Location-Based Spiral Clustering Algorithm for Avoiding Inter-Cluster Collisions in WSNs

Young-Uk Yun, Jae-Kark Choi and Sang-Jo Yoo
Graduate School of Information and Telecommunications, INHA University
253 Yonghyun-dong, Nam-gu, Incheon 402-751, Korea
[e-mail: younguk.yun@samsung.com, jkc@inha.edu, sjyoo@inha.ac.kr]
*Corresponding author: Sang-Jo Yoo

Received November 27, 2010; revised February 14, 2011; revised March 31, 2011; accepted April 8, 2011; published April 29, 2011

Abstract

Wireless sensor networks (WSN) consist of a large amount of sensor nodes distributed in a certain region. Due to the limited battery power of a sensor node, lots of energy-efficient schemes have been studied. Clustering is primarily used for energy efficiency purpose. However, clustering in WSNs faces several unattained issues, such as ensuring connectivity and scheduling inter-cluster transmissions. In this paper, we propose a location-based spiral clustering (LBSC) algorithm for improving connectivity and avoiding inter-cluster collisions. It also provides reliable location aware routing paths from all cluster heads to a sink node during cluster formation. Proposed algorithm can simultaneously make clusters in four spiral directions from the center of sensor field by using the location information and residual energy level of neighbor sensor nodes. Three logical addresses are used for categorizing the clusters into four global groups and scheduling the intra- and inter-cluster transmission time for each cluster. We evaluated the performance with simulations and compared it with other algorithms.

Keywords: Wireless sensor network, clustering, energy efficiency, network lifetime

A preliminary version of this paper appeared in ICACT 2010, February 7-10, Gangwon-Do, Korea. This version includes the updated descriptions for the location-based spiral clustering algorithm and intensive simulation results. This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2010-C1090-1011-0007). This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST) (No. 2010-0000362).

DOI: 10.3837/tiis.2011.04.003

1. Introduction

Recent advances in technology have witnessed an increasing in using wireless sensor networks (WSNs) in many applications, including environmental monitoring and military field surveillance [1]. Unlike general purpose data communication networks, WSNs are typically designed for a specific domain of applications. In these applications, hundreds to thousands of low cost sensors are deployed and periodically report physical information such as temperature, pressure, humidity, light, and chemical activities. Reports from sensors are collected by observers (called sinks). Many WSN applications require only the aggregated value at certain region. In this case, sensors in different positions in a certain region can collaborate to aggregate their data and more efficiently report their information. Moreover, data aggregation reduces the communication overhead in the network, leads to meaningful energy savings. In order to support such data aggregation or network topology control, nodes can be partitioned into a number of small groups called clusters. Clustering has been considered as an effective approach for organizing the network into a connected hierarchy. Besides achieving energy efficiency, a well designed clustering mechanism can reduce packet collisions between nodes so that it results in better network throughput under high load conditions. Many localization algorithms [2][3][4] and clustering approaches [5][6][7][8][9][10][11][12][13][14][15][16][17][18][19] for WSNs have been proposed.

Clustering in WSNs causes several issues, such as ensuring connectivity and scheduling inter-cluster communications [20][21]. In this paper, we propose a location-based spiral clustering (LBSC) algorithm to handle the problems. In LBSC algorithm, we iteratively select cluster heads (CHs) located along with the four spiral directions. In this paper, we assume that a node knows the location information of its one-hop neighbor nodes (NNs) by exchange hello messages. The contributions of LBSC are as follows. First, while generating new clusters by LBSC algorithm, the node sets of four different directions can execute LBSC clustering simultaneously without any control packet collisions. Therefore we can reduce the network-wide clustering latency. Second, after completing the cluster formation, each node will be assigned in one of the four global groups and all clusters assigned to the same global group can concurrently transmit data to their respective CHs without inter-cluster collision. Third, each CH is able to select a gateway (GW) node among one-hop NNs to configure the routing path to the sink node in the process of cluster formation unlike other location-based routing protocol [22][23].

The rest of this paper is organized as follows. In Section II, the motivation of our research is introduced. Section III presents our proposed LBSC algorithm in detail. In Section IV, the simulation results are presented. Finally, Section V concludes this paper.

2. Motivation and Approach

The essential operation of clustering algorithm in WSN is efficient cluster formation. Each cluster has a coordinator, referred to as a CH, and a number of member nodes (MNs). The MNs report sensing data to the respective CHs. The CHs not only perform sensing the environments, but also collect the data from MNs and relay the aggregated data to a sink through other CHs [5][11][12] or GWs [8][9]. The execution of a clustering algorithm can be carried out at a centralized authority or in a distributed way by each node. The centralized approach [5][6][7] is not desirable for the large scale networks where collecting all the necessary information at the central authority is both time and energy consuming work. On the other hand, distributed approach [8][9][10][11][12][13][14][15][16] is more suitable for large scale networks and either iterative or probabilistic in nature. In iterative approach [8][9][10], a node waits for a specific event to occur or certain nodes to decide their role (CHs) before making a decision. The problem is that their convergence speed is dependent on the network diameter and the performance is also highly sensitive to packet collision which can lead to consume more energy and packet transmission delay. Our protocol design goal is to prevent packet collisions during the clustering procedure. In probabilistic approach [11][12][13][14][15], nodes can independently decide on its role in a clustered network. However, if some of them are not contained in any cluster because of the limited cluster range as shown in **Fig. 1**, they need a rejoin method for network connectivity with multi-hop communication to the CHs. Such a multi-hop communication consumes more energy for relaying data, and thus is unsuitable for energy efficient design. Related with this problem, the objective of our proposal is to ensure the full connectivity by using one-hop clustering method. In both approaches, a CH can set the time schedule and inform such schedule to its MNs for intra-cluster transmissions. The problem is how to prevent the inter-cluster collisions ($b_2 \rightarrow CH_B$ and $c_2 \rightarrow CH_C$) and how to support the concurrent transmission ($a_1 \rightarrow CH_A$ and $b_1 \rightarrow CH_B$) as shown in **Fig. 1**. One possible way is to use different code division multiple access (CDMA) code to avoid inter-cluster collision problem as in LEACH [11] protocol. However, a relatively expensive cost is required for the usage of CDMA. Our aim is to ensure concurrent intra-cluster communications without inter-cluster collisions.

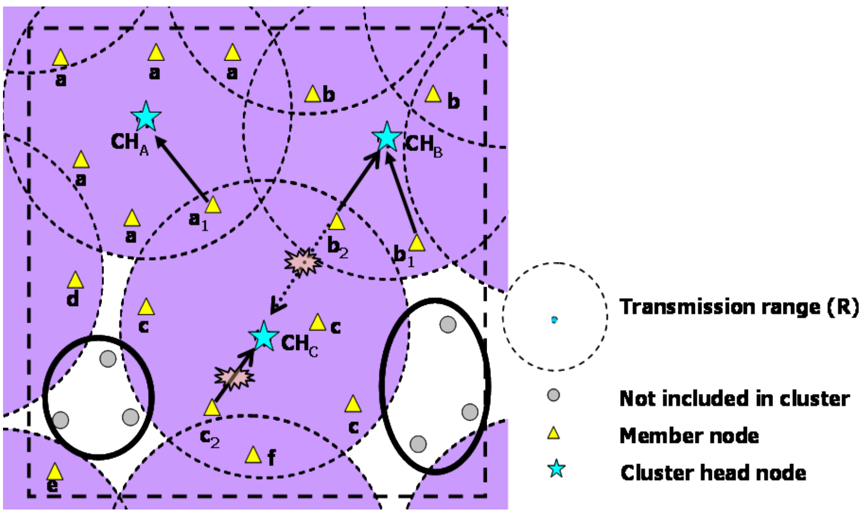


Fig. 1. General problem in clustering

For solving the above-mentioned problems, we propose a LBSC algorithm for improving connectivity and avoiding inter-cluster collisions. The basic idea is to iteratively construct clusters in four spiral directions from the center of the deployed sensor field using the location information as shown in **Fig. 2**. As shown in **Fig. 2**, the proposed distributed iterative clustering progresses to four spiral different directions from the center of the sensor field. A_1 and A_2 denote the order of clustering tier and the direction, respectively. In our approach network periodically performs new cluster formation to capture topology changes of the network and to provide fair energy consumption between nodes. Whenever the network needs the next cluster formation (at the beginning of each round), the spiral direction is moved by θ.

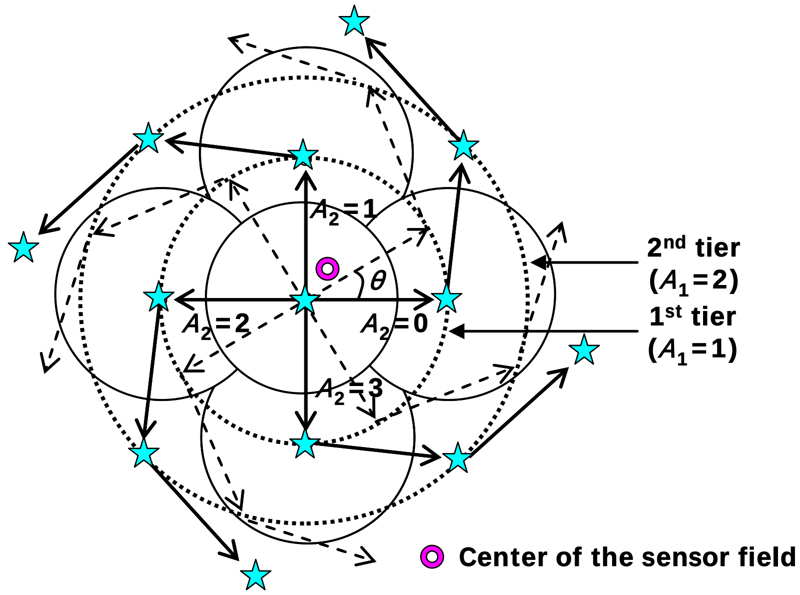


Fig. 2. Approach of LBSC algorithm

When a cluster head determines the next tier cluster head (CH_{NXT}), it is not possible that the CH selects next CH directly due to the limited transmission range. In order to select CH_{NXT} and perform LBSC algorithm with four directions properly, we define a new type of node so-called decision node (DN) which is one of the neighbor nodes of the current CH. A DN is selected by its CH, and it will decide a new CH_{NXT} in order to expand the spiral clustering. While the first CH selects (CH_{FST}) four DNs in order to make four new clusters, every other CH selects only one DN.

LBSC algorithm is to ensure that the each control packet transmitted by certain nodes at the same time have not suffered from packet collisions during the clustering procedure because destinations of those nodes are located out of the transmission range each other. Proposed algorithm also establishes the routing path from all CHs to the sink during clustering for inter-cluster communication. After successful cluster formation, each cluster is divided into each global group using proposed address and establish time schedule for concurrent intra-cluster transmission without inter-cluster packet collision. If the number of global groups is less than four, some of nodes are not contained in any cluster. However, if it is more than four, we need not only additional bits to distinguish each global group but also additional inter-cluster time schedule which can lead to packet transmission delay. Thus, we use four global groups.

As shown in **Fig. 3**, in order to distinguish four global groups, LBSC algorithm uses three logical addresses as follows. The first address, denoted by A_1, is initially set as 0 for the first CH and increases as a new cluster is generated in each spiral direction, in which A_1 represents clustering tier number from the center as shown in **Fig. 3**. The second address, denoted by A_2, distinguishes the four different spiral directions. The combination of the A_1 and A_2 distinguishes the four global groups for avoiding the inter-cluster packet collisions. The third address, denoted by A_3, is used for indexing each MN in a cluster. This address can be changed at every round by clustering procedure.

In order to balance the energy consumption between the CHs (including CH_{FST}), the cluster range of the CH which is closer to the center should be smaller than the others. Thus, our proposed algorithm adaptively calculates the proper cluster range (CR) of each CH as follows:

$$CR_{A_1} = \min(R, \alpha \times R + \beta \times A_1 \times R) \quad (1)$$

where R represents the predefined maximum transmission range and A_1 is the first address of a node. α and β are scaling parameters. As a cluster head is departing from the center, its cluster range is increasing up to the maximum value.

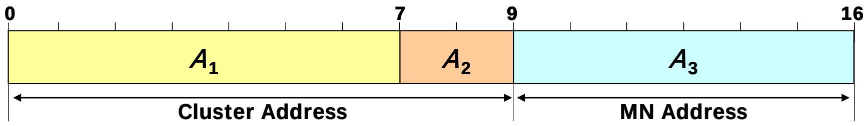


Fig. 3. Proposed address distinguishes four global group

In order to select proper CH, DN, and GW, we use $\mathbf{NN}_k$ as the set of neighbor nodes (NNs) of the node k and $\mathbf{MN}_k$ as the set of member nodes (MNs) of the node k. If a node is located within the CR from CH, it becomes a MN of CH. Note that $\mathbf{MN}_{CH} \subset \mathbf{NN}_{CH}$. In the following, the node that is nearest to a basis position BP among the node set $\mathbf{S}$ is derived as (2).

$$j^* = \arg\min_{\forall j \in S} \left(\sqrt{(BP.x - j.x)^2 + (BP.y - j.y)^2} \right) \quad (2)$$

where j represents a node contained in set $\mathbf{S}$; ($j.x$, $j.y$) represents the position of node j.

Rotating the role of CHs among sensor nodes is essential so as not to lay a burden on a few nodes with more duties than others. CH_{FST} can be replaced by its neighbors in every round and we can select different CH_{NXT} by using different degree (θ) as shown in **Fig. 2**. The rotation of cluster heads in our method can facilitate the evenly distributed energy consumption in the network, even if the node density varies from place to place.

3. Location-Based Spiral Clustering

In WSN, nodes are generally allowed to sleep when they are not sensing or communicating. Therefore each node can avoid idle listening that consumes similar energy to transmitting or receiving. We assume that all nodes are fixed and have the same transmission range R. We also assume that each node knows the location information of its one hop NNs by using the localization algorithm. In order to prolong the network lifetime, rotating the role of CHs is basically handled by a predefined timer.

3.1 Time Scheduling

As shown in **Fig. 4**, in every round, the proposed clustering method is performed at first, and then data transmission and sleep mode are followed repeatedly. As shown in **Fig. 5** and **Fig. 6**, after node deployment, the closest node from the center of the sensor field announced by the sink node becomes the CH_{FST}. Then it selects the four decision nodes (DNs) to determine the CH_{NXT} in four different directions. And each CH selects a DN simultaneously. This step will be performed recursively until all nodes are included in clusters.

After cluster formation is successfully performed using LBSC algorithm, we can distinguish four global groups and establish time schedule for concurrent data transmissions based on cluster addresses denoted by $A_1.A_2$=*odd.odd*, $A_1.A_2$=*odd.even*, $A_1.A_2$=*even.odd*, and $A_1.A_2$=*even.even* as shown in **Fig. 6**. Then clusters of the same group can simultaneously transmit the data to the respective CHs without inter-cluster collisions. In **Fig. 6** the shadowed clusters represent "$A_1.A_2$=*odd.even*" group and they can use the same time interval for data transmission.

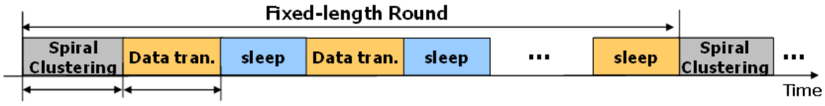


Fig. 4. Super frame structure

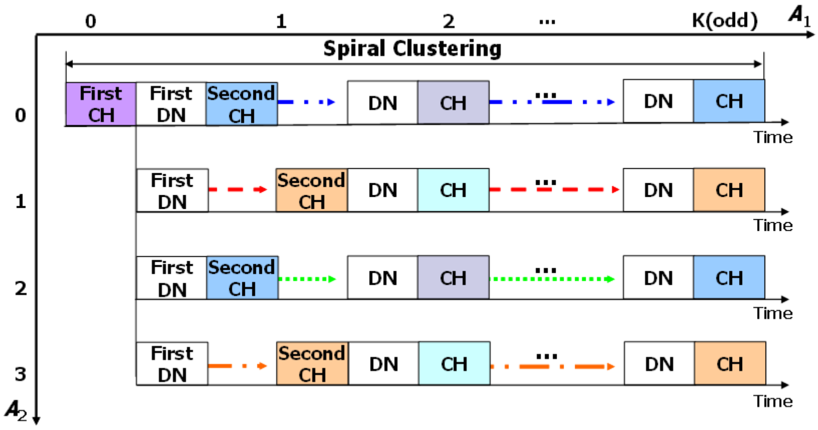


Fig. 5. Location based spiral algorithm

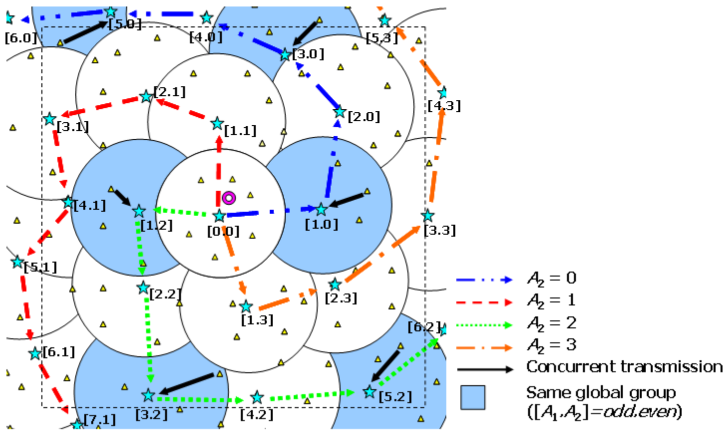


Fig. 6. Each global group concurrently transmits data from MN to respective CH without inter-cluster collision ([$A_1.A_2$]=*odd.even*)

For sensing data transmission to the sink node, each node sends its data to the respective CH. Then the CH aggregates the data and sends it to the sink through gateway nodes (GWs). Data transmission with inter-cluster coordination can be implemented as two different manners as shown in **Fig. 7**. First, the cluster time schedule is divided into four different time period for four global groups. Second, each global group uses different frequency channel for simultaneous data transmission. For both cases, the time period for single cluster is also partitioned to the mini slot assigned to each member node with address A_3 category.

3.2 Selection of the First Cluster Head

The first CH should be located around the center of the sensor field for successful cluster formation. At the beginning of the proposed clustering algorithm, the sink determines the center position of the sensor field, and then broadcasts a control message, which is carrying the center position and a predetermined range of the possible first CH from the center. If a node is inside the predetermined range of the center and it does not have any one hop neighbor that is closer to the center, then it will perform the role as the first CH (CH_{FST}) after random back-off time. CH_{FST} can be replaced by its neighbors in every round for fair energy consumption.

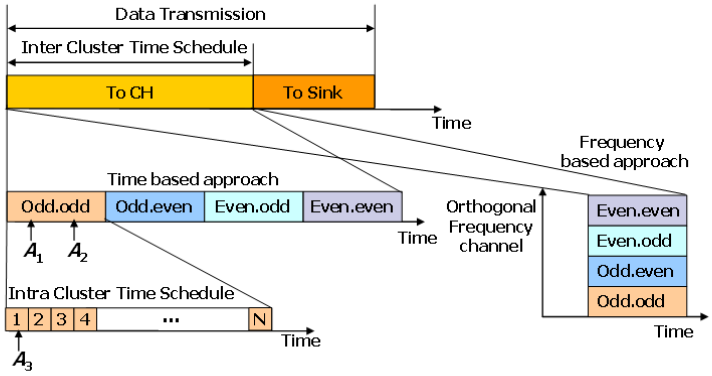


Fig. 7. Data transmission with proposed address

3.3 Selection of the First Tier Decision Nodes

The first work of CH_{FST} is to determine four decision nodes (DNs) located in four different directions. Each DN is responsible for selecting the new CH_{NXT} in each direction. CH_{FST} defines four coordinates, P_i (i = 0, 1, 2, 3), as the basis position of each direction as follows:

$$
\begin{aligned}
P_{i.x} &= CH_{FST}.x + R\cos(90°\cdot i + \theta) \\
P_{i.y} &= CH_{FST}.y + R\sin(90°\cdot i + \theta) \\
\theta &= \omega \times round / C_{center}
\end{aligned} \tag{3}
$$

where C_{center} is the number of neighbor nodes within predetermined range from center of the sensor field. We can select different next CH by using θ. In this paper, we use ω=25° . Therefore, at each cluster selection round, cluster selection base point is rotating by θ.

CH_{FST} can select the four DNs nearest to the respective P_i among its NNs as in (2) where **S** is $\mathbf{NN}_{CHFST}$ and BP is P_i. A_2 address of each selected DN follows the index i of each base position P_i as shown in **Fig. 8**.

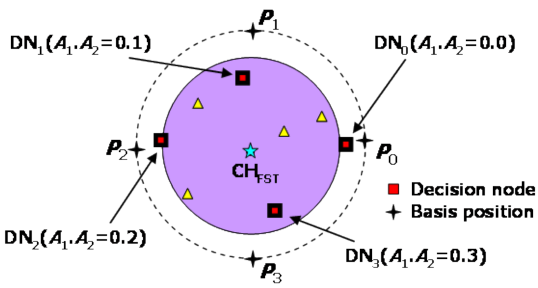


Fig. 8. Method for selecting the first decision node

3.4 Selection of the Second Tier Cluster Heads

As shown in **Fig. 9**, in order to select the second tier cluster heads, CH_{SND}, in four directions, each DN defines new coordinate, P_i (i=0,1,2,3), as the basis position for selecting second clusters.

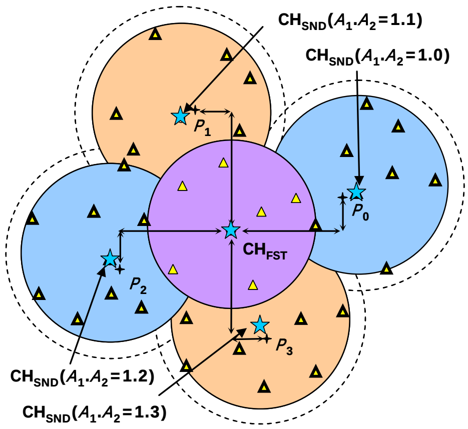


Fig. 9. Method of selecting the second tier cluster heads

$$P_i.x = CH_{FST}.x + \frac{4}{3} CR_{A_1} \cos(90° \cdot i + \theta)$$
$$P_i.y = CH_{FST}.y + \frac{4}{3} CR_{A_1} \sin(90° \cdot i + \theta) \quad (4)$$

The reason of calculating P_i is to avoid a hole which includes certain nodes that are not contained in any cluster. Then DN can select the CH_{SND} nearest to the respective P_i among the set $\mathbf{S}=\mathbf{NN}_{DN} - \mathbf{MN}_{CH\_FST}$.

To select the nearest node for second tier cluster, the residual energy of each node is considered. If we can know the residual energy of one hop NNs, then we can select the node j^* for CH as in (5).

$$j^* = \underset{\forall j \in S}{\arg\min}\left(\sqrt{(P.x - j.x)^2 + (P.y - j.y)^2} + \gamma \times R \times E_{\max} / E_{remain} \right) \quad (5)$$

where E_{remain} and $E_{\max}$ represent the remained energy and initial energy levels of each sensor node, respectively. In this paper, we use γ=0.5. In order to avoid control packet collision, DNs with A_2=(0 or 2) select CH_{SND} at first and then DNs with A_2=(1 or 3) select CH_{SND} next as shown in **Fig. 9**.

3.5 Selection of Decision Nodes for $A_1 > 0$

After selecting four CH_{SND}s, each CH_{SND} node concurrently repeats selecting DNs and CHs until no more nodes will be found in four direction. Method for selecting the DN is divided into two procedures as follow.

A. Calculating Intersection Points

As shown in **Fig. 10**, each CH_{CUR} (current cluster head) can find two intersection points of the line that crosses the line between CH_{CUR} and CH_{FST} at a right angle and the transmission range circle from CH_{CUR}. Each CH can calculate two intersection points $\{(x_1, y_1), (x_2, y_2)\}$ by using a quadratic formula. Then select one point P from two intersection points as in (6). The reason of calculating P is to select proper DN.

$$\begin{array}{l} if\ \ (CH_{CUR}.x > CH_{FST}.y) \\ \quad if\ \ (y_1 - slope(x_1 - CH_{FST}.x) - CH_{FST}.y > 0) \\ \quad\quad P.x = x_1, P.y = y_1 \\ \quad else\ \ P.x = x_2, P.y = y_2 \\ else\ \ if\ \ (y_1 - slope(x_1 - CH_{FST}.x) - CH_{FST}.y < 0) \\ \quad\quad P.x = x_1, P.y = y_1 \\ \quad else\ \ P.x = x_2, P.y = y_2 \end{array} \tag{6}$$

where *slope* is $(CH_{FST}.x - CH_{CUR}.x)/(CH_{FST}.y - CH_{CUR}.y)$.

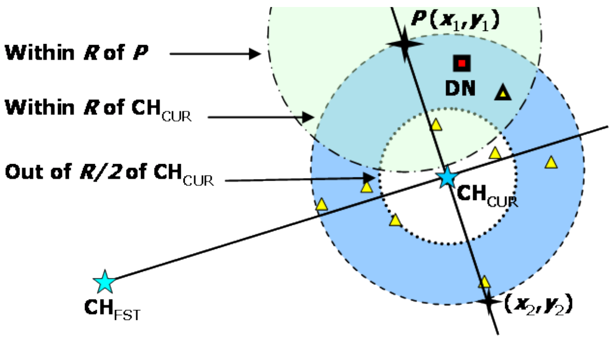


Fig. 10. Method for selecting the decision node ($A_1 > 0$)

B. Selecting Decision Nodes

As shown in **Fig. 10**, in order to select the DN, first CH_{CUR} constructs the $\mathbf{S}_{DN}$ (set of DN candidate nodes), in which to be a DN candidate node, any node N of $\mathbf{NN}_{CHCUR}$ should satisfy the following two conditions (7) and (8).

$$\sqrt{(CH_{CUR}.x - N.x)^2 + (CH_{CUR}.y - N.y)^2} > \frac{1}{2}R \tag{7}$$

$$\sqrt{(P.x - N.x)^2 + (P.y - N.y)^2} < R \tag{8}$$

Then CH_{CUR} can select the DN nearest to the respective P among the set as in (2) where **S** is $\mathbf{S}_{DN}$ and BP is P. If selected DN does not find CH_{NXT} appropriately, it sends a message to the CH_{CUR} to select another DN except it.

3.6 Selection of Cluster Heads for $A_1 > 1$

As shown in **Fig. 11**, the DN selects the CH_{NXT} with the location information of the NNs, CH_{CUR}, and CH_{FST}. In order to select the CH_{NXT}, DN constructs the set $S_{CH} = \mathbf{NN}_{DN} - \mathbf{MN}_{CHCUR}$.

3.7 Selection of Gateway Nodes

In WSN, all information sensed by sensor nodes only need to be sent to remote sink unlike ad hoc network that needs to communicate among the nodes for information exchange. As shown in **Fig. 12**, the proposed algorithm also establishes routing paths from all CHs to the sink during clustering. Each CH selects a gateway node with the location information of the $\mathbf{NN}_{CH}$, CH_{CUR}, and the sink. Each CH can select a GW node nearest to the sink using (2) where **S** is $\mathbf{NN}_{CH}$ and BP is the sink node position. Selection of gateway node based on the residual energy level can be helpful for the more evenly distributed energy consumption. In LBSC, as addressed in Section 2, the rotation of cluster heads supports the evenly distributed energy consumption among the sensor nodes. Since the cluster head selects its gateway node, the different cluster heads in different rounds select the different gateway nodes. Therefore, the gateway nodes are also evenly distributed.

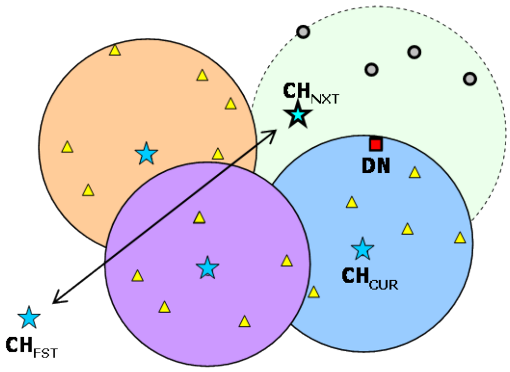


Fig. 11. Method for selecting the cluster head ($A_1 > 1$)

4. Simulation Results

In this section, we compare the performance of LBSC with other works such as LEACH[5], HEED[12], and M-LEACH[17]. For this simulation study, we implemented our LBSC and other algorithms in C++ programming language and winAPI (for the captured images from **Fig. 13** to **Fig. 15**). For the experiments, we consider a sensor network with randomly distributed sensor nodes. The area of network field and the number of sensor nodes in the field can be varied in experiments. The parameters used in simulation study are summarized in **Table 1**.

4.1 Connectivity and Inter-Cluster Packet Collision

From **Fig. 13** to **Fig. 15**, the network field size is 900m×900m and the number of nodes in the network is set as 200. The transmission range of a node is set as 150m. **Fig. 13** shows that every node in the sensor network is included in one of the clusters generated by LBSC algorithm. We can also see that all the clusters are successfully categorized into one of four global groups. Moreover, we can see that the LBSC algorithm is well formed in spiral directions by following the four colored arrows in **Fig. 13**. It should be noted that, in original LEACH protocol, there is no limitation for the transmission range, which is impractical for the battery-powered sensor networks. By using the upper-bounded transmission range, we can observe that a number of nodes were not able to join any cluster after clustering procedure in LEACH protocol, because the limited transmission range of a sensor node and the probability-based self-election for cluster head selection as shown in **Fig. 14**.

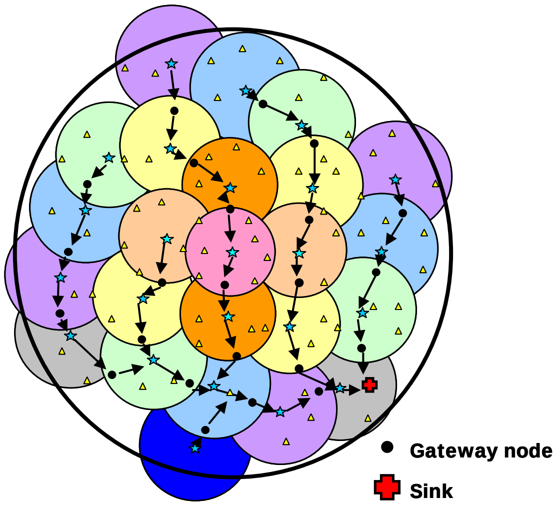


Fig. 12. Method for selecting the gateway

Table 1. Simulation parameters

Parameter	Value
Node type	Static node
Transmission range of a node (*R*)	25m ~ 150m
α, β	0.8, 0.05
γ	0.5
ω	25˚

The result of successful cluster formation using LBSC, each set of clusters are engaged in one of the four global groups as shown in **Fig. 15**. Thus each set of clusters engaged in a global group can concurrently perform the packet transmission without inter-cluster packet collision using the inter-cluster time schedule.

In **Fig. 16**, we simulated relatively long time (100 rounds) to evaluate performance. In case of LEACH, we vary the expected number of clusters from 10 to 50. Note that in LEACH, the expected number of clusters is a design parameter, which should be set before sensor node deployment. Unlike LEACH, in LBSC appropriate numbers of clusters are automatically generated in accordance with node's transmission rage. In **Fig. 16-(a)**, the generated number of clusters in LBSC is about 30 for every round. However in LEACH, even though the expected number of clusters was set to 30, the actually generated clusters are about 40 because LEACH algorithm requires the cluster re-forming procedure for the nodes that fail to participate in any cluster as shown in **Fig. 16-(b)**. It represents that LEACH requires additional packet exchanges, and thus may have the lengthy network-wide clustering latency.

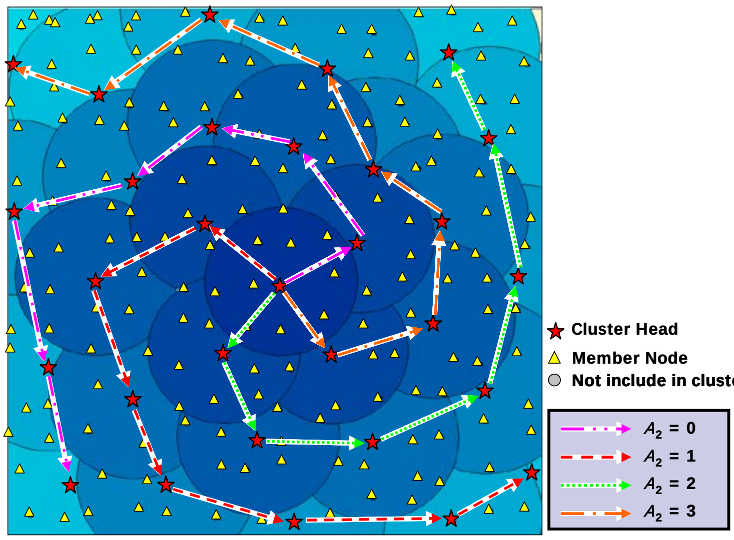


Fig. 13. After cluster formation using LBSC

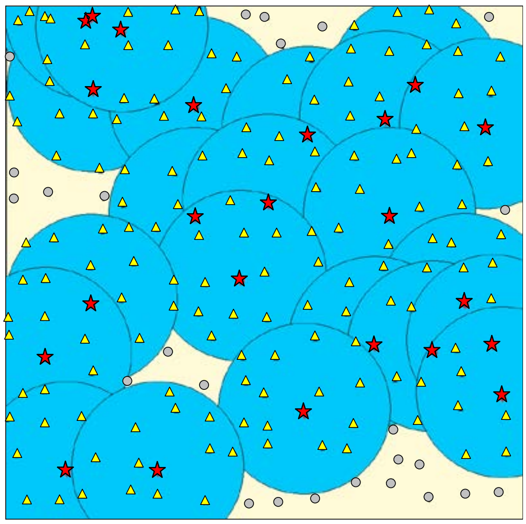
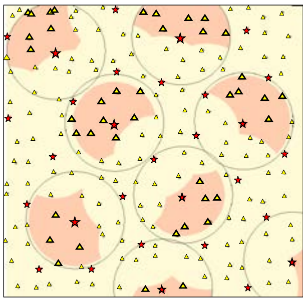
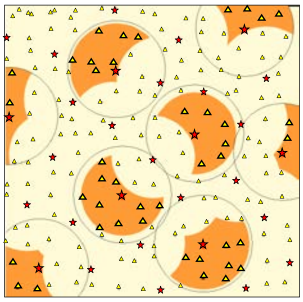
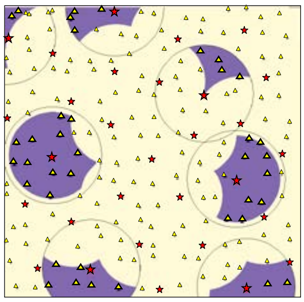


Fig. 14. After cluster formation using LEACH

Fig. 16-(c) shows the required inter-cluster time slot separations (or the required frequency channels) to avoid inter-cluster collisions. LBSC always needs four inter-cluster time slots whereas LEACH needs much more time slots. Moreover, in LEACH to avoid inter-cluster collisions, neighboring CHs need to negotiate their time slots to prohibit using the same time slot. LEACH requires complex computation work to guarantee that in the entire sensor network there is no inter-cluster collision. LBSC generally has the stable number of nodes in a cluster than that of LEACH as shown in **Fig. 16-(d)**.


(a) $A_1.A_2 = odd.odd$


(b) $A_1.A_2 = odd.even$


(a) $A_1.A_2 = even.odd$

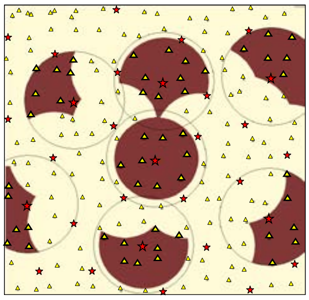
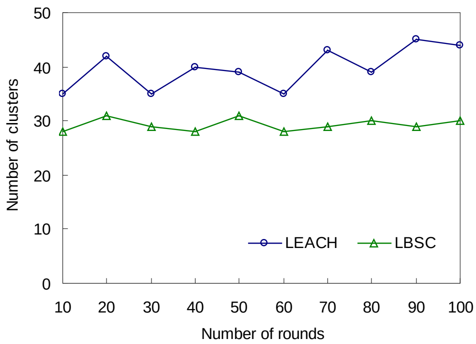
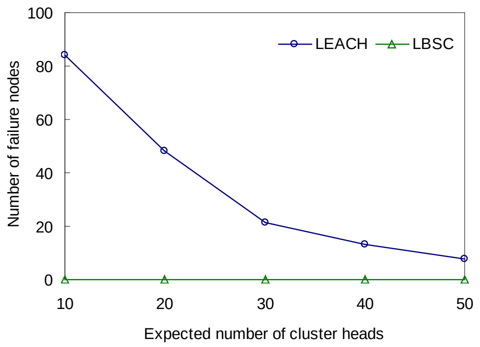

(b) $A_1.A_2 = even.even$

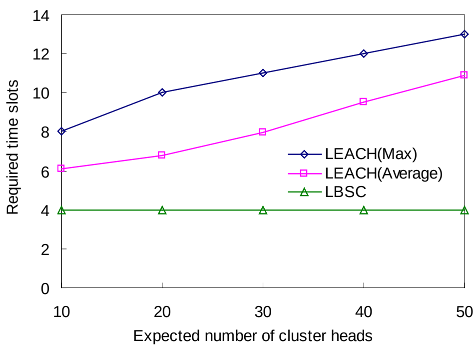
Fig. 15. Concurrent packet transmission without inter-cluster packet collision



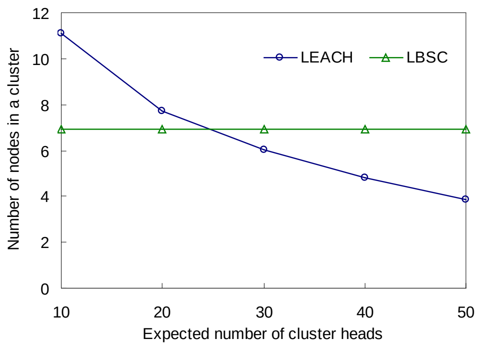
(a) Number of clusters



(b) Number of clustering failure nodes



(c) Number of required time slots

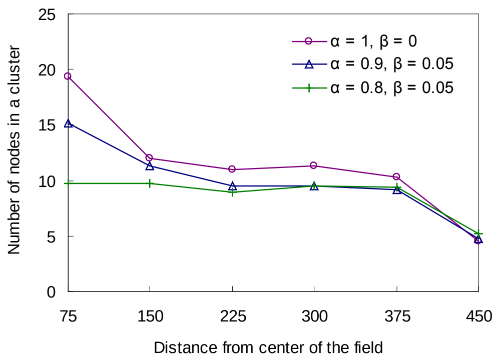


(d) Average number of nodes in a cluster

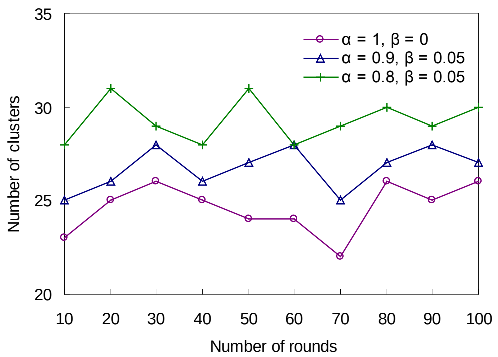
Fig. 16. Performance comparison with LEACH

4.2 Cluster Range

In order to balance the energy consumption between the CHs, we calculate the proper cluster range of each CH as specified in Section 2. To manage the energy consumption of all nodes evenly, in the proposed LBSC the cluster range is slightly increasing as a cluster is going away from the center. In **Fig. 17**, we compared the number of nodes in a cluster and the required number of clusters at different rounds for three different set of (α, β). (α=1, β=0) means the cluster range is the same for all clusters. As we can see when the cluster range is not adaptive to the location of cluster head (i.e., α=1, β=0), the first CH located around the center of the sensor field has larger member nodes than the others as shown in **Fig. 17-(a)**. This is because the area of the first cluster fully covers the radius of the transmission range of the first CH, while the subsequent clusters covers only part of the transmission range of the respective CH as shown in **Fig. 15**. For this reason, it makes the difference of number of member nodes in a cluster, and thus some nodes located around the center of the sensor field consumed more energy than other nodes. Thus in the proposed LBSC, the CR of a CH which is closer to the center should be smaller than the others in order to ensure that CHs are not overburdened. **Fig. 17-(b)** shows variation of the number of clusters. As α is decreasing, the number of clusters to cover the entire field increases. However α is too small, it can make a hole including certain nodes that are not contained in any cluster. There is a tradeoff between even energy consumption and number of clusters (i.e., cluster size).



(a) Average number of nodes in a cluster from center of the field



(b) Number of clusters

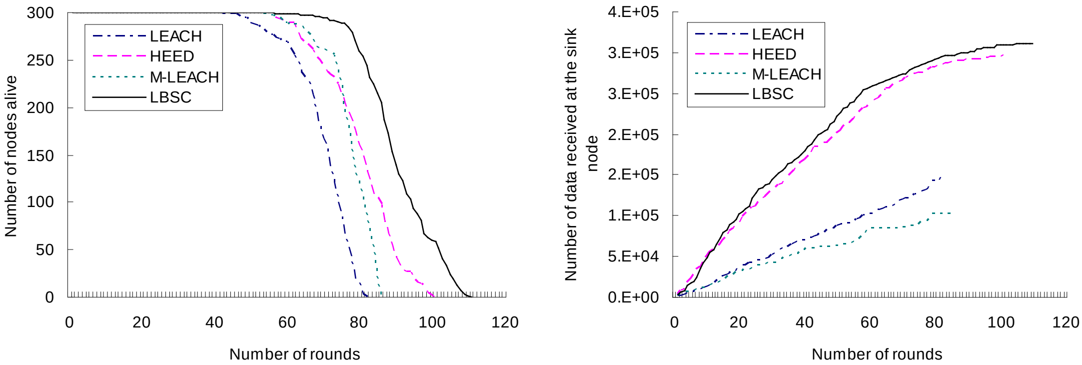
Fig. 17. Influence of cluster range

4.3 Energy Efficiency

In this section, we investigate the energy efficiency of LBSC. We refer to [5] to set the parameters for energy model, as shown in Table 2. In **Fig. 18**, the simulated network consists of 300 nodes randomly distributed within 100m×100m sensor field. The length of a round is set to 5 seconds and the bandwidth of the channel is set to 1Mbps. Each packet consists of data message of 475 bytes long and packet header of 25 bytes long. The cluster heads are in charge of receiving data from their member nodes, and relay the aggregated data to the respective gateway nodes during a round. By assuming all nodes are continuously sending reports [12], the maximum number of packets received by a cluster head during the predetermined length of a round can be determined. The network lifetime is relatively less affected by the transmission data size, but the larger the data size, the less data signals are received by sink node. In cases of HEED and our LBSC algorithms, transmission range of a node is set to 25m, but is assumed to be adjustable when there is no relay node alive between a distant cluster head and sink node. On the other hand, in LEACH and M-LEACH, the transmission range of a node is assumed to be adjustable based on the distance between the sender and receiver. The expected number of cluster heads is set to be 5% of deployed number of nodes in LEACH, M-LEACH, and HEED. It should be noted that, in any algorithm, the clustering phase is relatively very short compared to the time for a round (e.g., 5 seconds with 1Mbps). Thus, the network lifetime is mainly affected by the intra-cluster communications and aggregated data forwarding at each CH. In LEACH and M-LEACH, the elected CHs transmit the aggregated data to sink node directly so that relatively large energy consumption is expected (i.e., distance between each CH and sink node is likely to be larger than reference distance d_0=75m, and thus ε_{mp} is used). On the other hand, HEED and LEACH use multi-hop transmission using the gateway nodes (we assume the gateway nodes in HEED are appropriately determined similar to our LBSC), which results in relatively small energy consumption by CHs when they forward the data. As shown in **Fig. 18**, our LBSC algorithm outperforms the others in terms of network lifetime and total amount of data received at sink node. LBSC is more stable than HEED, since it triggers only required control messages and achieves the distribution of cluster head geographically, while, in HEED, the messages occur several times at each node during clustering phase.

Table 2. Simulation parameters for energy model

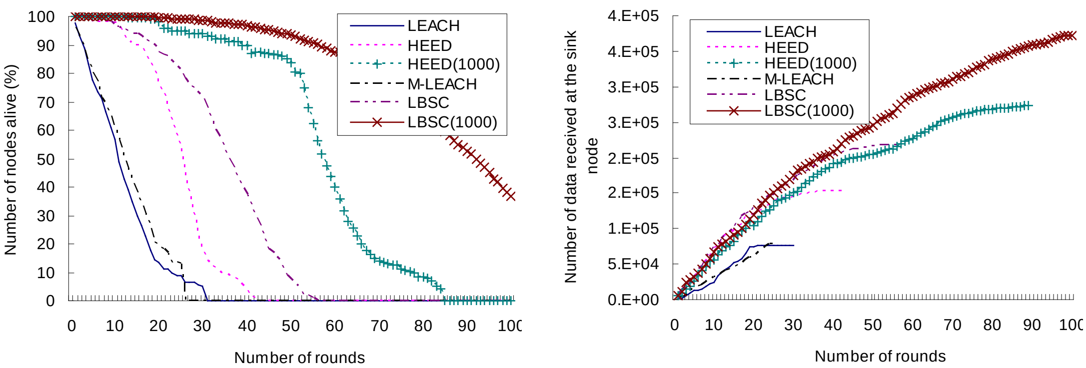
Parameter	Value
The initial energy of sensor node	2 J
Electronics energy ($E_{elec.}$)	50 nJ/bit
Free space transmit amplifier (ε_{fs})	10 pJ/bit/m^2
Multipath transmit amplifier (ε_{mp})	0.0013 pJ/bit/m^4
Energy for data aggregation	5 nJ/bit/signal
Reference distance (d_0)	75m



(a) Number of nodes alive over time **(b)** Number of data received at sink node

Fig. 18. Performance for the small network

Fig. 19 shows similar results with those in **Fig. 18**. The field size of 500m×500m is experimented. The transmission range of a node in HEED and LBSC is set to 100m. We can see that, the network lifetime and the amount of packet transmitted are reduced compared to those in **Fig. 18**. It is mainly because of the increased distance between senders and receivers. If the density of sensor nodes in the entire field increases (e.g., 1000 nodes in HEED and LBSC, HEED(1000) and LBSC(1000) in **Fig. 19**, respectively), the reduced transmission distance improves the network performance.



(a) Number of nodes alive over time **(b)** Number of data received at sink node

Fig. 19. Performance for the large network

4.4 Non-uniform Node Distribution

In LBSC, we assume that sensors are uniformly distributed in the entire sensor field. In all of the experiments presented above, evenly deployed sensor nodes are considered. The proposed method constructs clusters based on the pre-defined rule regardless the different node density so that the average number of clusters is almost constant. However, the density of sensor nodes can be different from place to place and there may be some skewed sensor nodes. For example 20% area has 40% of nodes of the entire sensor field. The higher node density results in the more nodes in a cluster. As shown in **Fig. 7**, all MNs in a cluster should transmit their data to the CH within the globally scheduled inter-cluster time period. Thus, the amount of time period allocated to each MN in a dense cluster is relatively smaller than those in the less dense clusters. This unfairness in terms of throughput is a drawback of the proposed mechanism for the non-uniform node density case.

Let A_{dense} and n_{dense} denote the percentage of dense area to the entire field and the percentage of the number of nodes in the dense area to the total number of nodes, respectively. Imbalance index denotes the ratio of n_{dense} to A_{dense} (i.e., n_{dense}/A_{dense}) and fairness index denotes the ratio of the minimum throughput in the dense area to the maximum throughput in the coarse area. In **Fig. 20**, the fairness index is investigated according to the imbalance index for the case of non-uniform node distribution, where A_{dense}=20% and the simulation environments used in subsection 4.1 are applied.

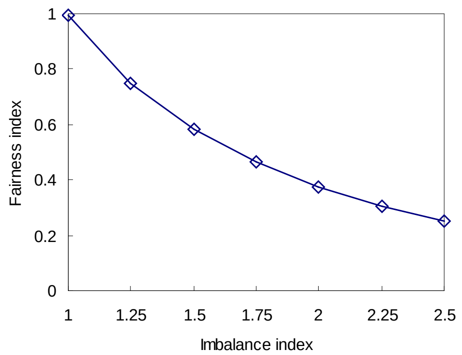


Fig. 20. Fairness index according to imbalance index

The imbalance index =1 (i.e., $A_{dense}=n_{dense}$) represents that nodes are uniformly distributed in the entire field. When the imbalance index equals 1.5, which indicates that 20% of area has 30% of nodes, the minimum throughput is about 60% of the maximum throughput. As we can see in **Fig. 20**, nodes transmit data to CHs unfairly as the network is the more non-uniformly distributed.

5. Conclusion

In this paper, we proposed a location-based spiral clustering algorithm which assures that all nodes are included in clusters without control packet collisions during the cluster formation period. It also provides the routing path from CHs to the sink. We categorize all generated clusters into the four global groups and establish time schedule by using three logical addresses. Clusters belonging to the same global groups can share the same time schedule so that we can increase network throughput and avoid data packet collisions. Simulation results show that each global group concurrently transmits data from MN to respective CH without inter-cluster collisions using four inter-cluster time schedule.

References

[1] I. F. Akyildiz, S. Weilian, Y. Sankarasubramaniam and E. Cayirci, “A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, no. 8, pp.102-114, Aug. 2002. Article (CrossRef Link)

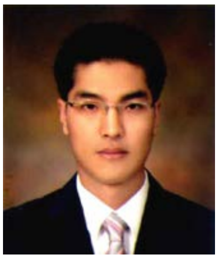
[2] T. He, C. Huang, B. M. Blum, J. A. Stankovic and T. Abdelzaher, “Range-free localization schemes for large scale sensor network,” in *Proc. of 9th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pp. 81-95, Sept. 2003. Article (CrossRef Link)

[3] D. Niculescu and B. Nath, “Ad hoc positioning system (APS) using AOA,” in *Proc. of 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM*), pp.

1734-1743, 2003. Article (CrossRef Link)
[4] G. Mao, B. Fidan and B. D. O. Anderson, "Wireless sensor network localization techniques," *Computer Networks*, vol. 51, no. 10, pp. 2529-2553, Jul. 2007. Article (CrossRef Link)
[5] W. B. Heinzelman, A. P. Chandrakasan and H. Balakrishnan, "An application specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660-670, Oct. 2002. Article (CrossRef Link)
[6] H. O. Tan and I. Korpeoglu, "Power efficient data gathering and aggregation in wireless sensor networks," *ACM SIGMOD Record*, vol. 32, no. 4, pp. 66-71, Dec. 2003. Article (CorssRef Link)
[7] S. D. Muruganathan, D. C. F. Ma, R. I. Bhasin, and A. O. Fapojuwo, "A centralized energy-efficient routing protocol for wireless sensor networks," *IEEE Communications Magazine*, vol. 43, no. 3, pp. 8-13, Mar. 2005. Article (CrossRef Link)
[8] S. Basagni, "Distributed clustering for ad hoc networks," in *Proc. of 4th International Symposium on Parallel Architectures, Algorithms, and Networks (I-SPAN)*, pp. 310-315, Jun. 1999. Article (CrossRef Link)
[9] A. D. Amis, R. Prakash, T. H. P. Vuong and D. T. Huynh, "Max-min d-cluster formation in wireless ad hoc networks," in *Proc. of 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pp. 32–41, Mar. 2000. Article (CrossRef Link)
[10] H. Chan and A. Perrig, "ACE: An emergent algorithm for highly uniform cluster formation," *Lecture Notes in Computer Science*, vol. 2920, pp. 154–171, Jan. 2004. Article (CrossRef Link)
[11] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy efficient communication protocol for wireless microsensor networks," in *Proc. of 33rd Annual Hawaii International Conference on System Sciences (HICSS)*, pp. 1-10, Jan. 2000. Article (CrossRef Link)
[12] O. Younis and S. Fahmy, "Distributed clustering in ad-hoc sensor networks: a hybrid, energy-efficient approach," in *Proc. of 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM*), pp. 629–640, Mar. 2004. Article (CrossRef Link)
[13] F. Kuhn, T. Moscibroda and R. Wattenhofer, "Initializing newly deployed ad hoc and sensor networks," in *Proc. of 10th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pp. 260–274, 2004. Article (CrossRef Link)
[14] P. Ding, J. Holliday and A. Celik, "Distributed energy-efficient hierarchical clustering for wireless sensor networks," *Lecture Notes in Computer Science*, vol. 3560, pp 322-339, Jun. 2005. Article (CrossRef Link)
[15] A. Youssef, M. Younis, M. Youssef and A. Agrawala, "Distributed formation of overlapping multi-hop clusters in wireless sensor networks," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1-6, 2006. Article (CrossRef Link)
[16] Y. Zhou, M. R. Lyu and J. Liu, "An index-based sensor-grouping mechanism for efficient field-coverage wireless sensor networks," in *Proc. of IEEE International Conference on Communications (ICC)*, pp. 2409-2415, May 2008. Article (CrossRef Link)
[17] Y. Liu, Y. Zhao and J. Gao, "A new clustering mechanism based on LEACH protocol," in *Proc. of International Joint Conference on Artificial Intelligence(IJCAI)*, pp. 715-718, , Apr. 2009. Article (CrossRef Link)
[18] Y. Liu, J. Gao, L. Zhu and X. Li, "Distributed routing for wireless sensor networks based on cluster maintenance," in *Proc. of 2nd International Conference on Future Generation Communication and Networking (FGCN)*, pp.373-376, Dec. 2008. Article (CrossRef Link)
[19] S. Lindsey, C. Raghavendra and 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, September 2002. Article (CrossRef Link)
[20] O. Younis, M. Krunz and S. Ramasubramanian, "Node clustering in wireless sensor networks: recent developments and deployment challenges," *IEEE Network*, vol. 20, no. 3, pp. 20-25, May 2006. Article (CrossRef Link)
[21] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, vol. 30, no. 14-15, pp. 2826–2841, Oct. 2007. Article (CrossRef Link)
[22] Y. Xu, J. Heidemann and D. Estrin, "Geography-informed energy conservation for ad hoc routing," in *Proc. of 7th Annual International Conference on Mobile Computing and Networking (MobiCom)*,

pp. 70-84, Jul. 2001. Article (CrossRef Link)
[23] R. Zheng, J. S. Walia, and L. Sha, "Local group communication-aware MAC protocol in wireless sensor networks," *International Journal of Wireless Information Networks*, vol. 13, no. 4, pp. 275-287, Oct. 2006. Article (CrossRef Link)

Young-Uk Yun received the B.S. degree in computer engineering and the M.S. degree in information and telecommunication engineering from Inha University, Incheon, Korea, in 2008 and 2010, respectively. He is currently an assistant engineer of Mobile Communication Division at Samsung Electronics, Suwon, Korea. His research interests include wireless sensor networks and wireless network QoS.



Jae-Kark Choi received the B.S. degree in electronic engineering and the M.S. degree in information and telecommunication engineering from Inha University, Incheon, Korea, in 2006 and 2008, respectively. He is currently working toward the Ph. D. degree at Multimedia Network Laboratory in the Graduate School of IT & T, Inha University. His research interests include wireless sensor networks, seamless mobility control and cognitive radio networks.

Sang-Jo Yoo received the B.S. degree in electronic communication engineering from Hanyang University, Seoul, Korea, in 1988 and the M.S. and Ph. D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), in 1990 and in 2000, respectively. From 1990 to 2001, he was a member of technical staff at Korea Telecom Research and Development Group, where he worked in communication protocol conformance testing and network design fields. From September 1994 to August 1995 and from January 2007 to January 2008, he was a guest researcher at National Institute Standards and Technology (NIST), USA. Since 2001, he has been with the Graduate School of Information Technology & Telecommunications, Inha University, where he is currently a professor. His current research interests include cognitive radio network protocols, seamless network mobility control, wireless network QoS, and wireless sensor networks.