheme (OS), the center scheme (CS), and the nearest scheme (NS) through analytical and simulation results. We implemented the four schemes in Network Simulator Qualnet 4.0. The models of sensor nodes are followed by the specification of MICA2. The transmitting and receiving energy consumption rates of sensor nodes are 42mW and 29mW, respectively. The transmission range of sensor nodes is 50m. The size of the sensor network is set to 500m*500m where 1000 nodes are uniformly distributed. As default setting, the dimension of an event region is $7500m^2$, the distance from the event region to the sink is 150m, the size of a sensed data is 30 bytes, and the data aggregation ratio is 0.3.

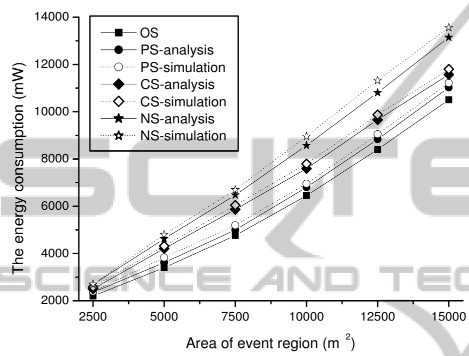

Figure 2: The energy consumption for the area of event region.

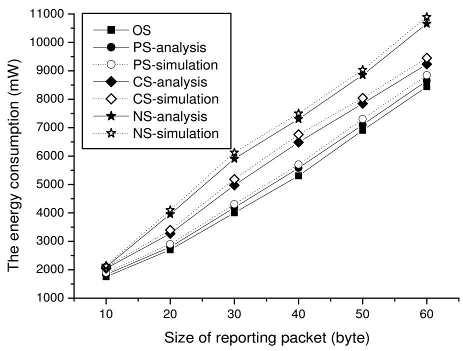

Figure 3: The energy consumption for the size of sensed data packet.

We present four analysis and simulation results for four parameters in relation to total energy cost for data dissemination from an event region to a sink. Fig. 2 and 3 show analysis and simulation results for the area of an event region and for the size of sensed data in relation to data gathering cost of a DAN. The optimal scheme consumes least energy because it selects an optimal DAN in ideal conditions with network topology information. Since the proposed scheme considers both data gathering cost and aggregated data forwarding cost, its energy consumption approximates to that of the optimal scheme irrespective of the area or the size. When the area or the size is small, the center scheme and the nearest scheme have similar energy consumption because the effect for reducing data gathering cost happens very little. However, if the area or the size increase, the energy consumption of the center scheme is less than that of the nearest scheme and approximates to that of the optimal scheme as that of the proposed scheme. Because, the center scheme minimizes the data gathering cost which is much greater than aggregated data forwarding cost.

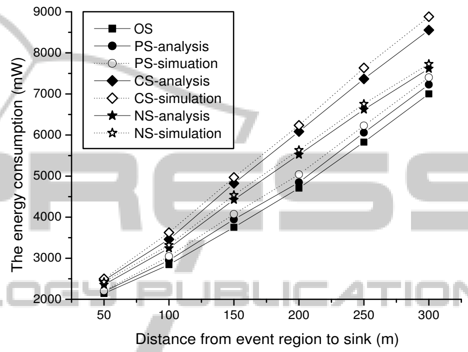

Figure 4: The energy consumption for the distance from event region to sink.

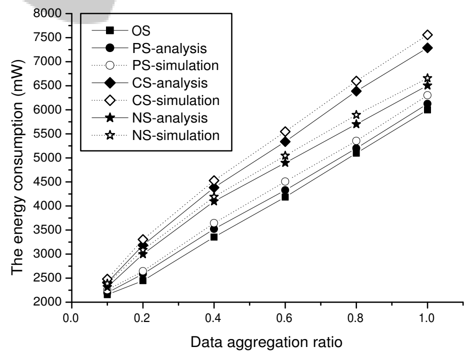

Figure 5: The energy consumption for data aggregation ratio.

Fig. 4 and 5 show analysis and simulation results for the distance from an event region to a sink and for the data aggregation ratio in relation to aggregated data forwarding cost from a DAN to the sink. The optimal scheme consumes least energy and the proposed scheme consumes energy approximate to the optimal scheme, because they considers both sensed data gathering cost and aggregated data forwarding cost. When the distance or the ratio is small, the energy consumption of the nearest scheme is similar to that of the center scheme because the effect for reducing data gathering cost happens very little. However, if the distance or the ratio increases, the energy consumption of the nearest scheme approximates to that of the optimal scheme as that of the proposed scheme. Because, the nearest scheme minimizes the aggregated data forwarding cost which is much greater than the data gathering cost.

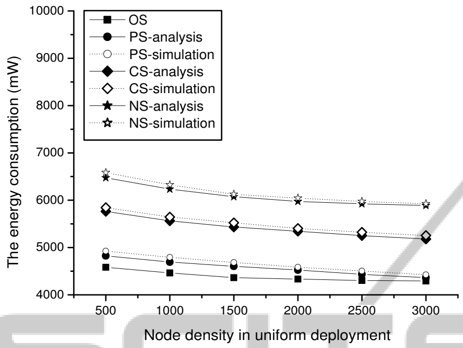

Figure 6: The energy consumption for the node density in uniform deployment.

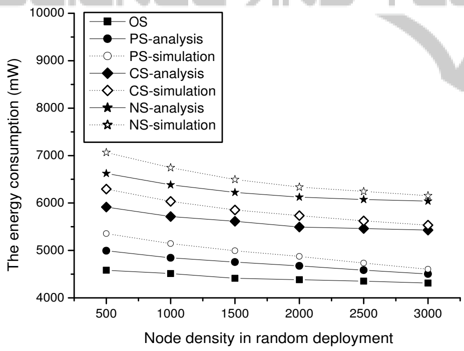

Figure 7: The energy consumption for the node density in random deployment.

Next, to justify our analytical model, we compare analysis and simulation results for the density of sensor nodes (namely, the number of sensor nodes) in their uniform and random deployment. Fig. 6 and 7 show analysis and simulation results in uniform and random deployment, respectively. In uniform deployment, the difference between analytical and simulation results is small. If the node density increases, analytical results is almost similar to simulation results. In random deployment, if the node density is small, the difference between analytical and simulation results is bigger than that in uniform deployment. However, if the node density increases, analysis results is approximate to simulation ones. It is because the real hop counts in real sensor network and the geographical hop counts in our analytical model between any two nodes are almost the same in high node density.

5 CONCLUSIONS

This letter presents a scheme to select a Data Aggregation Node (DAN) for minimizing the energy consumption for data dissemination from an event region to a sink. In the analytical model described herein, the proposed scheme selects a sensor node in the event region as the DAN, by which the total energy consumption for gathering sensed data from sensor nodes and for forwarding aggregated data to the sink is minimized. The slight difference between analytical and simulation results proves that our analytical model is well designed. Analytical and simulation results show that the proposed scheme is more energy-efficient than the center and nearest schemes.

REFERENCES

Akyildiz, I., Su, W., Sankarasubramaniam, Y., and Cayirci, E. (2002). A survey on sensor networks. In *IEEE Communication Magazine*. IEEE.

Chen, D., Deng, J., and Varshney, P. (2007). Selection of a forwarding area for contention-based geographic forwarding in wireless multi-hop networks. In *IEEE Transactions on Vehicular Technology*. IEEE.

Heinzelman, W., Chandrakasan, A., and Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. In *IEEE Transactions on Wireless Communications*. IEEE.

Karp, B. and Kung, H. T. (2000). Gpsr: Greedy perimeter stateless routing for wireless networks. In *ACM Mobicom*. ACM.

Kim, H., Abdelzaher, T., and Kwon, W. (2003). Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks. In *ACM SenSys*. ACM.

Luo, H. and Liu, Y.and Dasi, S. (2007). Routing correlated data in wireless sensor networks: A survey. In *IEEE Network*. IEEE.

Niculescu, D. and Nath, B. (2003). Dv based positioning in ad hoc networks. In *Journal of Telecommunication Systems*. Springer.

Petrovic, D., Shah, R., Ramchandran, K., and Rabaey, J. (2003). Data funneling: Routing with aggregation and compression for wireless sensor networks. In *IEEE International Workshop on SNPA*. IEEE.

Stojmenovic, I. (2004). Geocasting with guaranteed delivery in sensor networks. In *IEEE Wireless Communications*. IEEE.

Zhang, W. and Cao, G. (2004). Dctc: Dynamic convoy tree-based collaboration for target tracking in sensor networks. In *IEEE Transactions on Wireless Communications*. IEEE.