Multi-Chain Data Gathering in Wireless Sensor Networks

Shiow-Fen Hwang[†], Kun-Hsien Lu[††], Zen-You Huang[††], Chyi-Ren Dow[††]
Department of Information Engineering and Computer Science
Feng Chia University, Taichung, Taiwan 40724
Email:[†]sfhwang@mail.fcu.edu.tw, [††]{p9217988, m9503100, crdow}@fcu.edu.tw

***Abstract*—Data gathering is a fundamental operation in wireless sensor networks (WSNs). Due to the limited battery power, energy conservation is an essential issue to reduce the total energy consumption of sensor nodes and prolong the network lifetime. In addition, the transmission delay of a round is also an important factor in data gathering applications. In this paper, we propose a Multi-Chain Data Gathering (MCDG) protocol for WSNs. MCDG organizes all sensor nodes to form a single chain that is further divided into several subchains of unequal size. In each subchain, an aggregation node is assigned for reducing the transmission delay and some of the other nodes take turns to be a subchain leader for balanced energy consumption. All subchain leaders cooperate to forward data to the base station. The objective of MCDG is to achieve energy-efficient data gathering with delay reduction. Simulation results show that MCDG outperforms PEGASIS and COSEN in terms of the number of rounds and the *energy* × *delay* cost.**

***Index Terms*—wireless sensor networks, data gathering, chain, energy-efficient, delay reduction.**

I. INTRODUCTION

In recent years, with advances in micro-electro-mechanical systems (MEMS) technology, wireless communications and digital electronics have enabled the development of wireless sensor with low cost processor and low power consumption[1][2]. A wireless sensor network (WSN) usually consists of numerous sensors with light-weight and wireless communication ability and limited battery power. These sensors are deployed over an area of interest to collect useful information and transmit the collected data to a base station (BS) or a sink for post-processing. Therefore, wireless sensor networks have been generally applied to various military and civil applications, such as environment monitoring, disaster prediction, intrusion detection and so on.

Data gathering is an fundamental operation in WSNs. A simple strategy to accomplish data gathering is direct transmission. Each node collects data from its surroundings and transmits data to the BS directly in a round of communication. Since the BS is usually located far away from the sensor field, the energy cost to transmit to the BS for any node is high so that nodes will die quickly. Therefore, energy conservation is an essential issue for data gathering applications. There are two approaches that have been widely used for saving energy in WSNs. One approach is multi-hop relay, which utilizes cooperation among nodes. Since the energy cost for data transmission is proportional to the square of the distance of transmission, the use of multi-hop relay reduces the total energy consumption of nodes efficiently. The other approach is data aggregation[3], which combines two or more packets into a same-size packet. Hence, the amount of data packets transmitted in the network can be reduced.

Another important factor for data gathering applications, such as forest fire detection, is the transmission delay. As described in literature[4], the transmission delay can be measured as the number of transmissions to accomplish a round of data gathering. Therefore, by employing data gathering schedule and simultaneous data transmissions, the transmission delay can be reduced.

Various data gathering protocols, such as LEACH[5], PEGASIS[4][6], and PEDAP[7], have been proposed to minimize the total energy consumption of nodes and prolong the network lifetime. In these protocols, chain-based architecture is widely adopted to efficiently achieve the above goals. In general, the existing chain-based data gathering protocols group all nodes into a single chain[6][8] or multiple chains [4][9][10][11][12][13][14][15]. The protocols using a single chain can conserve energy because only one node becomes the leader to transmit data to the BS and every other node just transmits data to its next node. However, it also causes large transmission delay as the number of nodes becomes large. On the contrary, the protocols using multiple chains can reduce the transmission delay because all chains can perform data gathering simultaneously. But the use of multiple chains also increases the total energy consumption of nodes. Therefore, there is a tradeoff between the transmission delay and the energy consumption.

In this paper, we propose a Multi-Chain Data Gathering (MCDG) protocol in WSNs. MCDG first organizes all nodes to form a single chain by using the proposed chain formation algorithm. The single chain is then further divided into several subchains of unequal size. In each subchain, an aggregation node is assigned to reduce the transmission delay and some of the other nodes take turns to be a subchain leader for balanced energy consumption. These selected leader nodes cooperate to forward data to the BS. The objective of MCDG is to achieve energy-efficient data gathering with delay reduction. Simulation results indicate that MCDG outperforms PEGASIS and COSEN in terms of the number of rounds and the $energy \times delay$ cost.

The rest of this paper is organized as follows: Section II reviews the literatures about some chain-based data gathering protocols. The network model and energy model are introduced in Section III. Section IV presents the detail of the proposed multi-chain data gathering protocol. In Section V, we evaluate the proposed protocol and compare it with PEGASIS and COSEN. Finally, Section VI draws our conclusions.

II. RELATED WORK

Many chain-based data gathering protocols have been proposed in recent years. In this section, we briefly introduce PEGASIS[6], GSEN[10], and COSEN[11] which are related to our study.

A. PEGASIS

PEGASIS (*Power-Efficient GAthering in Sensor Information Systems*)[6] is a representative chain-based data gathering protocol. It organizes all nodes into a single chain by using greedy algorithm. Each node communicates with a close neighbor node, and takes turns to be a leader. The leader is responsible for transmitting the aggregated data of the chain to the BS in a round. Although PEGASIS reduces the total energy consumption of nodes than cluster-based data gathering protocols, such as LEACH[5], but it has some deficiencies. First, greedy algorithm is simple to construct a chain, however, long link between a pair of nodes may be generated. This is because a node in the chain may need to connect to another node that is not in the chain and far away from itself if most of nodes are added to the chain. Moreover, nodes far away from the BS consume much energy when transmitting data to the BS. Therefore, uneven energy consumption among nodes still exists in PEGASIS. Second, PEGASIS constructs a long chain and may result unacceptable transmission delay as the network size becomes large.

B. GSEN

N. Tabassum *et al*. proposed a two layer hierarchial routing protocol, called *Group-based SEnsor Network* (GSEN)[10]. GSEN divides all nodes into several groups. In each group, a chain is formed by using greedy algorithm as presented in PEGASIS. A group leader is rotated in a random order among nodes in a chain, and collects data. All groups leaders then form a high-level chain, and one random chosen high-level leader is responsible for transmitting the aggregated data to the BS. GSEN not only makes more number of rounds than PEGASIS, but also introduces reasonable delay in a round of data gathering. However, long link problem also exists in GSEN because of the use of greedy algorithm.

C. COSEN

N. Tabassum *et al*. also proposed a hierarchical chain-based protocol, called COSEN (*Chain Oriented SEnsor Network*)[11]. Similar to GSEN, COSEN groups all nodes into several low-level chains. Unlike the criterion of leader selection using a random fashion in GSEN, COSEN selects leaders according to the residual energy of nodes. In each low-level chain, the node with the maximum residual energy is elected to be the low-level leader. Besides, one high-level leader is elected from several low-level leaders based on the proposed formulation and transmits the aggregated data to the BS.

III. NETWORK AND ENERGY MODELS

Our proposed protocol, MCDG, is based on the typical data gathering applications of sensor networks where data are collected periodically from all sensor nodes to the BS. The network model of MCDG has the following assumptions.

1. All sensor nodes are homogeneous, energy constrained, and have the same capabilities.
2. The BS is located far away from the sensor field.
3. The BS has global knowledge of the network for constructing chain topology.
4. No mobility of the BS and sensor nodes.
5. Sensor nodes have the ability of power control to adjust their transmission power to communicate with any other sensor node and the BS.

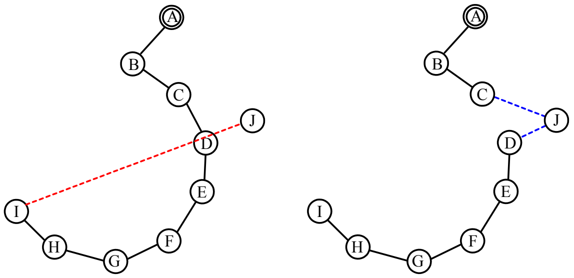
MCDG uses the first order radio model presented in LEACH[5] to calculate the energy consumption of a node. The energy consumed in transmitting a *l*-bit data packet over a distance *d* is shown in Equation (1).

$$E_{Tx}(l,d) = E_{elec} \times l + \varepsilon_{amp} \times l \times d^2 \qquad (1)$$

Equation (2) shows the energy consumed in receiving a *l*-bit data packet.

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

Here, a radio dissipates E_{elec} = 50 nJ/bit to run the transmitter or receiver circuitry, and ε_{amp} = 100 pJ/bit/m^2 to run the transmitter amplifier. It is obvious that a fixed energy consumption (i.e. $E_{elec} \times l$) is needed for transmitting or receiving a data packet, and an additional energy consumption for data transmission which depends on the distance of transmission d. Besides, A sensor node also consumes 5 nJ/bit/packet for data aggregation.



(a) PEGASIS (b) MCDG

Fig. 1. Single chain formation.

Algorithm 1 Single Chain Formation Algorithm

```
$V$ : the set of deployed nodes.
$N'$ : the set of nodes not in the chain yet.
$CHAIN$ : the set of nodes already added to the chain.
$s_a$ : the node with the index $a$ in the chain.
$d(a,b)$ : the distance between node $a$ and node $b$.
$AvgDist$ : the average length of all links in the chain.
$\alpha$ and $\beta$ : constants.
/*Initialization phase*/
$N' \leftarrow V$; $AvgDist \leftarrow 0$;
$s_0 \leftarrow$ the node closest to the BS;
$CHAIN \leftarrow \{s_0\}$; $N' \leftarrow N' - \{s_0\}$;
/*Chain formation phase*/
for $i = 0$ to $|V| - 2$ do
    Let $x \in N'$ with $d(s_i, x)$=min$\{d(s_i, p)|p \in N'\}$;
    $temp \leftarrow d(s_i, x)$;
    if $\frac{|CHAIN|}{|V|} \geq \alpha$ and $d(s_i, x) \geq \beta * AvgDist$ then
        for $j = 0$ to $i - 2$ do
            $Len[x, s_j]$=$d(s_j, x) + d(x, s_{j+1}) - d(s_j, s_{j+1})$;
        end for
        Select a node $s_j$ with minimum $Len[x, s_j]$;
        if $Len[x, s_j] < d(s_i, x)$ then
            Insert $x$ between $s_j$ and $s_{j+1}$ in $CHAIN$;
            $temp \leftarrow Len[x, s_j]$;
        else
            $s_{i+1} \leftarrow x$; /*$x$ becomes $s_{i+1}$.*/
        end if
    else
        $s_{i+1} \leftarrow x$;
    end if
    $CHAIN \leftarrow CHAIN \cup \{x\}$; $N' \leftarrow N' - \{x\}$;
    $AvgDist \leftarrow \frac{AvgDist*i+temp}{i+1}$;
end for
```

IV. MULTI-CHAIN DATA GATHERING

In this section, we present the proposed data gathering protocol, MCDG, which contains four phases: 1)single chain formation phase, 2)subchain formation phase, 3)aggregation node and leader node scheduling phase, and 4)data gathering and transmission phase. Besides, we also analysis the transmission delay incurred in a round of data gathering.

A. Single Chain Formation Phase

A simple method to construct a chain connecting all nodes is the greedy algorithm that is adopted by PEGASIS. The greedy chain starts with the furthest node from the BS. The closest neighbor not in the chain to this node is elected as the next node in the chain. This process is repeated until all nodes are added to the chain. However, the greedy chain may consists of long links among nodes. For example, in Fig. 1(a), a greedy chain, which starts with node A, is constructed by PEGASIS. It is obvious that the length of the link between node I and node J is more longer than those of the others. Node I and node J will consume more energy than other nodes when transmitting data to each other. Therefore, in order to reduce the effect of long links, we simply modify greedy algorithm as shown in Algorithm 1. The proposed single chain formation algorithm aims to eliminate long links as much as possible. The detail of single chain formation is as follows.

Initially, the node that is closest to the BS is assigned to construct the chain. This node selects its closest neighbor not in the chain as the next node and adds it to the chain. Hence, a new link (or a new edge) is created between both of them. The process of selecting next node is the same as the greedy algorithm used in PEGASIS and is repeated until the two following conditions are satisfied.

1. The ratio of the number of nodes added to the chain to the number of deployed nodes is larger than or equal to a threshold value α.
2. The length of the new link is larger than or equal to a value β times the average length of all links of the chain.

Here, the values, α and β, are set to determine whether a new link should be eliminated or not. In line 14-26 of Algorithm 1, a node x not in the chain is elected as the next node by node s_i that is at the end of the chain. When both of these two conditions are satisfied simultaneously, a node s_j $(j \leq i-2)$ with minimum increased chain length in the chain is elected if we insert node x between node s_j and node s_{j+1}. The increased chain length for node s_j is $d(s_j, x) + d(x, s_{j+1}) - d(s_j, s_{j+1})$. Therefore, if the minimum increased chain length is smaller than the length of the new link between node s_i and node x, node x will be inserted between node s_j and node s_{j+1}. The new link between node s_i and node x is thus eliminated. Otherwise, node s_i still selects node x as the next node and adds it to the chain.

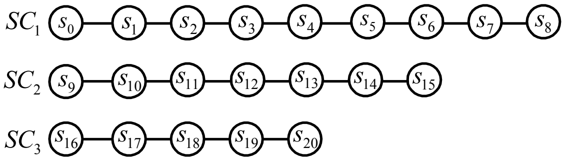


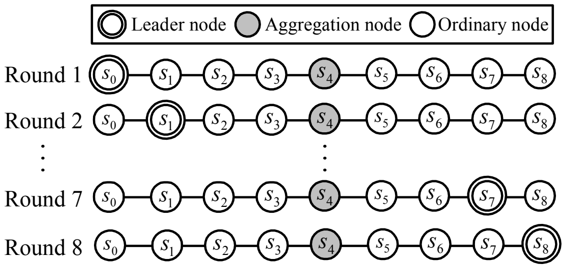
Fig. 2. An example of subchain formation.

Figure 1(b) shows a single chain constructed by MCDG. The long link between node I and node J is eliminated, and node J is inserted between node C and node D to minimize the increased chain length. Therefore, MCDG reduces the chain length compared to PEGASIS (in Fig. 1(a)) and can save energy in data transmissions. Moreover, the selections of α and β values will affect the chain length of MCDG compared to that of PEGASIS. We will evaluate these two values in Section V.

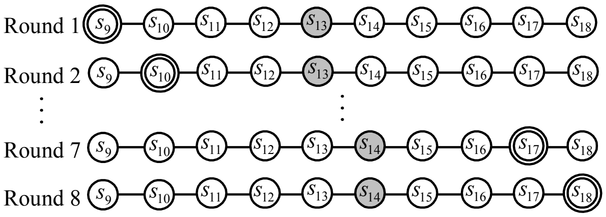
B. Subchain Formation Phase

Let the single chain produced by Algorithm 1 be $\{s_0, s_1, \ldots, s_{n-1}\}$, where s_0 is the node closest to the BS and s_{n-1} is the last node added to the chain. MCDG divides the single chain into k subchains of unequal size. The set of subchains is defined as $\{SC_1, SC_2, \ldots, SC_k\}$, where k is an odd number and $|SC_{i+1}| = |SC_i| - 2$, for all $1 \leq i \leq k-1$. Therefore, the number of nodes in the subchain SC_i is $|SC_i| = \frac{n}{k} + (k - 2 \times i + 1)$, for all $1 \leq i \leq k$. The first subchain SC_1 has maximum number of nodes while the last subchain SC_k has minimum number of nodes. The main objective of dividing a single chain into several subchains is to reduce the transmission delay of a round because all subchains can perform data collection simultaneously. Figure 2 is an example of dividing a single chain with 21 nodes into three subchains, SC_1, SC_2 and SC_3 with nodes 9, 7, and 5, respectively.

The above approach of dividing the single chain into k subchains is the case that n is divisible by k. However, when n is not divisible by k, we can simply assign the remainder to each subchain from



(a) $L = 9$



(b) $L = 10$

Fig. 3. Aggregation node and leader node scheduling.

SC_1 to SC_i in order, for some $i < k$. For example, if n is 23 and k is 3, the remainder is 2 (nodes). Therefore, three subchains, SC_1, SC_2, and SC_3 will have 10, 8, and 5 nodes, respectively.

C. Aggregation Node and Leader Node Scheduling Phase

In each subchain, an aggregation node and a leader node are assigned to collect and relay the aggregated data of the subchain. The other nodes is called ordinary node. Suppose that a subchain contains L nodes and starts with a node s_i, where i is the minimum index in the subchain. The selections of aggregation node and leader node are as follows.

When L is an odd number, a node s_j with the index $j = \lfloor \frac{L}{2} \rfloor + i$ becomes the aggregation node. The other nodes in the subchains take turns to be the leader node. Figure 3(a) shows an subchain containing 9 nodes. Node s_4 is elected as the aggregation node of the chain and the other nodes, $s_0 \sim s_3$ and $s_5 \sim s_8$, rotate to be the leader node from round to round. However, when L is a even number, we assign two nodes, s_j and s_{j+1}, where $j = \frac{L}{2} + i - 1$, to be the candidates for the aggregation node. The other nodes also take turns to be the leader node. Here, the aggregation node is elected depending on the index of the leader node. Assume that s_x is the leader node in current round, if x is small than j, node s_j then becomes the leader node in this round. However, if x is larger than $j+1$, node s_{j+1} will be the leader node. Figure 3(b) illustrates an subchain consisting of 10 nodes. Node s_{13} and node s_{14} are the candidates for the aggregation node. The other nodes, $s_9 \sim s_{12}$ and $s_{15} \sim s_{18}$, take turns being the leader node. When the leader node is one of nodes $s_9 \sim s_{12}$, node s_{13} is the aggregation node. Otherwise, node s_{14} becomes the aggregation node.

Since the aggregation node is fixed at or near to the center of the subchain, the transmission delay of a round can be further reduced by applying simultaneous data transmissions. Besides, the role of the leader is rotated among nodes in the subchain. When the number of nodes in the subchain increases, the average distance between the aggregation node and the leader node may also increase. This results in the increase of energy consumption for data transmission between the aggregation node and the leader node. Therefore, for each subchain, we set a value γ to limit the number of nodes to rotate the role of leader. For example, in Fig. 3(a), we set γ as 0.5. Therefore, 4 (=(9-1)*0.5) nodes, s_2, s_3, s_5, s_6, will take turns to be the leader node. The value γ is also evaluated in Section V.

D. Data Gathering and Transmission Phase

In a round, the process of gathering data from all nodes contains two parts: 1)intra-subchain communication and 2)inter-subchain communication.

First, intra-subchain communication takes place within each subchain. In a subchain, the aggregation node is responsible for collecting the aggregated data sent from ordinary nodes. In the beginning of a round, two ordinary nodes that are at the end of the subchain start to transmit their sensed data toward the aggregation node simultaneously. An ordinary node, which received data from its previous ordinary node, aggregates the received data with its own into a single data packet, and then transmit it to the next node until the aggregation node. As the aggregation node has collected the aggregated data from all ordinary nodes, it then transmits to the leader node. Figure 4(a) shows that an example of intra-subchain communication. In the example, the subchain contains 5 nodes. Node s_2 serves as the aggregation node, and the other nodes take turns to be the leader from round to round. In round 1, after the aggregation node s_2 has collected the

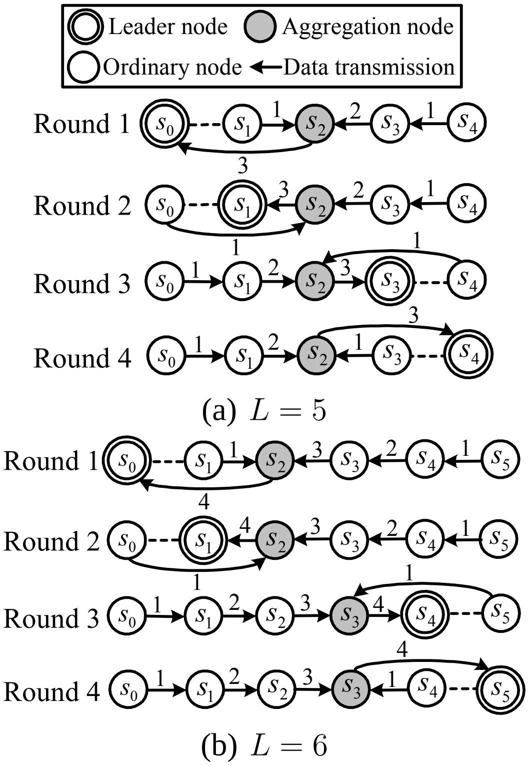


Fig. 4. Two examples of intra-subchain communication. The number near to a solid arrow indicates the time slot for a node to forward data.

aggregated data sent via two routes, $s_1 \rightarrow s_2$ and $s_4 \rightarrow s_3 \rightarrow s_2$, it then forwards data to the leader node s_0. The number near to a solid arrow indicates the time slot for a node to transmit data. In this case, the transmission delay for the subchain in a round is 3 units. Figure 4(b) illustrates another example of intra-subchain communication, where a subchain contains 6 nodes.

For inter-subchain communication, all leader nodes cooperate to forward the aggregated data toward the BS. As described in Section IV-B, there are k subchains, $SC_1, SC_2, \ldots, SC_k$, in the network. The leader node of the subchain SC_k forwards the aggregated data toward the leader node of the subchain SC_1 via the leader-to-leader route. Each leader node also performs data aggregation during inter-subchain communication. Finally, the leader node of the subchain SC_1 transmits the aggregated data of all nodes to the BS and a round is completed. Although the leader node in the subchain SC_1 does long-distance data transmission toward the BS, however, the subchain SC_1 has maximum number of nodes to rotate the role of the leader among nodes for sharing the energy consumption. Figure 5 shows an example of inter-subchain communication. After the leader node s_{16} of SC_3 received the aggregated data sent from the aggregation node s_{18}, it performs data aggregation and then forwards data to the next leader node s_9. Note that the leader node s_{16} has to wait 2 units of delay to start inter-subchain communication for collision avoidance. After the leader node s_0 received data from the leader node s_9, it then sends the aggregated data of the network to the BS. In this example, the transmission delay incurred in a round is 7 units.

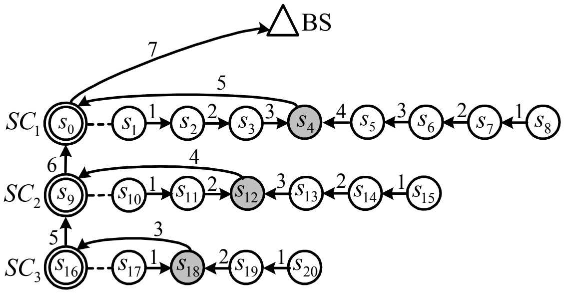


Fig. 5. An example of inter-subchain communication. The transmission delay of round is 7 units. The number near to a solid arrow indicates the time slot for a node to forward data.

E. Delay Analysis

In this subsection, we discuss the transmission delay of a round in MCDG. Suppose that n nodes is deployed in the network, and k subchains are required. Since the maximum number of subchains is restricted by the number of nodes, thus we have the following lemma.

Lemma 1: Given a network with n nodes. If k subchains are required in MCDG, then n has to be at least $k^2 + 3k$.

Proof: In MCDG, the number of nodes in a subschain is required to be at least 4, which are an aggregation node, a leader node, and two ordinary node. The set of the number of nodes in subchains is $\{2k+2, 2k, \ldots, 6, 4\}$. Hence, the sum n is k^2+3k.□

Based on the above relationship between n and k, and the following lemmas, we can obtain the upper bound of the transmission delay of MCDG in Theorem 1.

Lemma 2: Let n be divisible by k. If k is odd and $\frac{n}{k}$ is an integer, then the transmission delay of a round is $\lfloor\frac{n}{2k}\rfloor + \lfloor\frac{k}{2}\rfloor + 3$ units.

Proof: the transmission delay of a round, called T_{total}, is generated from three parts: 1)intra-subchain communication, 2)inter-subchain communication, 3)data transmission between the leader node and the BS. First, the transmission delay for intra-subchain communication, called T_{intra}, is determined by the units spent by the subchain with maximum number of nodes. Hence, T_{intra} is $\lfloor\frac{n}{2k}\rfloor + \lfloor\frac{k}{2}\rfloor + 1$ units. Second, the transmission delay for inter-subchain communication, called T_{inter}, is $k-1$ units. Third, the transmission delay for the leader node to the BS, called T_{BS}, is 1 unit. Here, note that some time slots in intra-subchain communication and inter-subchain communication overlap. For example, in Fig. 5, intra-subchain communication and inter-subchain communication overlap in time slot 5. The units for overlap, called T_o, is $k-2$ units. Consequently, we can get $T_{total} = T_{intra} + T_{inter} + T_{BS} - T_o = \lfloor\frac{n}{2k}\rfloor + \lfloor\frac{k}{2}\rfloor + 3$.□

When n is not divisible by k, then two cases that $\lfloor\frac{n}{k}\rfloor$ is either odd or even, are described in the following Lemma 3 and Lemma 4.

Lemma 3: Let $\frac{n}{k}$ be not an integer. If k is odd and $\lfloor\frac{n}{k}\rfloor$ is odd, then the transmission delay of a round is $\lfloor\frac{n}{2k}\rfloor + \lfloor\frac{k}{2}\rfloor + 4$ units.
Proof: similar to the proof of Lemma 2. □

Lemma 4: Let $\frac{n}{k}$ be not an integer. If k is odd and $\lfloor\frac{n}{k}\rfloor$ is even, then the transmission delay of a round is $\lfloor\frac{n}{2k}\rfloor + \lfloor\frac{k}{2}\rfloor + 3$ units.
Proof: similar to the proof of Lemma 2. □

Theorem 1: The transmission delay of a round in MCDG is at most $\lfloor\frac{n}{2k}\rfloor + \lfloor\frac{k}{2}\rfloor + 4$ units, where n is the number of nodes and k is the number of subchains.
Proof: the result follows from Lemma 2∼4 directly. □

We also briefly show the transmission delay of a round in COSEN for performance comparison. As described in Section II-C, COSEN groups n nodes into k low-level chains. When n is divisible by k, the transmission delay of a round is $\frac{n}{k} + k - 1$ units. However, if n is not divisible by k, the transmission delay of a round is $\lfloor\frac{n}{k}\rfloor + k$ units. From the above results of delay analysis, MCDG has lower transmission delay of a round than COSEN in all situations.

V. PERFORMANCE EVALUATION

We developed a custom simulator written in C++ Language. The proposed protocol, MCDG, is compared with two chain-based data gathering protocols, PEGASIS and COSEN. We first simulate the effect of the values, α, β, and γ on the performance of MCDG. Three metrics are then evaluated in the simulations, which are 1) the number of rounds until first node dies (FND), 2) the energy consumption per round (energy cost), and 3) the *energy* × *delay* cost.

Six simulation environments, from E1 to E6, are utilized to simulate the three metrics and are listed in Table I. The BS is located far away from the sensor field. All sensor nodes are randomly deployed. The size of a data packet is 2,000 bits. The initial energy for each node is 0.5 Joule. The first order radio model is used to evaluate the energy consumption of nodes.

We first compare the length of the chain constructed by MCDG with that constructed by PEGASIS. Figure 6 shows the percentage of decreased chain length for various α and β in a $100m\times100m$ network with 500 nodes. The result shows that the proposed single chain formation algorithm reduce the chain length up to 7% when α is 0.8 and β is 4. Figure 7 illustrates the number of rounds before first node dies (FND) for various γ under six environments. The value γ is used to limit the number of nodes to rotate the role of leader in a subchain. We vary γ from 0.2 to 1 and set k as 5. The result shows that when the number of nodes increases, the value γ should appropriately decrease to achieve highest number of rounds. We observe that γ can be set

TABLE I
SIMULATION ENVIRONMENTS.

Sensing area BS location	$100m\times100m$ (50,200)	$200m\times200m$ (100,400)
250 nodes	E1	E4
500 nodes	E2	E5
750 nodes	E3	E6

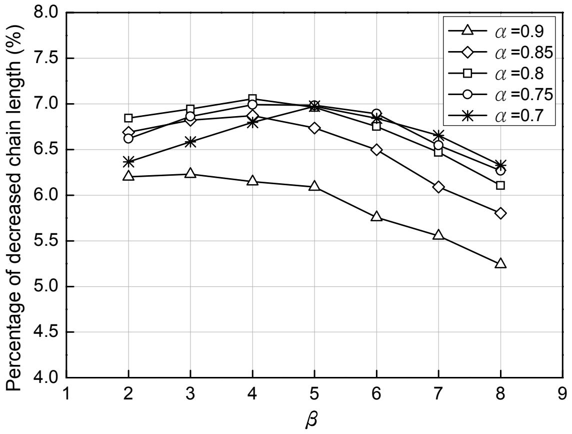


Fig. 6. The percentage of decreased chain length versus α and β values.

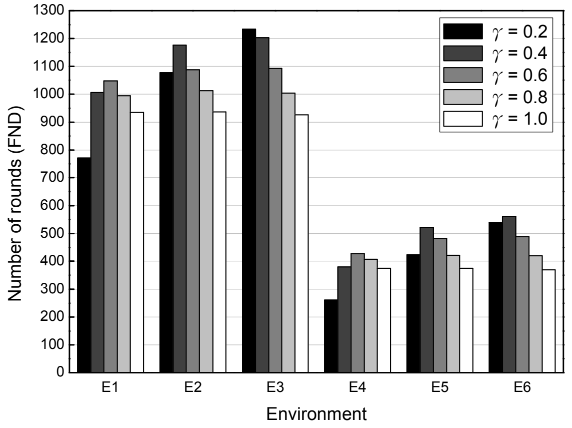


Fig. 7. The impact of the value γ under six environments.

as 0.4 for good performance in most environments. Hence, based on the above results, we set α, β, and γ as 0.8, 4, and 0.4, respectively for MCDG in the following simulations.

Figure 8 shows the number of rounds for PEGASIS, COSEN, and MCDG under six environments. We set k as 5 for COSEN and MCDG. Due to the existence of long links produced by using the greedy algorithm in PEGASIS, some nodes consume much energy in transmitting data to each other. Besides, nodes far away from the BS also consume more energy than nodes near to the BS when transmitting data to the BS. Hence, the first death node appears quickly in PEGASIS. Although COSEN selects leader nodes based on the residual energy of nodes, but it also suffers from the same problems in PEGASIS. On the contrary, the proposed single chain formation algorithm eliminates long links as much as possible. Therefore, MCDG achieves more number of rounds than PEGASIS and COSEN under all environments.

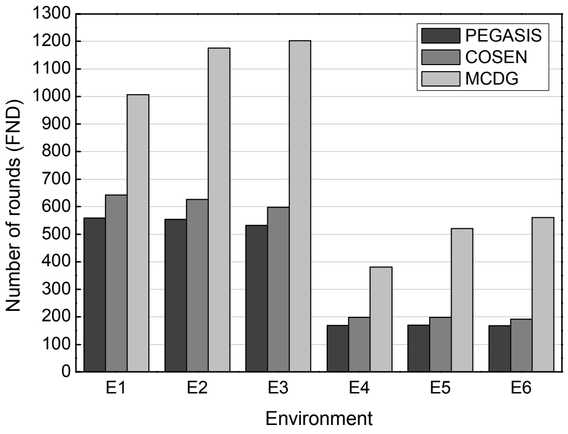


Fig. 8. The number of rounds for three protocols under six environments.

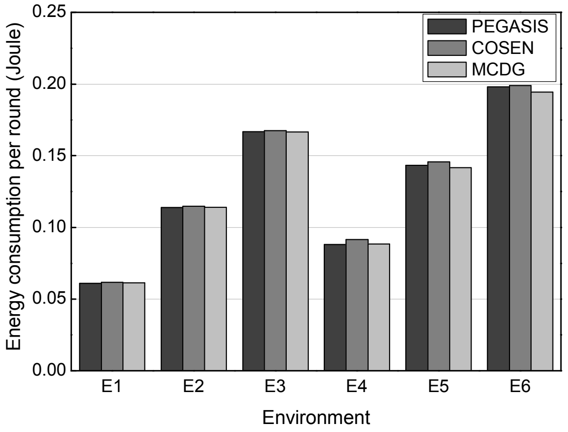


Fig. 9. The energy consumption per round for three protocols under six environments.

Figure 9 illustrates the energy consumption per round for three protocols under six environments. The energy consumption per round is the amount of energy consumed by all nodes to accomplish a round of data gathering and is treated as the energy cost. The result shows the energy consumption per round for three protocols are approximately same. Figure 10 shows the *energy* $\times$ *delay* cost, which is equal to the energy cost multiplied by the delay cost. Here, the delay cost means the transmission delay of a round. We list the delay cost for three protocols in Table II. The delay cost of PEGASIS is equal to the number of nodes since all nodes transmit their aggregated data one by one. The delay cost of

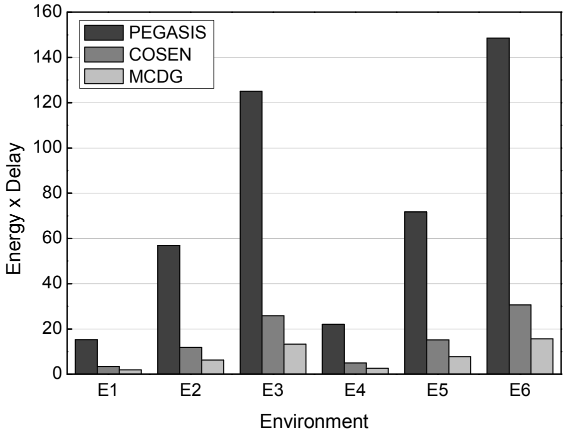


Fig. 10. The *energy* × *delay* cost for three protocols under six environments.

TABLE II
THE DELAY COST FOR PEGASIS, COSEN AND MCDG.

	250 nodes	500 nodes	750 nodes
PEGASIS	250	500	750
COSEN	54	104	154
MCDG	30	55	80

COSEN and MCDG can be calculated by equations described in Section IV-E. MCDG reduces the delay cost by employing several subchains of unequal size and simultaneous data transmissions in each subchain, hence, it achieves lower *energy* × *delay* cost than PEGASIS and COSEN.

Figure 11 shows the *energy* × *delay* cost of MCDG over the number of subchains. The maximum number of subchains for 250, 500, 750 nodes are 13, 19, and 25, respectively according to Lemma 1. When the number of subchains increases, the number of leader nodes participate in inter-subchain communication increases. Although this causes the slight increase of the energy consumption per round and the number of rounds drops slightly as shown in Fig. 12, but we observe that the delay cost is also reduced due to the increase of the number of subchains. Therefore, the *energy* × *delay* cost of MCDG is thus reduced.

VI. CONCLUSIONS

In this paper, we propose a Multi-Chain Data Gathering (MCDG) protocol that aims to reduce the transmission delay of a round and achieves a lower *energy* × *delay* cost. In our protocol, all nodes first form a single chain using the proposed chain formation algorithm. The single chain is then further divided into several subchains of unequal size. For intra-subchain communication, an aggregation node and a leader node are elected to collect the aggregated data of the chain in each subchain, and the transmission delay can be reduced. For inter-subchain communication, several leader nodes cooperate to forward the aggregated data of all nodes to the BS. Simulation results show that MCDG outperforms PEGASIS and COSEN in terms of the number of rounds and *energy* × *delay* cost.

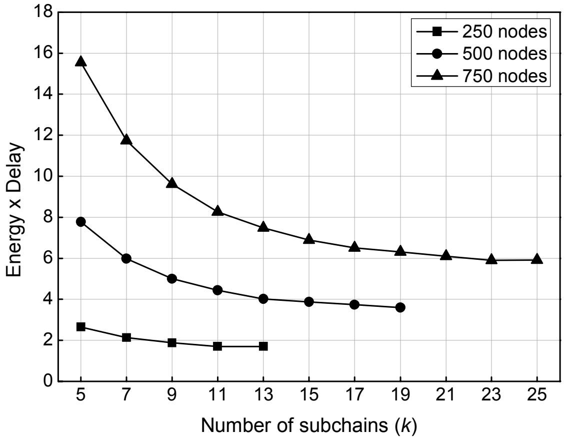


Fig. 11. The energy × delay cost for MCDG over the number of subchains in a $200m \times 200m$ network.

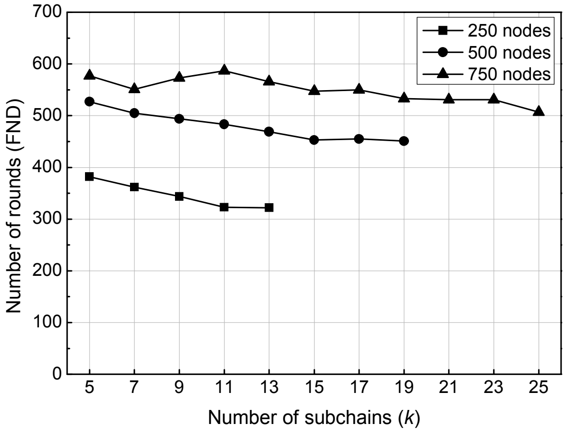


Fig. 12. The number of rounds for MCDG over the number of subchains in a $200m \times 200m$ network.

ACKNOWLEDGMENT

This research is supported by the National Science Council of the Republic of China under grant number NSC-98-2221-E-035-048.

REFERENCES

[1] I.F. Akyildiz, Weilian Su, Y. Sankarasubramaniam, E. Cayirci, ”Wireless sensor networks: a survey,” *Computer Networks*, vol. 38, no.4, pp. 393-422, 2002.

[2] I.F. Akyildiz, Weilian Su, Y. Sankarasubramaniam, E. Cayirci, ”A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102-114, 2002.

[3] R. Rajagopalan, P.K. Varshney, ”Data-aggregation techniques in sensor networks: a survey,” *IEEE Communications Surveys & Tutorials*, vol. 8, no. 4, pp. 48-63, 2006.

[4] S. Lindsey, C. Raghavendra, K. M. Sivalingam, ”Data Gathering Algorithms in Sensor Networks Using Energy Metrics,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 924-935, 2002.

[5] W. R. Heinzelman, A. Chandrakasan, H. Balakrishnan, ”Energy-Efficient Communication Protocol for Wireless Microsensor Networks,” in *Proc. of the 33rd Hawaii International Conference on System Sciences*, 2000.

[6] S. Lindsey, C. S. Raghavendra, ”PEGASIS: Power-Efficient Gathering in Sensor Information Systems,” in *Proc. of IEEE Aerospace Conference*, pp. 1125-1130, 2002.

[7] H. O. Tan, I. Korpeoglu, ”Power Efficient Data Gathering and Aggregation in Wireless Sensor Networks,” ACM SIGMOD Record, vol. 32, no. 4, pp. 66-71, 2003.

[8] Jisoo Shin, Changjin Suh, ”Energy-Efficient Chain Topology in Ubiquitous Sensor Network,” in *Proc. of 10th International Conference on Advanced Communication Technology*, pp. 1688-1693, 2008.

[9] K. Du, J. Wu, D. Zhou, ”Chain-Based Protocols for Data Broadcasting and Gathering in the Sensor Networks,” in *Proc. of the 17th International Symposium on Parallel and Distributed Processing*, 2003.

[10] N. Tabassum, Y. Urano, A.K.M Ahsanul Haque, ”GSEN: An Efficient Energy Consumption Routing Scheme for Wireless Sensor Network,” in *Proc. of ICNICONSMCL*, 2006.

[11] N. Tabassum, Q.E.K.M. Mamun, Y. Urano, ”COSEN: A Chain Oriented Sensor Network for Efficient Data Collection,” in *Proc. of Third International Conference on Information Technology: New Generations*, pp. 262-267, 2006.

[12] D. Mandala, X. Du, F. Dai, C. You, ”Load balance and energy efficient data gathering in wireless sensor networks,” *Wireless Communications and Mobile Computing*, vol. 8, no. 5, pp. 645-659, 2007.

[13] S.-M. Jung, Y.-J. Han, T.-M. Chung, ”The Concentric Clustering Scheme for Efficient Energy Consumption in the PEGASIS,” in *Proc. of 9th International Conference on Advanced Communication Technology*, pp. 260-265, 2007.

[14] Lingyun Yuan, Yunlong Zhu, Tianwei Xu, ”A Multi-Layered Energy-Efficient and Delay-Reduced Chain-Based Data Gathering Protocol for Wireless Sensor Network,” in *Proc. of International Conference on Mechtronic and Embedded Systems and Applications*, pp. 13-18, 2008.

[15] K.-H. Chen, J.-M. Huang, C.-C. Hsiao, ”CHIRON: An Energy-Efficient Chain-Based Hierarchical Routing Protocol in Wireless Sensor Networks,” in *Proc. of Wireless Telecommunications Symposium*, pp. 1-5, 2009.