ERP-DDA: Event Reporting Protocol Based on Distributed Data Aggregation for Wireless Sensor Networks

Raed T. AL-Zubi, Mohammad Q. Alawad, Abdulraheem A. Kreishan
Department of Electrical Engineering; The University of Jordan; 11942, Amman, Jordan; e-mail: r.alzubi@ju.edu.jo

Khalid A. Darabkh
Department of Computer Engineering; The University of Jordan; 11942, Amman, Jordan; e-mail: k.darabkeh@ju.edu.jo

Corresponding author: k.darabkeh@ju.edu.jo

Wireless Sensor Networks (WSNs) have a wide range of applications in human life. Accordingly, WSNs have been thoroughly considered in research community to improve their performance and address their challenges. Event reporting in WSNs has always been a major challenge in terms of energy consumption. Event reporting requires sensing the event and sending a reporting packet from the sensor to the centralized base station (BS). However, sensor’s energy is limited and stored in a non-rechargeable battery. Therefore, several event-reporting (or data collection) protocols have been proposed to improve energy consumption in WSNs, and consequently, to extend their lifetime. In this paper, we propose an efficient event reporting protocol for WSNs called Event Reporting Protocol Based on Distributed Data Aggregation (ERP-DDA). This protocol mainly aims at reporting any event by no more than one sensor node such that energy saving is satisfied in the whole network. To achieve this goal, ERP-DDA is mainly based on the following features: it is a cluster-based protocol, it is a multi-hop routing protocol, it applies distributed data aggregation, and it employs variable clustering and cluster head selection. Simulations show that ERP-DDA significantly extends the lifetime of WSNs compared with other related protocols.

KEYWORDS: Event Reporting, Cluster-based, Data Aggregation, Multi-hop Routing, WSNs, ERP-DDA.
1. Introduction

As a result of the great improvement and progress in communications, digital electronics, and Micro-Electro-Mechanical systems, applications on wireless sensor networks (WSNs) have become notably diverse including healthcare applications, environmental applications, military applications, etc [10]. The WSN is a network that has a number of small and smart nodes called sensors. There are different types of sensors including thermal sensors, radar sensors, acoustic sensors, infrared sensors, visual and magnetic sensors, and many others [38]. These sensors differ in terms of some features such as function, power consumption, sensing range, size, and cost. However, the sensors almost have the same basic and common structure; a power unit (i.e., a small non-rechargeable battery), communications unit (i.e., a wireless transceiver), sensing unit, and data processing unit [9].

To monitor a remote region, the sensors are randomly distributed over this region. The function of the sensors is summarized as follows: sensing the surrounding environment, converting the sensed information into data packets, and sending the data packets over a wireless channel to a base station (BS). After that, the BS may send these data packets to other destinations for more analysis [35]. Based on the application, the BS may be inside or outside the monitored region. Even though WSNs have different important and useful applications, they have many issues and challenges that have been considered by many researchers. Some of these issues are node deployment [36], energy constraint [46], scalability issue [34], data aggregation [3], security, and network heterogeneity. In this paper, we consider the issue of energy constraint in WSNs. This issue is disturbing both academia and industry researchers in the WSNs field [27]. The sensors require a lot of energy in order to perform different tasks such as environment sensing, data analysis and aggregating, and data sending. However, in sensors, the main source of energy is a non-rechargeable battery which almost cannot be changed after the deployment of the sensors in the network [42]. Therefore, WSNs require well-designed energy-aware protocols for event reporting and data collection in order to guarantee reduced energy consumption and extended battery lifetime.

Accordingly, several event reporting and data collection protocols (i.e., routing protocols) for WSNs have been proposed in literature. The proposed protocols can be generally classified into three categories; flat-based [6, 23, 30], tree-based [26, 29], cluster-based [1, 11, 44]. The last category (i.e., the cluster-based) has been proven to be efficient energy-aware routing protocol for WSNs [37, 43]. Therefore, in this paper, we mainly consider and thoroughly study this category of protocols. According to this type of protocols, the nodes in the network are classified into two types: Cluster Head (CH) nodes and normal nodes. Some protocols (as in [5]) have another type of nodes called Relay Nodes (RNs) that receive data from CHs and then send it to BS. The normal nodes in each cluster send their data to their corresponding CH. After that, each CH aggregates the overall data (i.e., removes redundancy in the collected data) and sends it to the BS. The network lifetime is divided into time intervals called rounds. The formation of clusters may be done in a fixed manner as in [13, 18] (i.e., the clustering is done after sensors deployment and it is not changed during the network lifetime) or in a variable manner as in [14, 15, 16] (i.e., it is changed every round).

1.1. Motivation

After studying different event reporting and data collection protocols proposed for WSNs, we have determined some of their elegant features that could be employed to propose a new energy-aware event reporting protocol for WSNs. These features mainly are the cluster-based, the multi-hop, and the variable clustering. The cluster-based feature allows data aggregation, which, by itself, reduces redundancy in data generated by the nodes that are located close to each other. Moreover, different research works have shown that clustering is efficient in WSNs in terms of energy saving [28, 40]. For multi-hop feature, it has been shown in literature that transmitting data over multi-hop route could be more energy efficient than that over single-hop route [2]. For variable clustering, we see that it is more efficient than fixed clustering in terms of load balancing. To illustrate, assume that in a certain region of a WSN, there is a number of events occur frequently more than that in other regions. It is clear that in a fixed clustering, these events will be reported by only
certain CHs during the network lifetime. This will lead to early death for these CHs before others. However, in variable clustering, the task of reporting these events will be distributed over different nodes in the network.

In addition to the above good features, we found that most cluster-based routing protocols (fixed or variable clustering) have a critical problem that severely degrades their performance in terms of energy saving. We call this problem as Intra/Inter-Cluster Event-Reporting Problem (IICERP). This problem appears on intra transmission and inter transmission levels. In the intra transmission level, if an event occurs in a certain area in the WSN, then different nearby nodes from the same cluster or different clusters will unnecessarily send similar packets to their CHs in order to report this event. In the inter transmission level, if an event occurs in a certain area in the network, then more than one nearby CH will unnecessarily send similar packets to inform the BS about the event. Figure 1 illustrates IICERP. Nodes N1, N2, and N3 unnecessarily send similar data packets to inform CH1 about event A. Moreover, CH1 and CH2 unnecessarily send similar data packets to the BS in order to inform about event B. Thus, to inform about event A, three similar data packets are sent from three different nodes. Moreover, CH1 and CH2 will perform data aggregation for the similar three data packets which leads to energy wasting. Also, sending similar packets by CH1 and CH2 to the BS to report the same event (event B) will consume more and more energy.

Figure 1
IICERP on both intra and inter transmission levels

1.2. Contribution

In this paper, we propose an Event Reporting Protocol Based on Distributed Data Aggregation for WSNs (ERP-DDA). The design of ERP-DDA is based on the features discussed in the motivation section; it is a multi-hop and variable clustering protocol as well as it considers and solves the problem of IICERP via distributed data aggregation. For distributed data aggregation, we propose an algorithm which strictly implies that any node senses an event should, first, listen to its neighbors which have higher residual energy and could sense the same event. Then, if the node hears that one of its neighbors reported the same event, it will ignore this event and will not send any packet to report the event. Therefore, the redundancy in packet reporting is removed in a distributed manner (i.e., distributed data aggregation). The process is not as in most cluster-based protocols where data aggregation is centralized (i.e., it is only done by the CH in each cluster). Hence, distributed data aggregation ensures that any event is reported by no more than one sensor. This leads to reducing the energy consumption and extending the network lifetime. Simulation results show that ERP-DDA significantly saves energy and so extends the network lifetime compared with other related protocols. It is worthwhile to mention that the novelty of our work can be summarized as follows: to the best of our knowledge, no previous work considered the integration of multi-hop feature, variable clustering feature, and specifically the problem of IICERP into one protocol.

1.3. Organization

The rest of the paper is organized as follows. Section 2 presents some related work. Section 3 presents and explains ERP-DDA. In Section 4, we use simulation to evaluate and compare the performance of ERP-DDA with other protocols. Finally, concluding remarks are drawn in Section 5.

2. Related Work

Many data collection and routing protocols for WSNs have been proposed. Interestingly, the Low Energy Adaptive Clustering Hierarchy (LEACH) [27] protocol is currently the most influential and most inspiring among the established hierarchical routing proto-
cols. In the LEACH protocol, in an effort to minimize the number of nodes that directly interact with BS, nodes are systematically grouped into clusters. As stated in [27], LEACH's operation is divided into several rounds. Two phases, the set-up and steady-state phases, constitute each round. A distributed LEACH-based CH selection strategy called LEACH with Distance-based Thresholds (LEACH-DT), aimed at achieving energy balance, was suggested by Kang and Nguyen [31]. However, according to their distances to the BS, sensor nodes in the LEACH-DT protocol are self-chosen to become CHs with dissimilar probabilities. LEACH-DT utilizes the same threshold probability for CH selection, similar to the LEACH protocol, with the major difference that the percentage of CHs is not a predefined dimension. Particularly, LEACH-DT operates a separated percentage of CHs, which is parametrized by the distance to the BS. A Centralized Energy Efficient Distance (CEED) protocol was suggested by Gawade and Nalbalwar [25]. In this protocol, not only the circular clusters are considered, but also the appearance of the energy depleted in the phases of cluster formation, CH selection as well as steady-state phase is considered. The optimal number of CHs can be easily formulated by doing this. The CH selection approach in the CEED protocol is fully distributed, taking into account the distance between the sensor node and BS and the depleted energy of the sensor.

Two novel CH selection protocols for WSNs, called EDB-CHS and EDB-CHS-BOF, which are key extensions of the LEACH-DT and CEED protocols, were proposed by Darabkh et al. [21]. For the first proposed protocol (EDB-CHS), there are mainly two major results. The first belongs to interesting analysis which ends up providing a tight formula for effectively defining the optimum number of CHs assuming that the hexagonal cluster shape is employed, which is more practical and different from the circular shape, which is widely used in the literature and specifically in both LEACH and LEACH-DT protocols. The other one refers to the CH selection algorithm update. A new optimal likelihood for each sensor node to function as a CH is extracted from the EDB-CHS-BOF protocol. A new energy-aware and density-based clustering and routing protocol (EA-DB-CRP) was proposed by Darabkh et al. [20] for data collection in WSNs, which essentially aims to spread the load between sensor nodes, which results in balancing the energy consumption of the network and thus extending the life of the network. More specifically, they presented a network model that ends up with empirical expressions that explain how the network area can be effectively divided into layers and sub-layers of equal-size. By proposing a new protocol, namely, Life Time Maximizing Based on Analytical Hierarchical Process and Genetic Clustering (LiM-AHP-G-C) protocol, Darabkh et al. [19] addressed the problem of IoT networks as the integration of energy-restricted sensor devices. Specifically, the proposed protocol introduces a new optimal clustering algorithm for unrechargeable IoT sensor devices operated by batteries, an effective algorithm for selecting IoT heads, a heuristic approach for selecting optimal hops, and a powerful design for getting rid of both inter- and intra-cluster interferences.

To this extent, we present and discuss the works that are quite related to our proposed protocol such as Load Balancing Cluster Head (LBCH) [4], Round-Robin Cluster Header (RRCH) [33], LEACH-Fixed (LEACH-F) [27], Adaptive LEACH-F [7], An Adaptive Energy-aware Fixed Clustering Protocol (AEA-FCP) with continuous data (AEA-FCP-CD) [12], AEA-FCP with event-based (AEA-FCP-EB) [12], and Self-incentive and Semi Re-clustering (SISR) [8]. In Section 4, we conduct simulation to compare the performance of ERP-DDA with all of these protocols. Al-Zubi et al. [4] proposed the LBCH protocol. LBCH applies fixed clustering technique. Each cluster has one cluster head (CH) node, one Relay Node (RN), and many member nodes. The member nodes, in each cluster, sense the events and then send reporting packets to their corresponding CH. After that, the CH aggregates the received packets and resends these packets to the RN. In a multi hop manner, RNs send to the BS the received packets. According to this protocol, the lifetime is divided into setup phase and steady state phase that is divided into time periods called rounds. In the setup phase, the BS sends a hello message to all nodes in the network. After that, the nodes send their locations and energy levels to the BS. Then, the BS divides the network area into subfields called clusters. These clusters are fixed over the lifetime of the network. After that, the BS chooses one RN and one CH node for each cluster. The selection of the RN and CH is based on two calculated values called magnitude and weight, respectively. After that, the BS determines for each RN the next hop RN. For the fol-
lowing rounds, the selection process of RNs and CHs is repeated in each round to ensure load balancing in the network. Nam and Min [33] proposed the RRCH protocol. RRCH is a fixed clustering protocol which achieves load balancing and a high energy efficiency in WSNs. In the setup phase, clustering process is done only once. In the steady state phase, there is no RNs selection. However, for each round, the selection of CH nodes for each cluster is done according to round robin method. Heinzelman et al. [27] proposed the first fixed clustering routing protocol that is based on a well-known protocol LEACH. The new protocol is called LEACH-F. The clusters are constructed during the setup phase by applying a centralized cluster formation algorithm. In the steady state phase, the CHs selection is done in a round robin manner. Azim and Islam [7] proposed the dynamic round time-based fixed LEACH scheme (we call it Adaptive LEACH-F). It was proposed to reduce the problem of the fixed round time in LEACH-F. The round time in the Adaptive LEACH is adaptively changed according to the current energy of the member nodes (not to their initial energy) and the total energy consumption in the cluster for that round. This leads certainly to reducing the possibility of early death of CHs in addition to extending the network lifetime. Darabkh and Al-Jdayeh [12] proposed the AEA-FCP protocol. This protocol is based on adaptive fixed clustering. In the setup phase, the network is divided into a number of clusters that are fixed during the network lifetime. In the steady state phase, the RNs route over a multi-hop path the aggregated data from the CHs to the BS. The selection of RNs and CHs is performed in each round based on the residual energy in the nodes. The CH and RN in a cluster switch their roles if the residual energy in them is less than certain percentages of the average residual energy of the cluster when they start working as a CH and RN nodes. AEA-FCP is evaluated under two scenarios; continuous availability of data and event-based availability of data. Baek et al. [8] proposed the SISR protocol. It is a fixed clustering protocol. In the setup phase, each node in the network elects itself as a candidate CH with probability P and then broadcasts an ADVERTISE MESSAGE with the initial radio range Ri. Then, the node gradually increases Ri until it receives at least one ADVERTISE MESSAGE from other nodes. Accordingly, the node checks the probability (P) values of other nodes. If one of the nodes has higher value of P, it considers this node as its CH. Otherwise, the node gives up the competition. The selected CHs send an INVITE MESSAGE and wait for a feedback from the normal nodes. When the normal node finds its CH, it sends a JOIN_REQ_MESSAGE to its CH to inform it about its decision. After the construction of clusters, each CH determines its CH sequence based on the signal strength of JOIN_REQ_MESSAGE from the normal nodes in its cluster. In the steady state phase, HEARTBEAT_MESSAGEs are broadcasted by the CHs to their member nodes. The member nodes that do not respond by HEARTBEAT_ACK_MESSAGE are considered as dead nodes. However, alive nodes send their HEARTBEAT_ACK_MESSAGEs that include their incentive values to be CHs.

3. ERP-DDA Protocol

ERP-DDA mainly aims at reducing the energy consumed for reporting events in WSN in order to extend the network lifetime. The main features of ERP-DDA that help in achieving this goal are summarized as follows: it is a variable clustering and multi-hop protocol as well as it considers and solves IICERP problem. These features are explained by presenting the design of ERP-DDA.

3.1. ERP-DDA Design

According to ERP-DDA, the network lifetime is divided into multiple time intervals called rounds. Each round is divided into setup phase and steady state phase. In the setup phase of each round, the following steps are executed:

**Step 1.** All the alive nodes send to the BS its identification number, residual energy, and position (assuming the nodes have GPS or they apply a localization algorithm).

**Step 2.** The BS runs the algorithm proposed in LEACH protocol [27] for CHs selection. Other algorithms can be employed but in our work, we select the one proposed in LEACH protocol [27] because it is an energy efficient and satisfies load balancing.

**Step 3.** The BS runs Dijkstra’s algorithm [22] to find the best route (in terms of minimum energy) between each CH and the BS through other CHs.

**Step 4.** The BS divides the nodes into clusters where each node is associated to the nearest CH. To prevent interference between transmissions from different clusters, the BS generates a unique CDMA code for
each cluster. All the nodes of one cluster use the same code for transmission.

**Step 5.** To solve the problem of sending more than one packet in order to report the same event (we call this problem as Intra/Inter-Cluster Event-Reporting Problem -IICEP), the BS runs our proposed algorithm called Listening and Sending Scheduling Algorithm (LSSA) that is explained in Section 3.2. This algorithm generates a TDMA schedule that is followed by all the nodes during the steady state phase. Based on this schedule, each node knows at which time slot (we call it sending slot) and on which CDMA code it will send its packets. Also, it knows at which time slots (we call them listening slots) and on which CDMA codes it will listen to other nodes. At this point, if a node senses an event and the time of its sending slot passes without receiving any announcement from its neighbors about the same event, then the node will report this event. Otherwise, at the moment of receiving any announcement about the event, the node ignores this event and turns off its transceiver electronics to save its energy.

**Step 6.** Finally, the BS announces to all nodes information about the result of LSSA, the corresponding CH of each node, and the next hop of each CH.

### 3.2. Listening and Sending Scheduling Algorithm (LSSA)

The LSSA mainly relies on the fact that if there is a group of sensor nodes where their sensing regions are overlapped, then they will unnecessarily send similar packets for reporting an event occurs in the intersection region. The sensing region of a node is the region around the node where if an event occurs within it, the node will sense that event. The radius of the sensing region is called the sensing range Rs of a node [32]. Figure 2 shows a group of sensor nodes that will generate similar data packets while sensing the same event.

Accordingly, LSSA aims at reporting any event by only one packet generated by one node in order to save energy and achieve load balancing in the network. In other word, LSSA aggregates data in a distributed manner.

To achieve the objective of LSSA, firstly, it requires that if an event occurs in a certain region, then the node with the highest energy (among all the nodes from the same cluster or different clusters that sense the same event) will report this event by sending a packet to its corresponding CH. Secondly, the transmission of this packet should be done over at least 2Rs range such that all the possible nodes that could sense this event will know that there is a node that has higher energy than them reports this event instead of them. This means that a well-designed time-division multiple access (TDMA) schedule is needed to satisfy the following requirements:

1. The sensor nodes (from the same cluster or different clusters) that may sense the same event should be assigned different time slots. This is required to allow these nodes to hear the transmission of each other and hence avoid sending similar packets, and consequently, reducing the energy consumption.

2. The sensor node that has higher residual energy should be assigned the lower time slot index. This is required to ensure that the event is reported by the node that has the highest residual energy among its neighbors so load balancing is satisfied.

These requirements are considered in the design of LSSA. A pseudocode for LSSA is depicted in Algorithm 1 and illustrated in the following steps:

**Step 1.** The BS creates for all nodes in the network an initial TDMA schedule based on the residual energy in each node where the first time slot is assigned for the node that has the highest energy level in the network and the second time-slot is assigned for the node that has the second highest energy level and so forth. The length of the initial TDMA schedule is impractical and equals to the number of the nodes in the network. Therefore, in the next steps, the length of the initial TDMA will be reduced by assigning one time-slot for different nodes.
Algorithm 1. Listening and Sending Scheduling Algorithm (LSSA)

1: \[ \text{In} : \quad N \] // total number of nodes in the network
2: \[ \text{In} : \quad e(n), \quad 1 \leq n \leq N \] // residual energy in node \( n \)
3: \[ \text{In} : \quad \mathcal{N}(n) \] // neighbors of node \( n \) within \( 2Rs \) range
4: \[ \text{In} : \quad C(n) \] // group of nodes in the same cluster of node \( n \)
5: \[ \text{Out} : \quad G(s) \] // time schedule determines the transmission slot \( s \) of each node
6: \[ \text{Begin} \]
7: \[ \text{Set} \quad G(s) = \phi, \quad \text{for all} \quad s \text{ where} \quad 1 \leq s \leq N \]
8: \[ \text{Set} \quad N_g = \text{sorted list of nodes according to} \quad e(n) \text{ from the highest to the lowest.} \]
9: \[ \text{Set} \quad G(s) = N_g(s), \quad \text{for all} \quad s \text{ where} \quad 1 \leq s \leq N \quad \text{// assign each node in} \quad N_g \text{ in one time slot} \]
10: \[ \text{Set} \quad T_s(n) = \text{index of node} \quad n \quad \text{in} \quad N_g \quad \text{for all nodes} \quad \text{// time-slot index that assigned to node} \quad n \]
11: \[ \text{Set} \quad F_n(n) = T_s(n) \quad \text{for all} \quad n \text{ where} \quad 1 \leq n \leq N \quad \text{// flag number of node} \quad n \]
12: \[ \text{For} \quad n = 1 \quad \text{to} \quad N \]
13: \[ \text{If} \quad \text{the time-slot index of the node from} \quad \mathcal{N}(n) > F_n(n) \quad \text{then} \]
14: \[ \text{Change the flag number of the node in} \quad \mathcal{N}(n) \quad \text{to} \quad F_n(n) \]
15: \[ \text{End If} \]
16: \[ \text{End For} \]
17: \[ \text{For} \quad s = 2 \quad \text{to} \quad N \]
18: \[ \quad n = G(s) \quad \text{// this is the node in the time slot} \quad s \]
19: \[ \text{If} \quad F_n(n) = T_s(n) \quad \text{then} \]
20: \[ \quad \text{For} \quad i = 1 \quad \text{to} \quad s \]
21: \[ \quad \text{If} \quad G(i) \cap C(n) = \phi \quad \text{and} \quad G(i) \cap \mathcal{N}(n) = \phi \quad \text{then} \]
22: \[ \quad \text{Include} \quad n \quad \text{in} \quad G(i) \]
23: \[ \quad T_s(n) = i \]
24: \[ \quad F_n(n) = i \]
25: \[ \quad \text{If} \quad \text{the time-slot index of the node from} \quad \mathcal{N}(n) > F_n(n) \quad \text{then} \]
26: \[ \quad \text{Change the flag number of the node in} \quad \mathcal{N}(n) \quad \text{to} \quad F_n(n) \]
27: \[ \text{End If} \]
28: \[ \text{Loop Break} \]
29: \[ \text{End If} \]
30: \[ \text{End For} \]
31: \[ \text{Else If} \]
32: \[ \quad \text{For} \quad i = F_n(n) \quad \text{to} \quad s \]
33: \[ \quad \text{If} \quad G(i) \cap C(n) = \phi \quad \text{and} \quad G(i) \cap \mathcal{N}(n) = \phi \quad \text{then} \]
34: \[ \quad \text{Include} \quad n \quad \text{in} \quad G(i) \]
35: \[ \quad T_s(n) = i \]
36: \[ \quad F_n(n) = i \]
37: \[ \quad \text{If} \quad \text{the time-slot index of the node from} \quad \mathcal{N}(n) > F_n(n) \quad \text{then} \]
38: \[ \quad \text{Change the flag number of the node in} \quad \mathcal{N}(n) \quad \text{to} \quad F_n(n) \]
39: \[ \text{End If} \]
40: \[ \text{Loop Break} \]
41: \[ \text{End If} \]
42: \[ \text{End For} \]
43: \[ \text{End If} \]
44: \[ \text{End for} \]
45: \[ \text{End Main} \]
Step 2: Initially, each node is assigned a Flag Number (FN) equals to the index of its time-slot assigned in Step 1.

Step 3: Starting from the node with FN=1, the FNs of all its neighbors (within 2Rs range) are changed to 1. Also, for the neighbors of the node with FN=2, their FNs are changed to 2. This process is done for all nodes and the FN of a node should not be changed more than once in this step. In this step, for example, the first time slot is assigned for the node that has originally FN=1 and no one of its neighbors (that their FNs changed to 1) can use this slot for transmission.

Step 4: Starting from the second time-slot, if the node in the second time-slot and the node in the first time-slot are not neighbors and not members of the same cluster, then the node in the second time-slot is moved to the first time-slot and its FN is changed to 1 as well as the FNs of all its neighbors are changed to 1 instead of 2. This process is repeated for all nodes. To generalize this step, there are two cases should be clarified:

1. If the node has FN equals to the index of its time-slot, then this means that its FN is not changed and no neighbor in the previous time-slots exists. Therefore, it checks if it is possible to move to one of the previous time-slots (starting from the first one). If there is no member from the same cluster in the checked time-slot, then it can move to it. Otherwise, it cannot move.

2. If the FN of the node does not equal to the index of its time-slot, then this means that there is a neighbor in the time-slot of index equals to FN (that has the higher energy). Thus, the checking process should start from the time-slot of index equals to FN+1.

3.3. An Illustrative Example

LSSA can be clearly explained by the example shown in Figure 3. In this example, a WSN consists of 9 nodes grouped in 3 clusters. The neighbors and the cluster of each node are summarized in Table 1.

Table 1
Neighbors and clusters of 9 nodes WSN

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighboring nodes (within 2Rs)</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>E</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>F,G</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>E,F</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>A,C</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>C,B</td>
<td>2</td>
</tr>
<tr>
<td>G</td>
<td>B</td>
<td>2</td>
</tr>
<tr>
<td>H</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>I</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Step 1. The BS arranges all sensor nodes from the highest energy level to the lowest one and gives them an initial FNs. Assume that the nodes are arranged based on the residual energy as follows:

<table>
<thead>
<tr>
<th>Time-slot</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>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

| Node | E | G | F | D | H | A | I | C | B |

Step 2. The FNs of all neighboring nodes of node E (the highest residual energy node) will be changed to 1. Also, the FNs of all neighboring nodes of node G will be changed to 2. The same process is repeated for all nodes. The FN of the node should not be changed more than once in this step. The new schedule will be as follows:

<table>
<thead>
<tr>
<th>Time-slot</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>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

| Node | E | G | F | D | H | A | I | C | B |

Step 3. Check the possibility of moving the nodes to time-slots having smaller index than that of their current time-slots. For node G, its FN equals to the index of its time-slot. Therefore, all the time-slots before its time-slot will be checked. Therefore, node G will be compared with node E. Since nodes G and E are not neighbors and are not members in the same cluster, the FNs of node G and all its neighboring nodes will be...
changed to the FN of node E. The new schedule will be as follows:

<table>
<thead>
<tr>
<th>Time-slot</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>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Node</td>
<td>E, G</td>
<td>F</td>
<td>D</td>
<td>H</td>
<td>A</td>
<td>I</td>
<td>C</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

For node F, the index of its time-slot and its FN are equal. Hence, it will be compared with the nodes in all previous time-slots (1 and 2). Since nodes F and G are members in the same cluster, node F cannot be moved to the first time-slot but it can be moved to the second time-slot. Thus, the FNs of node F and its neighboring nodes (C and B) are changed to 2. The new schedule will be as follows:

<table>
<thead>
<tr>
<th>Time-slot</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>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Node</td>
<td>E, G</td>
<td>F</td>
<td>D</td>
<td>H</td>
<td>A</td>
<td>I</td>
<td>C</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

For node D, the index of its time-slot and its FN are equal. Hence, it will be compared with the nodes in all previous time-slots (1, 2, and 3). Since nodes D, G, and F are members in the same cluster, the node D will be moved to the third time-slot and the FN of node D will be changed to 3. The new schedule will be as follows:

<table>
<thead>
<tr>
<th>Time-slot</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>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Node</td>
<td>E, G</td>
<td>F</td>
<td>D</td>
<td>H</td>
<td>A</td>
<td>I</td>
<td>C</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

For node H, the same process will be repeated and the new schedule will be as follows:

<table>
<thead>
<tr>
<th>Time-slot</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>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Node</td>
<td>E, G</td>
<td>F</td>
<td>D</td>
<td>H</td>
<td>A</td>
<td>I</td>
<td>C</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

For node A, its FN and the index of its time-slot are not equal. Hence, it will be compared with the nodes in the previous time-slots starting from the time-slot that has index of 2 (i.e., FN of node A + 1). Since nodes A and F are not neighbors and are not members in the same cluster, the FN of node A will be changed to 2. Note that the FN of node E which is a neighboring node for node A will not be changed. The FNs of nodes, which their comparison processes have been finished, will not be changed again. The new schedule will be as follows:

<table>
<thead>
<tr>
<th>Time-slot</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>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Node</td>
<td>E, G</td>
<td>H</td>
<td>F, A</td>
<td>D</td>
<td>I</td>
<td>C</td>
<td>B</td>
<td>I, G, H</td>
<td></td>
</tr>
</tbody>
</table>

To this extent, we have five time-slots for this round. The fifth time-slot is assigned for CHs to send their packets to the BS. Then, the BS announces this schedule with the CDMA codes for each cluster. Then, each node in the network uses the proper CDMA code to send its packets in its assigned time-slot and listen to its neighboring nodes in their assigned time-slot. In our example, in the first time-slot, nodes E, G, and H send their packets and nodes A, B, C, and I listen to their neighbors. Nodes A and C adjust their receiver electronics on the CDMA code used by node E to lis-

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>The common simulation parameters</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data packet size</td>
<td>1024 bits</td>
</tr>
<tr>
<td>Control packet size</td>
<td>176 bits</td>
</tr>
<tr>
<td>Initial energy</td>
<td>2 J</td>
</tr>
<tr>
<td>ETX (Energy for transmitting one bit)</td>
<td>50 nJ/bit</td>
</tr>
<tr>
<td>ERX (Energy for receiving one bit)</td>
<td>50 nJ/bit</td>
</tr>
<tr>
<td>EDA (Energy for data aggregation)</td>
<td>5 nJ/bit</td>
</tr>
<tr>
<td>Transmit amplifier energy in free-space</td>
<td>10 PJ / bit / m²</td>
</tr>
<tr>
<td>Transmit amplifier energy in multi-path</td>
<td>0.0013 PJ / bit / m²</td>
</tr>
<tr>
<td>Simulation Area</td>
<td>100 m x 100 m</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>100</td>
</tr>
<tr>
<td>Distribution of nodes over the simulation area</td>
<td>Random</td>
</tr>
<tr>
<td>Base-station location</td>
<td>(50 m, 125 m)</td>
</tr>
<tr>
<td>Sensing range for each sensor</td>
<td>10 m</td>
</tr>
<tr>
<td>Percentage of CHs per round</td>
<td>5%</td>
</tr>
</tbody>
</table>
ten to node E since it is a neighboring node for each of them. Node B adjusts its receiver electronics on the CDMA code used by node G. The node I (i.e., CH node) receives the data packet of its member node E. In the second time-slot, nodes F and A send their packets if they did not receive similar packets from their neighbors in the first time-slot. Also, nodes B, C, G, and I turn on their receiver electronics to listen to their neighbors. Nodes B and C listen to node F; they adjust their receiver electronics on CDMA code assigned to F. The node G (i.e., CH node) receives the data packet of its member node F and the node I (i.e., CH node) receives the data packet of its member node A. In third time-slot, nodes G and H turn on their receiver electronics. The node G (i.e., CH node) receives the data packet of its member node D and the node H (i.e., CH node) receives the data packet of its member node C. In the fourth time-slot, only the node G (i.e., CH node) turns on its receiver electronics in order to receive the data packet of its member node B. In the fifth time-slot, all CH nodes send their packets to the BS using their CDMA codes.

4. Simulation Results and Discussion

In this section, we conduct simulations to evaluate and compare the performance of our proposed protocol ERP-DDA with different related protocols presented in Section 2. These protocols are LBCH [4], RRCH [33], LEACH-F [27], Adaptive LEACH-F [7], AEA-FCP-CD [12], AEA-FCP-EB [12], and SISR [8].

4.1. Simulation Environment

MATLAB program version (8.4.0.150421) R2014b is used to conduct simulations. The simulations are carried out using a computer with Intel Core i5 CPU working at 2.5 GHz with a 6GB RAM and is running on windows 10 Pro (64bits).

4.2. Simulation Parameters

Table 2 summarizes the parameters used in the simulation. The same parameters were used in [4] and [17].

4.3. Simulation Results

4.3.1. Scenario 1: Continuous Availability of Data

According to this scenario, the nodes in the network always have data to send. Therefore, the nodes will consume more energy to send their data packets and so short network lifetime is expected. To setup this scenario in simulation, the locations of events per round are distributed on a grid shape (i.e., the minimum distance between events is one sensing range). As a result, the total number of events per round is 50. This distribution of events is chosen such that each node, in each round, senses the closest event to it and has a packet to report this event. Since the nodes are randomly distributed over the network area, then it is possible that different nodes sense the same event. The performance of ERP-DDA is compared with other protocols that assume the same scenario, including: LBCH [4], RRCH [33], LEACH-F [27], Adaptive LEACH-F [7], and AEA-FCP-CD [12].

Figure 4 shows the total number of nodes alive per simulation round. From this figure, we can see that LEACH-F outperforms RRCH. This is an expected behavior because LEACH-F is employing a centralized clustering approach which ensures the load balancing among the nodes in the network. On the other hand, adaptive LEACH-F performs better than LEACH-F and RRCH as it overcomes the problem of fixed round time in LEACH-F by introducing the concept of dynamic round time. AEA-FCP-CD achieves higher performance compared with RRCH, LEACH-F and adaptive LEACH-F since it follows a centralized procedure to select the initial CHs and RNs. This procedure achieves load balancing among all the nodes in the network, limits the number of CHs, and ensures that CHs are well distributed over the network area. In addition, the use of RNs reduces the transmission distances and the energy depletion in CHs. For LBCH, we can see that it slightly achieves higher performance compared with AEA-FCP-CD. This is because LBCH employs a centralized procedure for selecting the initial CHs and RNs like AEA-FCP-CD but with different modifications. Furthermore, while other protocols employ just round robin, novel methodologies are evolved in LBCH protocol for rotating the role of CHs and RNs inside each cluster. Interestingly, it can be seen from Figure 4 that ERP-DDA shows a performance superiority over all the compared protocols. In the case of applying ERP-DDA, the number of nodes alive in the network is slowly decreasing with time which results in extending the lifetime of the network. This is because ERP-DDA integrates the good features in different protocols in one protocol as
well as it has an ability to overcome the weaknesses in all compared protocols. ERP-DDA is mainly based on four features that significantly reduce the energy consumed in reporting events in the network. First, ERP-DDA is a cluster-based protocol. This feature results in preventing the nodes in the same cluster to send the same data to the BS, and accordingly, reducing the consumed energy in the network. Moreover, different research works have shown that clustering is efficient in WSNs in terms of energy saving [45]. Second, ERP-DDA is a multi-hop routing protocol. Multi-hop routing between the CHs and the BS reduces the transmission distances, so reducing the transmission power, and leads to load balancing in the network. It has been shown in the literature that multi-hop routing protocols are more efficient than single-hop protocols in terms of energy consumption [24, 39, 41]. Third, ERP-DDA is based on variable clustering and variable cluster-head selection in each round which results in load balancing in the network, i.e., the task of reporting the events will be distributed over different nodes in the network. Fourth, as discussed before in Section 3, ERP-DDA considers and solves IIERP problem such that one packet is sent by one node to report one event. Thus, the consumed energy in the network is significantly reduced.

4.3.2. Scenario 2: Event-Based Availability of Data
In this scenario, the nodes generate data based on the surrounding events. For simulation, the number of events in each round is chosen to be random (i.e., uniformly distributed in the range of 1 to 50) and the locations of these events are randomly selected within the network area. Figure 5 shows that ERP-DDA achieves higher performance compared with the related protocols that assume the same scenario (i.e., LBCH [4], AEA-FCP-EB [12], and SISR [8]). This result is due to the same reasons discussed in scenario 1. Also, the figure shows that, in scenario 2, the performance of ERP-DDA is improved (compared with the other protocols) more than that in scenario 1. This can be explained as follows. The events in scenario 2 are more sparsely distributed than that in the first one (i.e., on average 25 events, instead of 50 events, are distributed randomly over the network area). This results in increasing the bad effect of the IIERP problem on the performance of the compared protocols. In Figure 6, we clearly illustrate this result. Figure 6 (a) depicts a scenario where two events occur near each other (i.e., dense events) and three nodes sense these events and announce them using the compared protocols by sending three packets to the CH. However, if ERP-DDA is employed, as shown in Figure 6 (b), then only one packet will be sent to the
Figure 5
Number of nodes alive per simulation round. A random number of events per round is selected and the locations of them are randomly distributed over the network area.

Figure 6
The effect of events distribution on the performance of ERP-DDA and the compared protocols (LBCH, AEA-FCP-EB, and SISR): (a) reporting dense events using the compared protocols, (b) reporting dense events using ERP-DDA, (c) reporting sparse events using the compared protocols, (d) reporting sparse events using ERP-DDA.

CH. In Figures 6 (c) and (d), we assume sparse events. As we can see, in Figure 6 (c), using the compared protocols to report these events requires sending eight packets to the CH. However, as shown in Figure 6 (d), ERP-DDA requires sending only two packets to CH. Therefore, the difference between the performance of ERP-DDA and that of the compared protocols is increased with increasing the distances between events.

5. Conclusion
In this paper, we proposed an efficient energy-aware data collection protocol for improving the lifetime of wireless sensor network. This protocol is called Event Reporting Protocol based on Distributed Data Aggregation (ERP-DDA). This protocol aims at reducing the consumed energy in the network by reporting the event that occurs in a certain region by only one sensor node that has the highest residual energy in that region. Moreover, ERP-DDA applies clustering and multi-hop routing schemes in order to reduce the total consumed energy in the network. Simulation results showed that ERP-DDA protocol, compared with other protocols, achieves higher performance in terms of extending the network lifetime.
References


19. Darabkh, K. A., Kasab, W. K., Khalifeh, A. F. Lim-AHP-G-C: Life Time Maximizing Based on Analytical Hier-


