Distributed clustering algorithms for data-gathering in wireless mobile sensor networks

Chuan-Ming Liua,∗, Chuan-Hsiu Llea, Li-Chun Wangb

aNational Taipei University of Technology (NTUT), Taiwan
bNational Chiao Tung University, Taiwan

Received 17 October 2005; received in revised form 27 June 2007; accepted 29 June 2007
Available online 19 July 2007

Abstract

One critical issue in wireless sensor networks is how to gather sensed information in an energy-efficient way since the energy is a scarce resource in a sensor node. Cluster-based architecture is an effective architecture for data-gathering in wireless sensor networks. However, in a mobile environment, the dynamic topology poses the challenge to design an energy-efficient data-gathering protocol. In this paper, we consider the cluster-based architecture and provide distributed clustering algorithms for mobile sensor nodes which minimize the energy dissipation for data-gathering in a wireless mobile sensor network. There are two steps in the clustering algorithm: cluster-head election step and cluster formation step. We first propose two distributed algorithms for cluster-head election. Then, by considering the impact of node mobility, we provide a mechanism to have a sensor node select a proper cluster-head to join for cluster formation. Our clustering algorithms will achieve the following three objectives: (1) there is at least one cluster-head elected, (2) the number of cluster-heads generated is uniform, and (3) all the generated clusters have the same cluster size. Last, we validate our algorithms through an extensive experimental analysis with Random Walk Mobility (RWM) model, Random Direction Mobility (RDM) model, and a Simple Mobility (SM) model as well as present our findings. © 2007 Elsevier Inc. All rights reserved.

Keywords: Wireless sensor networks; Mobility; Clustering; Data-gathering; Energy efficiency

1. Introduction

Wireless sensor networks have attracted much research attention in recent years and can be used in many different applications, including battlefield surveillance, machine failure diagnosis, biological detection, inventory tracking, home security, smart spaces, environmental monitoring, and so on [1,4]. A wireless sensor network consists of a large number of tiny, low-power, cheap sensor nodes having sensing, data processing, and wireless communication components. It has not only the ability to sense some phenomena in the interested region but also the network features, thereby representing an improvement over the traditional sensor systems.

The sensor nodes in a wireless sensor network are usually deployed randomly inside the region of interest or close to it. A remote base station (BS) connected to the Internet is engaged to give commands to all the sensor nodes and gather information from the sensor nodes. In addition to sensing, the wireless sensor nodes can process the acquired information, transmit messages to the BS, and communicate to each other. An architecture of the wireless sensor network is depicted in Fig. 1.

Since the sensor nodes are randomly scattered in the sensor field, a wireless sensor network should have the capability of self-organizing. This indicates that the wireless sensor nodes are more autonomous than the traditional sensors. Hence, it becomes more challenging to design related protocols or algorithms in such a distributed environment. Besides, wireless sensor nodes have many limitations, including modest processing power, little storage, short communication range, and limited power source. These limitations also make designing wireless sensor network protocols difficult. Because wireless sensor nodes are low-powered, the constraint on the power consumption is an important issue when designing wireless sensor network protocols. Furthermore, the node mobility imposes more

* Corresponding author. Fax: +886 2 87732945.
E-mail address: cmliu@csie.ntut.edu.tw (C.-M. Liu).
among the existing protocols, the cluster-based structure promises it back to the user. To collect the sensed information from the sensor nodes and commands by broadcasting. On the other hand, the BS needs has power supplied. So, it is possible for the BS to send the information it receives and its own sensed information may also consume much power. Therefore, automated methods of combining or aggregating the data into a small set of meaningful information are required [14]. Such methods depend on the applications. In this paper, we assume that the sensor nodes can do the data-aggregation (data-fusing) and focus on how the mobile sensor nodes route sensed information to the BS.

There are two categories for the existing data-gathering protocols: hierarchy (cluster-based) protocols and non-hierarchy protocols. The non-hierarchical protocols include Flooding [12,31], Gossiping [10], Directed Diffusion [15], SPIN [12], and REAR [9]. On the other hands, the hierarchical protocols include SMECN [19], SAR [29], Data funneling [24], LEACH [14], PEGASIS [20], HIT [6], LPT [17], and ACW [30,32]. Furthermore, the communication patterns for wireless sensor networks take one of two general forms:

- Time-driven (periodical) transmissions: periodical transmissions from all the sensor nodes, and
- Event-driven transmissions: reports from only those sensor nodes that observe a specific event.

In [18], the authors pointed out that the tree-based structure is a proper architecture for collecting data in real static wireless sensor networks. A tree-based structure can be organized by many clusters with multiple layers. However, in a wireless mobile sensor network, the topology of the network changes frequently. A tree-based structure for the static environment cannot be directly applied to a mobile environment. To retain the merits of the tree-based structure, we consider a cluster-based architecture which is a tree-based architecture with two layers for data-gathering. A cluster-based (or hierarchical)
data-gathering protocol consists of a series of rounds. In each round, there are two major phases: (1) clustering and (2) message transmission phases. In Phase (1), the entire sensor network will be partitioned into different clusters. Each cluster consists of one cluster-head and a number of sensor nodes. There are usually two steps in Phase (1): cluster-head election step and cluster formation step. After all the clusters are organized, the cluster-head aggregates data from the sensor nodes in the cluster and then transmits information to the BS directly in Phase (2). Fig. 2 captures the operations in a round and presents a cluster-based mechanism. This paper focuses on Phase (1).

We observe that the mobility impacts on the performance during the clustering phase. In [21,22], the authors discussed the effects of different cluster formation mechanisms in mobile sensor networks using a SM model. In the SM model, each sensor node randomly selects a speed and direction among four pre-defined directions initially. Then, a sensor node moves according to the fixed speed and direction. If a sensor node reaches the boundary of the sensing area, it then reverses its direction with the same speed. In ad hoc networks, many mobility models have been presented [3]. In this paper, we also consider the RWM model [7] and RDM model [26]. In the RWM model, a sensor node can randomly choose a speed within a given range and a direction to move from its current location to a new location. After moving a constant time or a constant distance traveled, each sensor node will calculate a new speed and direction to modify its current position. In the RDM model, each sensor node has a constant speed and can randomly choose a direction to travel. A sensor node can choose a new angular direction (between 0° and 180°), if the boundary of the sensing area is reached.

Although many cluster-based routing protocols have been proposed [13,20,30], the problem about clustering the sensor nodes is realized and discussed in recent years [32,34,35]. A better clustering algorithm may average the workload on each sensor node and results in a longer system lifetime. If the number of cluster-heads generated in each round changes dramatically, some sensor nodes will use up all of their energy quickly due to the heavy workload. Therefore, the system lifetime is deteriorated. If there is no cluster-head elected in a round, each sensor node will communicate directly to the BS and consumes a large amount of energy. The other factor that impacts the energy consumption is the cluster size. The cluster size is the number of sensor nodes in a cluster. If the difference in the cluster size is big among all the generated clusters, the cluster-head having large cluster size will consume a larger amount of energy than the others. As a result, the system lifetime becomes short. In summary, a good clustering algorithm should achieve the following three objectives:

1. there is at least one cluster-head elected in each round,
2. the number of cluster-heads generated in each round should be uniform, and
3. all the generated clusters should have the same cluster sizes.

3. Problems

In this paper, we consider the wireless mobile sensor networks where:

- The BS is fixed and located far away from the sensor nodes.
- All the sensor nodes are mobile, homogeneous, and power limited.
- Each sensor node is equipped with a Location Finding System [2,27].
- All sensor nodes are time-synchronized [4].

Fig. 3 shows the components of a mobile sensor node considered in this paper.

Recall that there are two steps in the clustering phase: one step is to elect the cluster-heads and the following step is to form the clusters. In this paper, we first propose two efficiently distributed algorithms to elect the cluster-heads among mobile sensor nodes in terms of the energy consumption. The proposed algorithms will achieve the objectives mentioned in Section 2. We then provide a mechanism to form the clusters among the mobile sensor nodes. If a sensor node neglects its mobility when deciding which cluster-head to join, it may select a cluster-head with a longer distance to it, thereby consuming more power.

The rule for deciding the cluster-heads in LEACH provides a distributed way to elect cluster-heads and can be fully applied to a distributed environment like sensor networks. However, when applying such a rule, the difference in the number of the cluster-heads between the rounds may be big and it also happens that there is no cluster-head at all in some rounds. The cluster-head election strategy in LEACH uses the following threshold to be the probability of being a cluster-head for each sensor node:

\[
T_v = \begin{cases} 
\frac{P}{1 - P \left( r \mod \frac{1}{P} \right)} & \text{if } v \in G, \\
0 & \text{otherwise},
\end{cases}
\]  

(1)

where \( r \) is the number of rounds that have passed, \( P \) the desired percentage of cluster-heads, and \( G \) the set of sensor nodes not
being a cluster-head yet. From Eq. (1), a sensor node is elected to be a cluster-head according to the accumulative times of not being a cluster-head.

The above threshold provides a good mechanism for each sensor node to determine to be a cluster-head or not independently. In each round, the threshold is independent from the actual number of cluster-heads elected in the previous rounds. However, it is not necessary to exactly have the desired percentage of cluster-heads in every round. Hence, the above threshold is an ideal condition. For example, consider that there are 100 sensor nodes and the number of cluster-heads elected in each round is expected to be five. The optimum number of cluster-heads that should be elected in a wireless sensor network had been discussed in [11,14]. We will not discuss this here. Suppose that, in the first round, there are only three cluster-heads elected. Then, there are 97 sensor nodes left in the second round. By applying (1), the expected number of cluster-heads elected is therefore greater than 5. Besides, we can further derive the probability of no cluster-head elected in a round as

\[ P_f(r) = \left( 1 - \frac{P}{1 - P \left( \frac{1}{1 + \frac{1}{P}} \right)} \right)^{|G|} \]  

By Eq. (2), we can conclude that, using such a threshold, there can be no cluster-head elected in some rounds and the number of cluster-heads elected in a round is related to the total number of cluster-heads elected in the previous rounds. In summary, the mechanism (i.e. Eq. (1)) used in LEACH for clustering the sensor nodes fails to achieve the objectives discussed in the previous section. In Section 4, we provide and discuss different distributed algorithms for cluster-head election which can achieve the mentioned objectives simultaneously.

In the experiments, we measure the energy efficiency by the system lifetime in terms of the number of rounds which a system experiences. We use the radio model used in [14] where the radio dissipates \( E_{\text{elec}} = 50 \text{nJ/bit} \) to run the transmitter or receiver circuitry and \( \varepsilon_{\text{amp}} = 100 \text{pJ/bit/m}^2 \) for the transmitter amplifier. The radios have power control and can expend the minimum required energy to reach the intended recipients. We also assume an \( r^2 \) energy loss due to channel transmission [25]. Using this radio model, the transmission cost and receiving cost for a \( l \)-bit message with a distance \( r \) are

- Transmission: \( E_{\text{Tx}}(l, r) = E_{\text{elec}} \times l + \varepsilon_{\text{amp}} \times l \times r^2 \), and
- Receiving: \( E_{\text{Rx}}(l, r) = E_{\text{elec}} \times l \).

Receiving data is also a high cost operation; therefore, the number of receptions and transmissions should be minimized to reduce the energy cost of an application. We use the data-fusion model discussed in [14] where the processor dissipates \( E_D = 5 \text{nJ/bit} \) to fuse the gathering data.

4. Algorithms for cluster-head election

In this section, we consider the first step in the clustering phase. Two simple but effective distributed algorithms for cluster-head election are proposed. Both of them can guarantee that there is at least one cluster-head elected and the number of cluster-heads in each round is equal. The basic idea of the first algorithm is to determine the cluster-heads by counting. The other distributed algorithm determines the cluster-heads by location. Recall that, in [11,14], the optimal number of cluster-heads has been discussed and suggested. We refer to that conclusion and will not discuss this further in this paper. The simulation in Section 6 will show that the proposed algorithms can achieve the three objectives mentioned in Section 2 and outperform the approach used in LEACH in terms of energy efficiency.

4.1. Cluster-head election by counting

This subsection describes the distributed algorithm of cluster-head election by counting. Algorithm ACE-C. Suppose that the number of sensor nodes in a mobile sensor network is \( N \) and we number the sensor nodes from 0 to \( N - 1 \). Each sensor node hence can use the assigned number as an unique identifier (ID) in the sensor network. We assume that there are \( C \) clusters in each round. Using the ID’s, algorithm ACE-C elects the sensor nodes as the cluster-heads in a round-robin fashion. In other words, in the first round, the sensor nodes with ID’s from 0 to \( C - 1 \) are the cluster-heads. In the second round, the sensor nodes with ID’s from \( C \) to \( 2C - 1 \) are the cluster-heads. The algorithm continues for the following rounds. After \( \frac{N}{C} \) rounds, each sensor node has been the cluster-head once, and the whole process starts over from the sensor node with ID \( 0 \).

To adapt this approach in a distributed way and allow each sensor node to decide independently whether it is a cluster-head in the current round, a sensor node will keep the total number of cluster-heads generated so far in each round and use its ID to decide. If it is the time for it to be one of the cluster-heads in the current round, it will claim that it is a cluster-head to all the other sensor nodes. Otherwise, it will be a non-cluster-head node. The algorithm ACE-C executed at each sensor node \( v \) hence can be stated as follows. Each sensor node \( v \) will copy

![Diagram](https://via.placeholder.com/150)
Algorithm of Cluster-head Election by Counting

Input: C: the number of cluster-heads in a round; N: the total number of sensor nodes

(1) \( N_{CH} = 0 \) /* number of cluster-heads */
(2) while \( N_{CH} < C \) do
   (2.1) \( v_{id} \) = \((vid+1) \mod N \)
   (2.2) if \((v_{id} = 0)\) then
      (2.2.1) \( v \) is a cluster-head
      (2.2.2) broadcast an advertisement message
      (2.2.3) increase \( N_{CH} \) by 1
   endif
   (2.3) wait a period of time to receive the advertisement message
   (2.4) if (a message is received) then
      (2.4.1) increase \( N_{CH} \) by 1
   endif
   endwhile
End

Fig. 4. Algorithm ACE-C used in each sensor node \( v \) for cluster-head election in each round.

its ID to a variable \( v_{id} \). This variable is used to determine if it is the turn for a sensor node to be a cluster-head. Initially, sensor node \( v \) sets the total number of cluster-heads, \( N_{CH} \), to be 0. Then, it repeatedly executes the following steps. The \( v_{id} \) is first increased by one. The algorithm then considers the remainder \( v_{id} \) divided by \( N \). There are two cases:

- When \( v_{id} = 0 \), this case indicates that it is the turn for sensor node \( v \) to be a cluster-head. Sensor node \( v \) then broadcasts an advertisement message to all the other sensor nodes.
- When \( v_{id} \) is not 0, this case denotes that sensor node \( v \) is not a cluster-head. It will wait a period of time for receiving an advertisement message.

After the decision, sensor node \( v \) increases the total number of cluster-heads generated, \( N_{CH} \), by one. The repetition stops when \( N_{CH} = C \) since there have been \( C \) cluster-heads elected. Algorithm ACE-C therefore will stop and the result will be used in the next step of the clustering phase.

Please note that, the value of \( v_{id} \) is kept and used continuously when algorithm ACE-C is invoked. Fig. 4 shows the high-level description of algorithm ACE-C and we present an example in Fig. 5. There are nine sensor nodes numbered from 0 to 8 as the identifiers. Suppose the number of cluster-heads in each cluster is \( C = 3 \). By following algorithm ACE-C, sensor nodes 6, 7, and 8 are the clusters-heads in the first round after three loops. The next three loops will decide the next three cluster-heads in the next round. The process continues for deciding the cluster-heads.

Algorithm ACE-C still works when some sensor nodes die (use up all of the energy) during the election process. Suppose sensor node \( u \) dies and it is its turn to be a cluster-head. Since sensor node \( u \) is gone, all the other sensor nodes cannot receive the advertisement message from \( u \). After waiting a period of time (i.e. Step (2.3) in Fig. 4), all the other sensor nodes will ignore \( u \) and continue the process until there are \( C \) cluster-heads elected. In Fig. 5, sensor node 6 dies in the third round. According to the algorithm, sensor node 6 should be a cluster-head in the fourth round. However, since it has died, no message will be sent at loop 12. All the other sensor nodes wait a period of time and receive nothing. They will assume that sensor node 6 has gone and then resume the loop again. At this time, sensor node 5 will be a cluster-head at loop 13. Therefore, the cluster-heads in the fourth round are nodes 5, 7, and 8 instead.

4.2. Cluster-head election with location

In this subsection, we present a distributed algorithm of cluster-head election with location, Algorithm ACE-L, which is especially for mobile sensor nodes. The basic idea is to use the node mobility to have each sensor node be a cluster-head in turns. Given some fixed reference points in the area of the mobile sensor network, the sensor nodes closest to these reference points will be the cluster-heads, respectively, when electing the cluster-heads. To achieve this, we consider to set the distance of a sensor node to a reference point as the metric of the delay time, which is used when a sensor node contends a channel. The decision is hence also a product of the channel contention among sensor nodes.

Suppose that there are \( C \) clusters in each round. Algorithm ACE-L defines \( C \) fixed reference points in the beginning. These \( C \) reference points will affect the priority for a node to contend a channel and the locations of the elected cluster-heads. Consider an arbitrary sensor node \( v \). We denote the reference point closest to \( v \) as the main reference point (MRP) of \( v \). The delay time of \( v \), which is the priority to contend a channel, is then set up according to the distance from \( v \)’s current location to its MRP. All the sensor nodes having the same MRP will contend a channel and the one with the shortest delay time will success. It turns out that the sensor node closest to an MRP will obtain the channel. We hence can elect such a sensor node as a cluster-head, which is closest to an MRP. When a sensor node has been elected as a cluster-head, it will use the channel obtained to broadcast a beacon of being a cluster-head to the other nodes. On the other hands, if sensor node \( v \) receives a beacon of being a cluster-head from other sensor node \( u \) during the delay time or the channel contention and node \( u \) has the
Algorithm of Cluster-head Election by Location

Input: C reference points
/* chb: cluster-head beacon */
(1) decide the MRP and set distance to be the distance to the MRP
(2) if receive chb from other node u and have the same main MRP as node u then
   (2.1) cluster-head ← true
   (2.2) exit
(3) else
   (3.1) delay time ← distance
   (3.2) while (delay time decreases one) do
      (3.2.1) if receive chb from other node u and have the same MRP as node u then
         (3.2.1.1) cluster-head ← true
         (3.2.1.2) exit
      endif
   endwhile
   (3.3) transmit its chb
endif
End

Fig. 6. Algorithm ACE-L used in each sensor node for cluster-head election.

Fig. 7. An instance of execution algorithm ACE-L in a sensor network having 15 mobile sensor nodes, where the dashed circles are the reference points.

same MRP as v, sensor node v stops the competition and will be not be a cluster-head in this round.

Fig. 6 presents algorithm ACE-L step by step. In Step (1), a sensor node v calculates the distances to all the reference points and uses the reference point having the shortest distance as the MRP. Step (2) then examines whether there is an elected cluster-head which uses the same MRP as v. If there is one, the algorithm stops. Otherwise, the algorithm moves to Step (3). In Step (3), a sensor node v sets up the delay time according to the distance to its MRP. After setting the delay time, node v again examines whether there is an elected cluster-head which uses the same MRP as in Step (2). If the delay time passes, sensor node v will successfully obtain the channel. This indicates that node v is the sensor node closest to its MRP and will be a cluster-head. It then transmits the beacon of being a cluster-head to the other sensor nodes.

An example of electing C = 4 cluster-heads among 15 mobile sensor nodes, v1, v2, ..., and v15, using algorithm ACE-L is shown in Fig. 7. Let the four reference points are rp1, rp2, rp3, and rp4, respectively. Consider sensor node v3. The distances from v3 to all the reference points rp1, rp2, rp3, and rp4 are d1, d2, d3, and d4, respectively, and d1 is the shortest distance. Reference point rp1 is the MRP of v3. Recall that we set the delay time according to the distance from a sensor node to its MRP. Hence v3 has the smallest delay time among all the sensor nodes which have rp1 as the MRP. After the channel contention, sensor node v3 succeeds and is the cluster-head by referring to rp1. Similarly, sensor nodes v5, v6, and v10 are the cluster-heads by referring to rp2, rp3, and rp4, respectively.

The correctness comes from the property of transmission among sensor nodes in a wireless environment. Under a fixed frequency, only one sensor node can use that frequency to transmit data at an arbitrary time instance. Therefore, at each time slot, there is only one sensor node can transmit the cluster-head beacon. Suppose there are C reference points, rp1, ..., rpC. By the property mentioned above, if a sensor node uses reference point rp1 as the MRP for some i ≤ C and obtains the channel, all the other sensor nodes use reference point rp1 as the MRP will not send any beacon and the cluster-head is thus elected. For the other reference points, the cluster-heads can be elected similarly. Therefore, the algorithm will generate C cluster-heads exactly unless there are fewer than C sensor nodes left.

4.3. Comparison between cluster-head election algorithms

We now compare algorithm ACE-C and algorithm ACE-L in three aspects. First of all, we consider the mobility of sensor nodes. Since algorithm ACE-C does not consider the locations of the sensor nodes, it fits for both mobile and static sensor networks. In contrast with algorithm ACE-C, algorithm ACE-L is designed especially for mobile sensor networks. When applying algorithm ACE-L to a static environment, some sensor nodes will use up all the energy quickly because the static reference points will make a sensor node being a cluster-head in a consecutive rounds until it dies. Hence, algorithm ACE-L does not fit for a static sensor network. To adopt algorithm ACE-L to a static environment, one can use dynamic reference points instead. However, every sensor node needs to know the moving pattern of the reference points.

Secondly, the mobility models on the sensor network will affect the performance of algorithm ACE-L since the algorithm refers to the fixed reference points. In general, the reference points are given uniformly in the area of the sensor network. In this paper, we consider that the sensor nodes are randomly scattered and each sensor node moves randomly under the three mobility models considered. The impact of the mobility models on the cluster-head election using algorithm ACE-L becomes less. For a mobility model where the sensor nodes are distributed uniformly, using algorithm ACE-L to elect cluster-heads may not be a good choice. However, since algorithm ACE-L does not consider the location of a sensor node, the mobility models will not affect algorithm ACE-L at all.

Last, we discuss these two proposed algorithms from the viewpoint of a good clustering algorithms mentioned in Section 2.
Recall the three objectives for a good clustering algorithm. It is not difficult to see that both algorithms can generate at least one cluster-head in each round and the number of cluster-heads elected in each round is uniform for these two algorithms. As for the cluster size, things are different for these two algorithms.

Since algorithm ACE-C only uses the ID to decide the cluster-heads without considering the locations of sensor nodes, the elected cluster-heads may be close to each other. As a result, the cluster sizes (i.e. the number of sensor nodes in a cluster) may be different among all the generated clusters. It turns out that some sensor nodes (being cluster-heads with a large number of sensor nodes in the corresponding clusters) may consume a large amount of energy and the system lifetime hence becomes short. Algorithm ACE-L uses the locations to elect the cluster-heads; hence, can avoid such a situation. Consider that there are 100 sensor nodes in a sensor network and the expected number of clusters in each round for the best energy consumption is 5 [11,14]. The expected cluster size thus is 20. We apply algorithms ACE-C and ACE-L to generate the cluster-heads, respectively, and both algorithms use the clustering mechanism, CM, discussed in next section to form the clusters. Fig. 8 shows the distribution of the standard deviation of the cluster size generated by algorithms ACE-C and ACE-L for 200,000 rounds. The results indicate that the cluster size produced by algorithm ACE-L is more uniform than the one generated by algorithm ACE-C. Indeed, in our experiments, algorithm ACE-L generates more percentage of clusters of size between 18 and 22 than algorithm ACE-C does. Therefore, algorithm ACE-L leads to a better performance in terms of energy consumption and the experimental results shown in Section 6 demonstrate this trend.

5. Clustering with mobility (CM)

After the cluster-heads have been elected, the next step in the clustering phase in a round is the cluster formation step. We observe that the node mobility plays an important role when forming the clusters. If a sensor node neglects its mobility when deciding which cluster-head to join, it may select a cluster-head with a longer distance to it; therefore, consumes more power. Recall that each sensor node is equipped with a Location Finding System. Having the location information from the attached location finding device, each mobile sensor node can calculate its speed and direction. The proposed mechanism, CM (Clustering with Mobility), uses the location information on each sensor node to decide which cluster this sensor node belongs to properly.

When the cluster-heads in the current round have been decided, the elected cluster-heads then broadcast an advertisement message to all the sensor nodes to recruit the cluster members. This advertisement message includes the cluster-head’s position, direction, and speed. Every non-cluster-head node must listen to such a broadcast after the cluster-head election has been done. After receiving the advertisement message, each sensor node needs to select a proper cluster-head to join.

To determine which cluster-head to join, each non-cluster-head node considers the relative position to each cluster-head and will select the nearest cluster-head to join. However, since all the sensor nodes are mobile, the currently nearest cluster-head may be the farthest one for the sensor node when the message transmission phase starts in the current round. Hence, the proposed mechanism, CM, uses the location information to predict the relative position when the message transmission phase starts. Nevertheless, it is not easy to predict when the message transmission phase will start in a round. We hence consider the time instance in the time period from the time, \( t_1 \), when the cluster-head election ends, to the time, \( t_2 \), when the next round starts. Suppose sensor nodes \( h_1, h_2, \ldots, h_n \) have been elected as the cluster-heads at time \( t_1 \). For a non-cluster-head sensor node \( v \), since the advertisement messages include the location information of the cluster-head nodes’, sensor node \( v \) can calculate all the distances to all the cluster-heads at time \( t_1 \) and \( t_2 \), respectively. Let \( d^1_i \) and \( d^2_i \) be the distances from sensor node \( v \) to the cluster-head node \( h_i \) at time \( t_1 \) and \( t_2 \), respectively, for \( i = 1, 2, \ldots, n \). At a given time instance \( t \in [t_1, t_2] \), the distance from sensor node \( v \) to each cluster-head \( h_i \) can be estimated with

\[
DH_i(t) = \frac{d^2_i - d^1_i}{t_2 - t_1}(t - t_1) + d^1_i, \quad i = 1, 2, \ldots, n. \tag{3}
\]

Fig. 9 gives a picture of the method we have described above. Then, suppose the time instance \( t \) is given, sensor node \( v \) will select a cluster-head \( h_j \) to join during the cluster formation step, where

\[
DH_j(t) = \min_{i=1}^{n}(DH_i(t)).
\]

Please note that Eq. (3) is an estimator for the distance from a sensor node \( v \) to a cluster-head at time \( t \). It is not the exact distance. One, of course, can use another way to estimate.
Recall that a cluster-based data-gathering protocol consists of a series of rounds and there are two major phases: clustering and message transmission phases in each round. The proposed algorithms ACE-C and ACE-L are used for cluster-head election and the mechanism CM is used for cluster formation in clustering phase. Using different algorithms in a step makes different protocols and different results. Hence, in our experiments, we consider the cluster-based protocols which employ the proposed distributed algorithms. We first implement the protocol CM using the mechanism, discussed in Section 5, which forms the clusters by considering the node mobility. The step for electing cluster-heads in protocol CM is the same as the one used in LEACH. Then, based on CM, we further implement two other protocols which differ in the way to elect cluster-heads as discussed in Section 4. Protocol ACE-C uses the cluster-head election with counting and protocol ACE-L uses the cluster-head election with location. The flow chart in Fig. 10 gives a picture about the process of a cluster-based protocol at each sensor node in our simulation. From now on, we will use the protocols instead of the algorithms when we discuss the performance.

The experimental results show that all these three protocols, CM, ACE-C, and ACE-L, make the whole system live longer in comparison with LEACH. For the experimental results, we first discuss the system lifetime and make a comparison between these protocols. Then, we demonstrate that the number of cluster-heads generated by the cluster-head election algorithms is uniform. Last, we show the impact of the clustering factor $T_c$ used in CM mechanism.

We consider three mobility models in the experiments, including SM model, RWM model, and RDM model. The simulation environment is set up as follows. The sensor network area considered in the simulation is $200 \times 200$ m. The BS is located at position (100, 300). There are 100 mobile sensor nodes in the sensing area. The speed of each sensor node is from 0 to 1 m/s and the initial energy in a sensor node is 1.0 J. We have each sensor node send a 2000-bits data packet to the BS in each round. The period of one round is 5 s. Initially, all the sensor nodes are assumed to be scattered randomly in the area. For the cluster-head election, the initial probability $P$ of a sensor node to be a cluster-head is 0.05 when we consider LEACH protocol. The expected number of cluster-heads in each round is therefore $100 \times 0.05 = 5$.

We list all the parameters used in the simulation in Table 1 for reference.

### 6.1. System lifetime

In the experiments, we use the system lifetime as the measure of energy consumption. The system lifetime is the total number of rounds which a wireless mobile sensor network experiences. We first consider the round at which the first dead sensor node occurs. Recall that, if the number of cluster-heads in each round is equal, the energy consumption will be similar among all the sensor nodes and hence the system lifetime can last longer. In particular, an equal number of cluster-heads in each round makes the first sensor node died in a latter round. Table 2 shows...
Fig. 10. The process of a cluster-based protocol at each sensor node in our simulation; using different approaches at each step making different protocols and results.

Table 1
Parameters used in the simulation

<table>
<thead>
<tr>
<th>Items</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network area</td>
<td>200 m × 200 m</td>
</tr>
<tr>
<td>Base-station location</td>
<td>(100,300)</td>
</tr>
<tr>
<td>Mobility models</td>
<td>SM, RWM, RDM</td>
</tr>
<tr>
<td>Number of sensor nodes</td>
<td>100</td>
</tr>
<tr>
<td>Speed of a sensor node</td>
<td>between 0 and 1 m/s</td>
</tr>
<tr>
<td>Size of the packet (in a round)</td>
<td>2000 bits</td>
</tr>
<tr>
<td>Period of each round</td>
<td>5 s</td>
</tr>
<tr>
<td>Transmission distance</td>
<td>&gt; 300 m</td>
</tr>
<tr>
<td>Transmission rate</td>
<td>&gt; 1 M/s</td>
</tr>
<tr>
<td>Reference points</td>
<td>(50,50), (150,150), (150,50), (50,150)</td>
</tr>
<tr>
<td>ID’s of sensor nodes</td>
<td>0–99</td>
</tr>
<tr>
<td>Expected number of cluster-heads</td>
<td>5 (or 4 for MRP)</td>
</tr>
</tbody>
</table>

the average system lifetimes with different mobility models in terms of the number of rounds.

In this experiment, we run each protocol 40 times and measure the round at which the first sensor node dies and the round at which the last sensor node dies (i.e. the total number of rounds the sensor network experiences). All the three proposed protocols perform better than LEACH does and can make the system live longer. Protocol CM can improve LEACH protocol under the mobile environments. Based on the CM mechanism, protocols using our cluster-head election algorithms, ACE-C and ACE-L, can further make the system live much longer due to an equal number of cluster-heads elected in each round. In general, the proposed protocols can also make the first dead sensor node occur in a later round. However, when using algorithm ACE-L on the SM model, because there are only four directions to select initially for a sensor node and a sensor node bounces back when it meets the boundary, a sensor node may be a cluster-head more often than the others. Therefore, a mobile sensor network with SM model using protocol ACE-L may have the first death node occur earlier as shown in Table 2.

We further plot the system lifetimes when each sensor node has initial energy 1.0 J with RDM model and RWM model in Figs. 11 and 12, respectively. Again, as shown in the plots, our proposed protocols make the system live longer than LEACH does, thereby result in a better performance than LEACH in terms of energy consumption. Protocols ACE-C and ACE-L use algorithms ACE-C and ACE-L to elect the cluster-heads in each round, respectively. As mentioned in Section 4.3, these two cluster-head election algorithms can generate at least one
Table 2
The performance about the system lifetimes of all the protocols with different mobility models

<table>
<thead>
<tr>
<th></th>
<th>First node dead round</th>
<th>Last node dead round</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LEACH</td>
<td>CM</td>
</tr>
<tr>
<td>SM</td>
<td>304</td>
<td>309</td>
</tr>
<tr>
<td>RWM</td>
<td>384</td>
<td>426</td>
</tr>
<tr>
<td>RDM</td>
<td>458</td>
<td>501</td>
</tr>
</tbody>
</table>

In general, protocol ACE-L is better than protocol ACE-C since the positions of the cluster-heads in each round are more uniformly distributed in the interested area. In some rounds, there could be five cluster-heads close to each other when applying protocol ACE-C. In such a case, some cluster-heads will consume the energy dramatically due to a large cluster size as mentioned in Section 4 and the cluster members also need to consume more energy due to a long distance to the cluster-heads. In our experiments, we further set the number of the reference points to be four and the reference points are located at the centers of the four quadrants of the sensing region. For five reference points, we add one reference at the center of the sensing region.

Besides, since the CM mechanism predicts the sensor node location more precisely in the RDM model, the performance of using the RDM model is better than the one of using the RWM model.

6.2. Number of generated cluster-heads

As mentioned in Section 2, by using the cluster-head election strategy of LEACH protocol, the difference in the number of cluster-heads between the rounds may be big and it happens that there is no cluster-head at all in some rounds. Our experiment results show that the range of the number of the cluster-heads in a round can be from 0 to 10 when applying the rule in LEACH and the percentage of the round in which there is no cluster-head elected is about 12–15%. Such a cluster-head election will deteriorate the performance in terms of the energy consumption. Table 3 shows the distribution of the number of the cluster-heads generated in each round for protocols ACE-C, ACE-L, and LEACH. The results also indicate that if one expects $C$ cluster-heads in each round, our distributed algorithms will elect $C$ cluster-heads exactly unless there are fewer than $C$ sensor nodes left in the system. Since algorithm ACE-C uses counting, it will generate more rounds having $C$ cluster-heads than algorithm ACE-L does.
6.3. Impact of clustering factor and the sensor node speed

To study the impact of clustering factor $T_c$, we consider the time period of a round is about 5 s. The value of $T_c$ ranges from 0 to 1. For each fixed $T_c$, we run 100 different cases and take the average system lifetime (total number of rounds) among these 100 cases. For a wireless mobile sensor network where each sensor node has an initial energy 1.0 J in an area of 200 m $\times$ 200 m with the RWM model, Figs. 13(a) and (b) show the effects of different $T_c$ values on the system lifetime for protocol ACE-L and ACE-C, respectively. In general, as the value of $T_c$ increases, the system lifetime increases. Recall that the message transmission phase is the second phase in a round and consumes most of the energy. As $T_c$ is larger than 0.5, $T_c$ is closer to the time when the message transmission phase starts and protocols using CM can have better prediction on the relative location to find the nearest cluster-heads. Hence, setting $T_c > 0.5$ leads to less energy consumption for transmission in the system.

We now further discuss the CM mechanism. Since mechanism CM works well when the location of a sensor node can be predicted well, we consider the SM model where one can fully predict sensor node’s location and speed. Fig. 14 shows the effects of different $T_c$ values on the system lifetime in a wireless mobile sensor network where each sensor node has an initial energy 1.0 J in a sensing area of 200 m $\times$ 200 m with the SM model. As shown in Fig. 14, as the clustering factor $T_c$ increases, the system lifetime increases and the best value of $T_c$ is between 0.75 and 0.9.

We observe that the sensor node speed impacts the performance of CM mechanism as well; therefore, impacts the system lifetime. Fig. 15 shows the effect of different maximum speeds ranged from 0 to 2.0 m/s on the system lifetime in a wireless mobile sensor network where each sensor node has an initial energy 1.0 J in an area of 200 m $\times$ 200 m on the RWM model. The cluster factor is set to $T_c = 0.875$ by following the conclusion in Fig. 14. When the maximum speed in a mobile sensor

---

**Table 3**
The percentage of different number of the cluster-heads elected in a round during a system lifetime using different cluster-head election algorithms

<table>
<thead>
<tr>
<th># of cluster-heads</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Over 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEACH (%)</td>
<td>12</td>
<td>3</td>
<td>6</td>
<td>15</td>
<td>14</td>
<td>15</td>
<td>12</td>
<td>10</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>ACE-C (%)</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.1</td>
<td>99.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ACE-L (%)</td>
<td>0</td>
<td>0.6</td>
<td>0.4</td>
<td>0.4</td>
<td>0.7</td>
<td>97.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

---

**Fig. 13.** The effects of different $T_c$ values on the system lifetime in a wireless mobile sensor network with Random Walk Mobility model. (a) ACE-L protocol, (b) ACE-C protocol.

**Fig. 14.** The effect of different $T_c$ values on the system lifetime in a wireless mobile sensor network with simple mobility model.
network increases, the change on the distance between a sensor node and the cluster-head becomes dramatic. Proper prediction on the location when forming the clusters will improve the performance. On the other hand, if the average speed is slow, using prediction may not help too much since the communication range between sensor nodes may not change too much. Fig. 16 shows the trend of the impact of both the clustering factor and the sensor node speed on the performance of CM mechanism. We can conclude that, in general, as the clustering factor and the maximum speed increase, the system lifetime increases for CM mechanism.

7. Conclusions

In this paper, we provide different distributed clustering algorithms which lead less energy dissipation for data-gathering in a cluster-based mobile sensor network. There are two steps when clustering the mobile sensor nodes: cluster-head election and cluster formation steps. For cluster formation, our CM (Cluster with Mobility) mechanism can achieve a better performance in terms of the energy consumption and system lifetime when the sensor nodes are capable of mobility. Two distributed cluster-head election algorithms, ACE-C and ACE-L, are proposed for cluster-head election in a wireless mobile sensor network. Based on the CM mechanism, the two proposed cluster-head election algorithms make different clustering algorithms. Our clustering algorithms achieve the three objectives: there is at least one cluster-head elected in each round, the number of cluster-heads generated in each round is always the same (except the final rounds), and all the generated clusters should have the same cluster sizes.

We consider the Random Walk Mobility and Random Direction Mobility models as well as the simple mobility model in our experimental analysis. The experimental results show that our algorithms yield a better performance in terms of energy consumption; therefore, lead a longer system lifetime. When applying CM mechanism, we conclude that as the maximum speed of a sensor node or the clustering factor increases, the system lifetime increases. The experimental results also indicate a better clustering factor which is from 0.75 to 0.9. Our future work tailors toward providing distributed clustering algorithms which fit different mobility models.

Appendix A. Discussion on the number of reference points

First, we define the following notations:

- $k$: the number of the reference points,
- $N$: the total number of sensor nodes,
- $l$: the packet size,
- $d_{toBS}$: the distance between cluster-head and BS, and
- $d_{toCH}$: the distance between a sensor node and cluster-head.

Suppose all the sensor nodes are uniformly distributed in the network area. Using the radio model and data-fusion model in [14], the energy consumption of a cluster-head in each round is

$$E_{CH} = E_{TX} + E_{RX} + E_{data-fusion}$$

$$= l \times E_{elec} + l \times \varepsilon_{amp} \times (d_{toBS})^2 + \left( \frac{N}{k} - 1 \right)$$

$$\times l \times E_{elec} + \left( \frac{N}{k} - 1 \right) \times E_D$$

$$= l \times \left[ E_{elec} + \varepsilon_{amp} \times (d_{toBS})^2 + \left( \frac{N}{k} - 1 \right) (E_{elec} + E_D) \right]. \quad (5)$$

On the other hand, the energy consumption of a non-cluster-head in one round can be derived as

$$E_{non-CH} = E_{TX}$$

$$= l \times E_{elec} + l \times \varepsilon_{amp} \times (d_{toCH})^2$$

$$= l \times (E_{elec} + \varepsilon_{amp} \times (d_{toCH})^2). \quad (6)$$
By combining Eqs. (5) and (6), the total energy consumption of all the sensor nodes in a cluster is

\[
E_{\text{cluster}} = E_{\text{CH}} + \left( \frac{N}{k} - 1 \right) E_{\text{non-CH}}
\]

\[
= I' \left \{ E_{\text{elec}} + \varepsilon_{\text{amp}} \times (d_{\text{BS}})^2 \right \} + \left ( \frac{N}{k} - 1 \right ) \left ( 2E_{\text{elec}} + E_D + \varepsilon_{\text{amp}} \times (d_{\text{CH}})^2 \right ) \right \}.
\]

(7)

Using Eq. (7), we can obtain the total energy consumption of all the sensor nodes in a round. We first consider the case where four reference points are used. Recall that the four reference points are located at (50,50), (50,150), (150,50), and (150,150), respectively, and the BS is located at (100,300) in the experiments. According to [11], let \((d_{\text{CH}})^2 = \frac{250^2}{2} \). Then, by Eq. (7), the total energy consumption of all the sensor nodes in a cluster for each round is

\[
E_{\text{cluster}} = 0.0127 + 0.0000002 \times (d_{\text{BS}})^2.
\]

(8)

By applying the distances between the cluster-heads and the BS to Eq. (8), we can conclude that, when there are four reference points located as above, the total energy consumption of all the mobile sensor nodes is a round is 0.0868J. Similarly, when there are five reference point located at (50,50), (50,150), (150,50), (150,150), and (100,100), respectively, we can derive the total energy consumption of all the mobile sensor nodes in a round as 0.083J. Therefore, using four reference points outperforms using five reference points in terms of energy consumption.

References


Chuan-Ming Liu is an assistant professor in the Department of Computer Science and Information Engineering, National Taipei University of Technology (NTUT), Taiwan. He received his Ph.D. in Computer Sciences from Purdue University in 2002 and B.S. and M.S. degrees both in Applied Mathematics from National Chung-Hsing University, Taiwan, in 1992 and 1994, respectively. He is a member of Upsilon Pi Epsilon Honor Society in Computer Science since 1998. His research interests include parallel and distributed computation, data management and data dissemination in wireless environments, ad hoc and sensor networks, and analysis and design of algorithms.

Chuan-Hsiu Lee received his B.S. and M.S. degrees in Electrical Engineering and Computer Science from National Taipei University of Technology (NTUT), Taiwan in 2003 and 2005, respectively. His research interests are in the areas of mobile sensor networks, ad hoc networks, and mobile computing. Currently, he is an engineer in ALi Co., Taiwan.

Li-Chun Wang received his B.S. degree from National Chiao Tung University, Taiwan, ROC in 1986, the M.S. degree from National Taiwan University in 1988, and the Ms. Sci. and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, in 1995, and 1996, respectively, all in electrical engineering. From 1990 to 1992, he was with the Telecommunications Laboratories of the Ministry of Transportation and Communications in Taiwan (currently the Telecom Labs of Chunghwa Telecom Co.). In 1995, he was affiliated with Bell Northern Research of Northern Telecom, Inc., Richardson, TX. From 1996 to 2000, he was a Senior Technical Staff Member in the Wireless Communications Research Department of AT&T Labs-Research. Since August 2000, he has been an Associate Professor in the Department of Communication Engineering of National Chiao Tung University in Taiwan and became a full professor in August 2005. His current research interests are in the areas of adaptive/cognitive wireless networks, radio network resource management, cross-layer optimization, and cooperative wireless communications networks.

Dr. Wang was a co-recipient (with Gordon L. Stuer and Chin-Tau Lea) of the Jack Neubauer Best Paper Award of the IEEE Vehicular Technology Society in 1997. He has published over 30 journals and 70 international conference papers and is holding three US patents. He served as an Associate Editor for the IEEE Transactions on Mobile Computing and Networking for IEEE Journal on Selected Areas in Communications in 2005 and the special issue on “Radio Resource Management and Protocol Engineering in Future IEEE Broadband Networks” for IEEE Wireless Communications Magazine in 2006.