Trends in Applied Sciences Research 8 (2): 55-72, 2013
ISSN 1819-3579 / DOI: 10.3923/tasr.2013.55.72

Data Dissemination and Collection Algorithms for Collaborative Sensor Devices Using Dynamic Cluster Heads

[1,2]Salah A. Aly, [1]Turki A. Alghamdi, [2]Mohamed Salim and [1,2]Adnan A. Gutub
[1]Smart Networks Laboratory, College of Computer and Information Systems (CIS),
[2]HajjCore, The Custodian of the Two Holy Mosques Institute of Hajj Research, Umm Al-Qura University, Makkah, KSA

Corresponding Author: Salah A. Aly, HajjCore, The Custodian of the Two Holy Mosques Institute of Hajj Research, Umm Al-Qura University, Makkah, KSA

ABSTRACT

Novel data dissemination and collection algorithms for Wireless Sensor Networks (WSNs) were developed in which n sensor nodes are distributed randomly in a certain field to measure a physical phenomenon. Such sensors have limited energy, short covering range, band width and memory constraints. It is desired to disseminate the sensed data throughout the network such that a base station will be able to collect the sensed data with high probability by querying a small number of nodes. Two data Dissemination and Collection Algorithms (DCA's) were proposed to solve the data collection and dissemination problem. In particular, data dissemination was achieved through dynamical selection of some nodes. The selected nodes will be changed after a time slot t and will be repeated after a period T. The simulation and performance results met the developed theoretical bounds.

Key words: Data dissemination, data collection, head nodes, PHDCA, RHDCA

INTRODUCTION

Wireless Sensor Networks (WSNs) are expanding rapidly due to various applications and ease of development. However, WSNs encounter several challenges to be deployed efficiently in a given environment. Such challenges are limited source of energy, limited transmission bandwidth, short covering range, data dissemination, data persistence, redundancy of defective nodes and data security. A typical Wireless Sensor Network (WSN) can be used in many applications such as monitoring physical phenomenon from the surrounding environment like temperature, gases, humidity, volcanoes and tornados. Also, it can be used in tracking animals, forest fire detection and military applications such as detection of enemy intrusion.

Many techniques are used in data dissemination (Imran *et al.*, 2010; Kokalj-Filipovic *et al.*, 2007) and cluster head election (Buttyan and Holczer, 2009; Liu *et al.*, 2007; Younis and Fahmy, 2004). Fountain codes and random walks have been used to disseminate data from k sources to a set of storage nodes n, (Kokalj-Filipovic *et al.*, 2008, 2009). LEACH algorithm (Handy *et al.*, 2002) is the most popular clustering algorithm. Lots of cluster head selection algorithms are based on LEACH architecture. The main drawback of the mentioned techniques is the requirement that all positions of all sensors must be known. Our algorithms don't use Fountain codes or random walks and independent on sensors positions.

This study considered a model for large-scale wireless sensor net- works with n identical sensing nodes distributed randomly and uniformly in a certain field. The nodes do not know the locations of the neighboring nodes as required (Dimakis *et al.*, 2006) and they don't maintain routing tables. In this work, two algorithms were proposed for data dissemination and data collection in wireless sensor networks. The first algorithm is Pre-known Heads for data Dissemination and Collection Algorithm (PHDCA). The second algorithm is Random Heads for data Dissemination and Collection Algorithm (RHDCA). The main aim was to develop an efficient method to randomly distribute and collect information from n sensors by querying 10-20% of nodes for retrieving information about all network nodes with a high probability. The main advantages of the proposed algorithms are as follows:

- No need for routing tables and the geographical positions of sensing nodes
- The possibility of controlling the amount of energy consumption in accordance with the desired application
- The algorithms are highly suitable for low data rates applications
- The base station can query 10-20% of the total nodes to retrieve information about all sensing nodes
- No synchronization is needed between network nodes

This study was organized as follows:

- **Section 1:** Background and short survey of the related work
- **Section 2:** Network model
- **Section 3:** Proposed DCA's algorithms
- **Section 4:** Some analysis for the DCA's algorithms
- **Section 5:** Simulation studies for the proposed algorithms. The conclusions were presented in Section IX

NETWORK MODEL

Network model was presented and described. For example: Consider a set of n identical sensing nodes distributed randomly in a field F of dimensions A = L×W, where L and W are the length and the width of F, respectively. It was assumed that each node has at least one neighboring node, meaning that with probability P = 1 there are no isolated nodes (Fig. 1):

- **Definition 1:** (Cluster head): The cluster head node (HN) is an arbitrary node among all network nodes N which exchanges its neighbors data with the other neighboring cluster head nodes
- **Definition 2:** (Node degree): The node degree d_{si} is the number of neighboring nodes to the node S_i within its coverage range. The average mean degree of all nodes in N is given by:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} d_{s_i} \tag{1}$$

The total period (T) is the period after which the sensed data has been disseminated in the network N and it is divided into equal time slots:

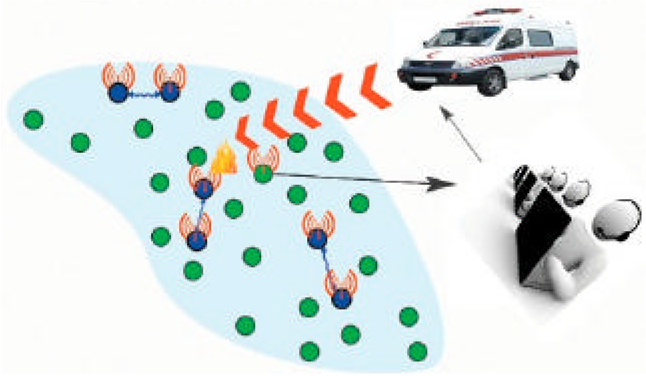

Fig. 1: A model for WSNs with n nodes distributed randomly and uniformly among them are k cluster head nodes with a blue color to illustrate data dissemination. The base station query some node to retrieve network data

$$T = \epsilon \times t \qquad (2)$$

for some integer number ϵ. The algorithm performance and simulation results confirm our theoretic bounds.

The head nodes consume more energy than other nodes due to excess transmissions needed for data dissemination and data collection. So, the head nodes are dynamically selected to apply fairness in energy consumption on all nodes. Also, the dynamical selection improves the performance of data dissemination in the network. The head nodes will be changed every time slot t. The number of head nodes in the network is k (where $k/n \cong \%10$). The selection of T depends on the desired application (i.e., T is small for high data rate applications and large for low data rate applications).

Assumptions: Let $S = \{S_1, \ldots .S_n\}$ be a set of n identical sensing nodes distributed randomly in a field F of dimensions $A = L \times W$, where L and W are the length and the width of F, respectively.

Let $H = \{h_1, \ldots .h_k\}$ be a set of k head nodes selected from the n sensing nodes to disseminate the data in the network and they will changed at each time slot t.

Let T be the period after which the sensed data has been disseminated in the network and it is divided into equal slots $t = \{t_1, \ldots ..t_e\}$.

The nodes use flooding to know their neighbors, as each node will send a message containing its ID_{s_i} to all neighboring nodes. Each node receives an incoming ID_{s_i} from any node s_i, it will consider the node of the incoming ID_{s_i} as its neighbor.

Each node in the network generates a packet P_{s_i} as follows:

$$P_{s_i} = (ID_{s_i}, x_{s_i}, \text{flag}) \qquad (3)$$

where, ID_{s_i} is the ID of the node s_i, x_{s_i} is the sensed data of the node s_i and flag is a variable set to 0 in flooding process or to 1 otherwise.

Each node has radio range coverage r_i.

The node s_i will be considered as a neighbor of s_j if and only if $d_{s_i,s_j} \le r_j$, where d_{s_i,s_j} is the distance between nodes s_i and s_j.

Initially, let the number of nodes n is known. Practically, the number of nodes in the network varies due to node energy depletion, failure nodes and added redundant nodes. Hence, it is important to estimate the number of nodes at each period T. The base station will consider the number of retrieved nodes when querying 10% of nodes as the total number of nodes n. The estimated number of nodes will be sent to the first survived head node (i.e., the first survived node from H) to be disseminated in the network. Figure 2 and 3 show the distribution of the head nodes at time slots t_0, t_1 in PHDCA and RHDCA, respectively.

DCA'S ALGORITHM

The DCA's algorithms were described in detail as follows.

PHDCA algorithm: This algorithm dynamically selected the *k* cluster head nodes that disseminate the data in the network according to a pre-known manner. The algorithm can be classified into four phases as follows:

- **Initialization phase:** In this phase, the head nodes are initially selected from $ID_{s_i} = 1:0.1\,n$ at the first time slot t_1

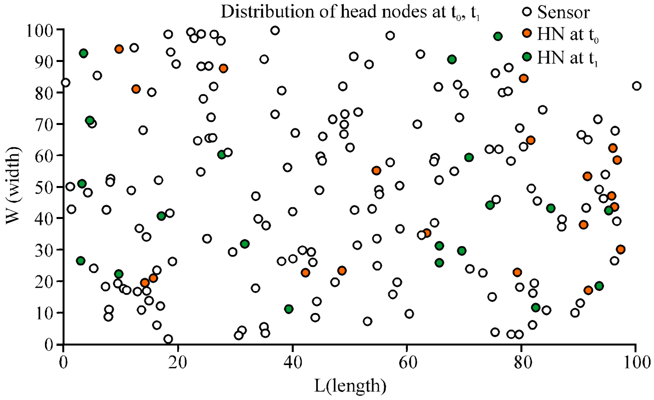

Fig. 2: The distribution of head nodes at time slots t_0, t_1 in PHDCA

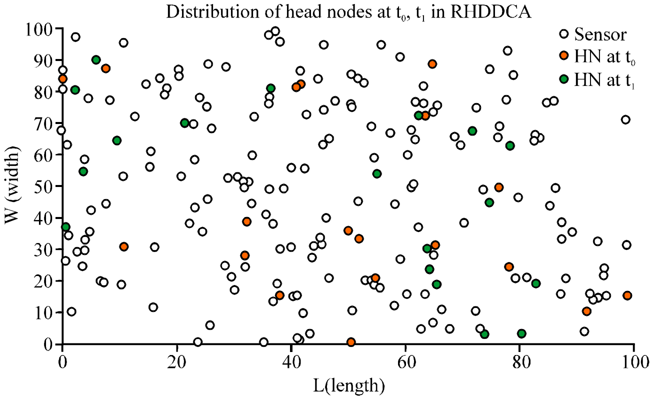

Fig. 3: The distribution of head nodes at time slots t_0,t_1 in RHDCA

- **Flooding phase:** In this phase, each sensor broadcasts a message containing its ID_{s_i} to be able to discover its neighbors to store them in its data base. If any node receives any incoming ID_{s_i}, it will consider the node of the incoming ID_{s_i} as its neighbor. Also, the broadcasting message containing a flag equal zero to indicate the flooding phase:

$$P_{s_i} = (ID_{s_i}, x_{s_i}, \text{flag} = 0) \tag{4}$$

- **Sensing and data dissemination phase:** In this phase, each sensor reads a new data, it will send this data to some of its neighboring nodes:

$$P_{s_i} = (ID_{s_i}, x_{s_i}, \text{flag} = 1) \tag{5}$$

The neighboring head nodes will disseminate data in the network by exchanging different data between them. The head nodes will be changed at each time slot and repeated each period T as shown in Algorithm 1. Therefore, dynamical selection of the head nodes improves the performance of data dissemination in the network.

- **Data collection phase:** In this phase, the base station can query small number of any nodes to retrieve the data sensed by the n sensing nodes and make estimation for n to send it to the first survived node

RHDCA algorithm: In PHDCA algorithm, it was assumed that the selection of head nodes is pre-known at each time slot t and the head nodes are repeated each period T. The disadvantage of this algorithm is the topology dependence. The performance of PHDCA depends on the distribution of head nodes. The PHDCA was extended to obtain RHDCA that randomly selects k head nodes at each time slot t. The performance of RHDCA is topology independent due to the randomly selection of head nodes. The main difference between PHDCA and RHDCA algorithms is the sensing and dissemination phase as follows:

- **Sensing and data dissemination phase:** In this phase, k head nodes are selected randomly at each time slot t. Each head node will have a status bit is set to 1 to indicate that the node is head node or to 0 otherwise:

$$P_{s_i} = (ID_{s_i}, x_{s_i}, \text{flag} = \times, \text{status} = 1) \tag{6}$$

Also, each sensor reads a new data which will be sent to its neighboring nodes. Neighboring head nodes will disseminate data in the network by exchanging the different data between them. Each current head node will select randomly one of its neighbors to be a head node at the next time slot t.

```
Input: A sensor network with $S = \{s_1, \ldots, S_n\}$ source nodes n source packet $x_{s_1}, \ldots, x_{s_n}$
Output: storage buffers $y_1, y_2, \ldots, y_n$ for all sensors S.
foreach node $u = 1{:}n$ do
Generate a packet containing $ID_u$, $flag = 0$ and broadcast this message to its set of neighbors;
$P_u = (ID_u, x_u, flag = 0)$
```

```
end
while there are surviving nodes do
    if t expired then
        Generate new k head nodes as follows:
        t + +;
        foreach node u = 1 ∈ n do
            if $(t-1)n/\in < u \leq tn/\in$ then
                u is a head node;
            end
        end
        if t = = ∈ then
            t = 0;
            Update n by the received estimated node number from base station.
        end
    end
    foreach node u = 1 ∈ n do
        if u sensed new data then
            u sends this data to some of its neighbors randomly;
            foreach node $v \in N(u)$ do
                if Rand(1)>0.5 then
                    u sends $P_u$ to v
                end
            end
        end
    end
    foreach head node h= 1∈k do
        h exchanges the different data with its neighboring head nodes;
        foreach head node $v \in N(h)$ do
            if Rand(1)>0.5 then
                h Sends different information to v;
            end
        end
    end
end
```

Algorithm 1: PHDCA algorithm: Data dissemination algorithm for WSNs using dynamic deterministic cluster head nodes.

DCA'S ANALYSIS

Here, analyzed the proposed DCA's algorithms.

Lemma 3: The probability that a set M of sensors has at least one cluster head node is given by:

$$\Pr(M \cap H) = 1 - \prod_{i=1}^{m}\left(1 - \frac{k}{n-i+1}\right) \tag{7}$$

where, m = |M| is the number of nodes in M.

```
Input: A sensor network with $S = \{s_1, \ldots, s_n\}$ source nodes, n source packets $x_{s_1}, \ldots, x_{s_n}$
Output: storage buffers $y_1, y_2, \ldots, y_n$ for all sensors S
foreach node u = 1 : n do
    Generate a packet contains $ID_u$, flag = 0 and broadcast this message to its set of neighbors.;
    $P_u = (ID_u, x_u, flag = 0, status = x)$;
end
```

```
while there are surviving nodes do
    if t expired then
        Generate new k head nodes randomly as follows:
        foreach current head node h = 1 ∈ k do
            h selects randomly a node v from its neighbors to be a new head node;
            $ID_v = round(Rand(1).|N(h)|)$;
            $status_h = o$;
            $P_h = (ID_h, x_h, flag = \times, status = 0)$;
            $status_v = 1$;
            $P_v = (ID_v, x_v, flag = \times, status = 1)$;
        end
    end
    foreach node u = 1 ∈ n do
        if u sensed new data then
            u sends this data to some of its neighbors randomly.;
            foreach head node h = 1∈k do
                h sends its neighbors data to its neighboring head nodes;
                foreach head node V∈N(h) do
                    if Rand (1)>0.5 then
                        h Sends different information to v;
                    end
                end
            end
end
```

Algorithm 2
RHDCA algorithm: Data dissemination algorithm for WSNs using dynamic random cluster head nodes.

Proof: Number of ways in which the m nodes can be drawn from the total number of nodes n is:

$$\binom{n}{m} = C_m^n = \frac{n!}{m!(m-n)!}$$

Number of ways so that no head nodes exist in the set M is $\binom{n-k}{m}$. So, the probability that the set M has no cluster head nodes is $\binom{n-k}{m}/\binom{n}{m}$. Hence, the probability that the set M has at least one head node is:

$$1-\frac{\binom{n-k}{m}}{\binom{n}{m}} = 1-\prod_{i=1}^{m}(1-\frac{k}{n-i+1})$$

Lemma 4: The probability that a set M of sensors has a set Z of cluster head nodes is given by:

$$\Pr(Z) = \frac{\binom{n-k}{m-z}\binom{k}{z}}{\binom{n}{m}} \qquad (8)$$

where, z = |Z| is the number of nodes in Z.

Proof: Number of ways in which the m nodes can be drawn from the n sensing nodes is $\binom{n}{m}$. From the Fundamental Counting Theorem, the total number of ways in which z head nodes and m-z non head nodes can be drawn from the n sensing nodes is $\binom{n-k}{m-z}/\binom{k}{z}$. So, the probability that a set of n sensor has z head nodes is:

$$\frac{\binom{n-k}{m-z}/\binom{k}{z}}{\binom{n}{m}}$$

Definition 5: (Head energy consumption (E_h)): is the energy consumption at the nodes n due to data dissemination in the network N when all nodes have the same coverage range and packet size.

Lemma 6: Let β be the probability that a node s_i has a set Z of neighboring head nodes. From Eq. 8, β can be given by:

$$\beta = \frac{\binom{n-k}{d_{s_i}-z_{s_i}}\binom{k}{z_{s_i}}}{\binom{n}{d_{s_i}}}$$

where, z_{si} is the number of neighboring head nodes to node s_i and d_{s_i} is the degree of node s_i when the set M represents the neighboring nodes of the node s_i.

The total energy consumption E_h is given by:

$$E_h = \frac{\in k}{n}(n\mu p_r + p_t \textstyle\sum_{i=1}^{n} \beta\alpha_i z_{s_i}) \tag{9}$$

where, α_i is the number of transmissions between the node s_i and its neighbors and p_t, p_r are the transmitted and received energy costs due to sending one packet.

Proof: Let σ be the received energy cost of nodes n due to data dissemination, so:

$$\sigma = \frac{k}{n}\sum_{i=1}^{n} d_{s_i} p_r = k\mu p_r$$

where, k/n is the probability that a node s_i is a head node. Let ξ be the transmitted energy cost of n nodes due to data dissemination, so:

$$\xi = \frac{k}{n}\sum_{i=1}^{n} \beta\alpha_i z_{s_i} p_t = \frac{kp_t}{n}\sum_{i=1}^{n} \beta\alpha_i z_{s_i}$$

Therefore, the total energy consumption due to data dissemination at time slots is given by:

$$E_h = \in(\sigma + \varsigma) = \frac{\in k}{n}\left(n\mu p_r + p_t \sum_{i=1}^{n} \beta\alpha_i z_{s_i}\right)$$

Lemma 7: The total energy consumption at the sensing nodes due to sending the sensed data to their neighbors is given by:

$$E_s = n(p_t+\mu p_r) \tag{10}$$

where all nodes have the same coverage range and packet size.

Proof: The energy consumption at nodes n due to sending its sensed data is np_t. The energy consumption at nodes n due to all received packets is. $\sum_{i=1}^{n} p_r \times d_{s_i}$. Hence, assuming that each node updates its data one time at each period T, the energy consumption at the n sensing nodes is:

$$E_S = \sum_{i=1}^{n} (p_t + d_{s_i} p_r) = n(p_t + \mu p_r)$$

Lemma 8: The optimum number of head nodes that gives minimum energy consumption is given by:

$$k_{opt} = \sqrt{\left| \frac{nE_h}{\in \lambda \sum_{i=1}^{n} \alpha_i z_{s_i}} \right|} \tag{11}$$

Where:

$$\lambda = \frac{d}{dk} \sum_{i=1}^{n} \beta$$

Proof: The optimum number of head nodes that gives minimum energy consumption can be driven by the differentiation of Eq. 9 as follows:

$$\frac{d}{dk_{opt}}(E_h) = 0$$

Lemma 9: The total number of nodes that achieves a certain value of the average mean degree of all nodes μ which indicates the network density, when they distributed randomly and uniformly in a certain field of region $A = L\times W$ is given by:

$$n \cong \mu \frac{L \times W}{a_2} \tag{12}$$

where, $\pi r^2 \leq a \leq (2r)^2$.

Proof: As the nodes are distributed uniformly in the field, an arbitrary area of dimensions $\pi r^2 \leq a \leq (2r)^2$ will contain μ nodes. Hence, the total number of nodes is:

$$n \cong \mu \frac{L \times W}{a^2}$$

Therefore, Lemma 9 illustrates the relation between total number of nodes n in the network N, mean degree of graph μ, coverage area of sensors and the field area.

SIMULATION AND PERFORMANCE EVALUATIONS

In this section we will demonstrate some simulation results to illustrate the performance of the proposed algorithms.

Definition 10: Decoding Ratio (η) is the ratio between the number of queried nodes $\hat{n}$ and the total number of sources n:

$$\eta = \frac{\hat{n}}{n} \tag{13}$$

Definition 11: Successful Decoding Probability (P_s) is the probability that the n source packets are all recovered from the $\hat{n}$ querying nodes.

Figure 4 and 6 show the relation between the successful decoding probability and the decoding ratio for different values of sensing nodes n in PHDCA and RHDCA algorithms. Increasing the number of network nodes n and fixing the covering radius r of all nodes will result in an improvement in the successful decoding probability as well. We can notice that as the number of nodes increases, the ratio of queried sensors will be decreased to recover data with a reasonable successful probability. Particularly, for n >500, we see that querying up to 10% will reveal about 85% of network data in PHDCA and about 92% of network data in RHDCA.

Figure 5 and 7 show the amount of energy consumption at each node after the dissemination of data in the network N in PHDCA and RHDCA algorithms. From these figures, it can be noticed that energy consumption in PHDCA algorithm is better than the obtained result in RHDCA algorithm. We assumed that energy consumption at the sensing node due to sensing the data itself is neglected and each sensor node is assumed to be of initial battery charge 5 Joule. The energy consumption was calculated according to Meghanathan *et al.* (2010). They assumed that the energy lost at a sensor node s_i due to transmission of one packet is given by:

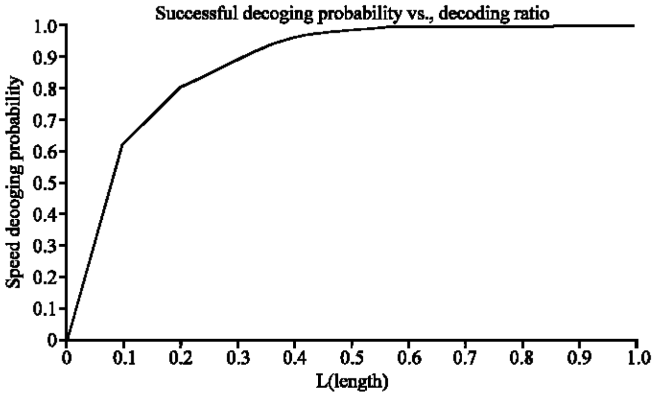

Fig. 4: Thes figure shows the relation between the successful decoding probability and the decoding ratio for n = 100, n = 200, n = 300, n =400, n = 500 when A = 100×100 and r = 10 in PHDCA

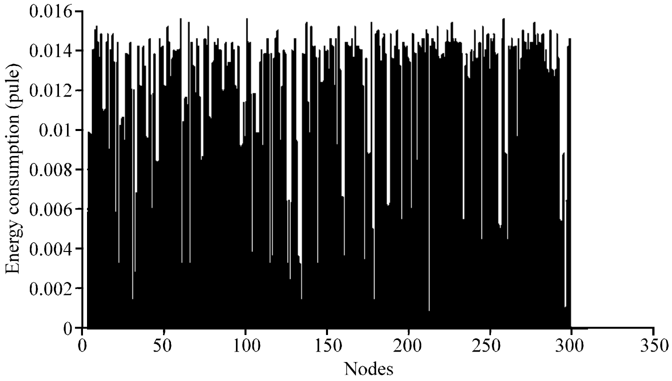

Fig. 5: Energy consumption at each node in network N in PHDCA after period T

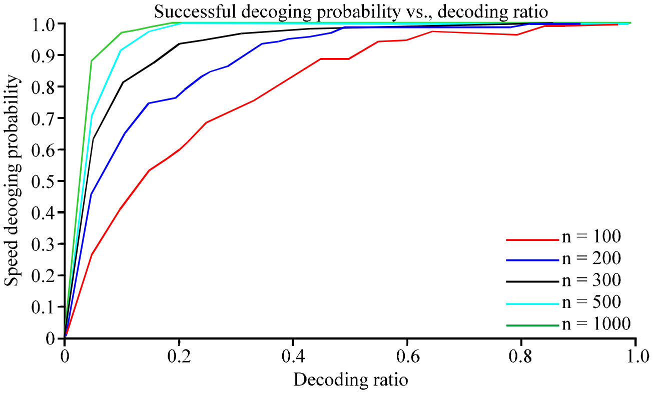

Fig. 6: This figure shows relation between the successful decoding probability and the decoding ratio for n = 100, n = 200, n = 300, n =500, n = 1000 when A = 100 * 100 and r = 10 in RHDCA

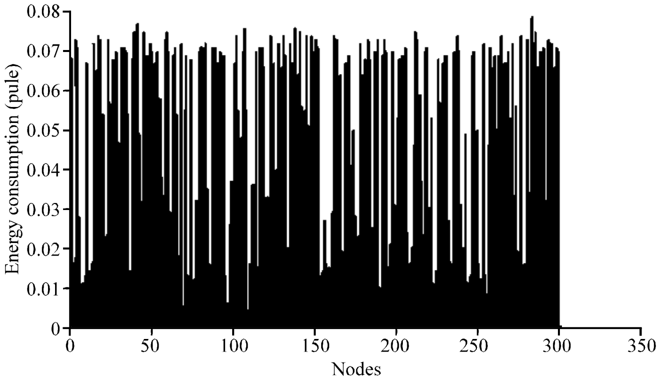

Fig. 7: Energy consumption at each node in network N in RHDCA after period T

$$p_t = (50\times10^{-9}+100\times10^{-12}\times r^2_{s_i})\times c_{s_i} \tag{14}$$

and the energy lost at a sensor node s_i due to receiving of one packet is given by:

$$p_r = 50\times10^{-9}\times c_{s_i} \tag{15}$$

where C_{si} is the packet size of node s_i.

Definition 12: Death Rate (DR) is the ratio between the number of dead nodes $\bar{n}$ and the total number of sensing nodes n:

$$DR = \frac{\bar{n}}{n} \tag{16}$$

Figure 8 and 9 illustrate the relation between the death rate and the total number of sensing nodes. Increasing the number of network nodes n and fixing the field area will result in an increasing in the death rate as well due to the excess transmissions needed for data dissemination.

Figure 10 and 11 show the relation between the performance of data collection and the elapsed time in DCA's algorithms. As the elapsed time increases, more nodes disappear from the network

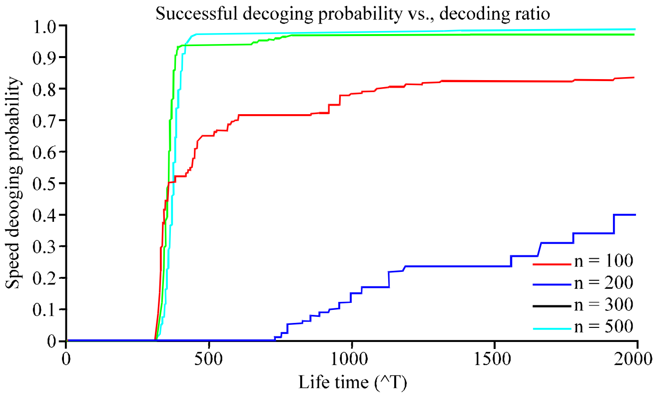

Fig. 8: Relation between the death rate and number of nodes n in PHDCA

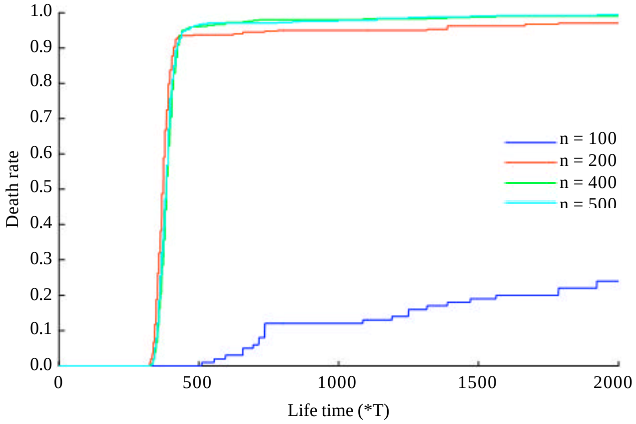

Fig. 9: Relation between the death rate and number of nodes n in RHDCA

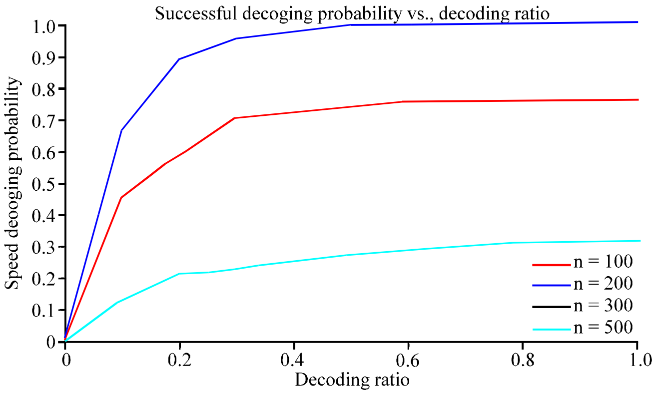

Fig. 10: Performance of PHDCA algorithm along time

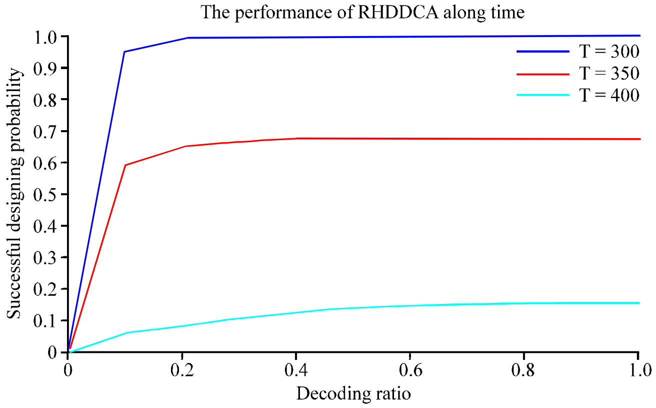

Fig. 11: Performance of RHDCA algorithm along time

N (i.e., DR increases) due to energy depletion. Hence, data dissemination and data collection performances will be negatively affected by the disappeared nodes $\bar{n}$. Although the performance of RHDCA algorithm is better than the performance of PHDCA algorithm at the beginning. The contrary will happen after a certain time (i.e., the elapsed time reaches a certain threshold). Also from these figures, we can deduce that the network life time depends on the period T, so if T is selected to be a large value, it will lead to a significant improvement in saving the amount of energy consumption. Therefore, the proposed algorithms are very suitable for low data rate applications such as temperature and humidity monitoring in a certain region. This is besides that the algorithms are also applicable to high data rate applications (i.e., T may be in seconds for high data rate applications and in minutes for low data rate applications). Hence, the period T can be used to control the amount of energy consumption in accordance with the intended applications.

COMPARISON

This section provided evaluation and comparison analysis between PHDCA and RHDCA algorithms. PHDCA has low energy consumption and fixed data dissemination performance over

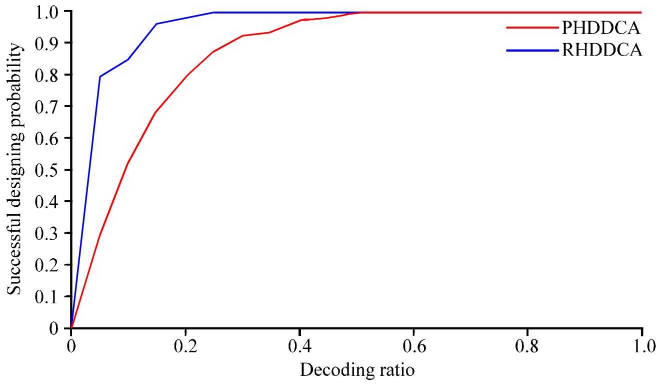

Fig. 12: A comparison between the performance of PHDCA and RHDCA algorithms

different periods on account of the periodic selection of cluster head nodes. This is in contrast to RHDCA which consumes more energy and has non predictable data dissemination performance over different periods due to the random selection of cluster head nodes. PHDCA has frustrating data dissemination performance with regard to RHDCA on account of the dependence on the topology of the pre-known cluster head nodes. RHDCA is suitable for the applications that require high data dissemination performance and PHDCA is suitable for the applications that have energy limitations.

Figure 12 shows a comparison between the data dissemination performance of PHDCA and RHDCA. The figure shows that the data collection performance of RHDCA is better than the performance of PHDCA as the probability of successful de- coding in RHDCA is higher than the probability of successful decoding in PHDCA for the same decoding ratio.

RELATED WORK

In this section we will indicate the related work to our work.

- The authors in Aly *et al.* (2011) proposed a distributed data collection algorithm to store and forward information obtained by wireless sensor networks. They used n-k storage nodes to collect the sensed data from the network, where *k* is the sensor nodes, n is the total number of nodes and (n-k)/n is 20%
- The authors in Aly *et al.* (2009), Kong *et al.* (2010), Aly *et al.* (2008) suggested two distributed storage algorithms for large-scale wireless sensor networks. They assigned a time-to-live counter to each node packet depending on its degree. According to this counter, each packet can travel to a certain number of hops. Each node chooses randomly one of its neighbors to send its data to another neighbor. The receiver node will decide with a random probability if it will accept the incoming message or not. The base station can query about 20%- 30% of the network nodes in order to retrieve the data collected by the n sensing nodes
- The authors in Dimakis *et al.* (2005) used a decentralized implementation of Fountain codes that uses geographic routing and every node has to know its location
- The authors in Kamra *et al.* (2006) proposed a novel technique called growth codes to increase data persistence in wireless sensor networks, i.e. increasing the amount of information that can be recovered at the sink

- The authors in Dimakis *et al.* (2010) presented a general theoretic framework that can determine the information that must be communicated to repair failures in encoded systems and identified a tradeoff between storage and repair bandwidth
- Authors in Heinzelman *et al.* (2002) used a central controller to select CH nodes throughout the network. The main drawbacks of this algorithm are non-automatic cluster-head selection and the requirement that the position of all sensors must be known
- Authors in Handy *et al.* (2002) extended LEACH stochastic algorithm with a deterministic cluster-head selection, which utilizes the remaining energy level of each node to determine the threshold
- The authors in Younis and Fahmy (2004) proposed a distributed clustering scheme HEED (Hybrid Energy- Efficient Distributed Clustering) in which CH nodes are picked from the deployed sensors. HEED considers a hybrid of energy and communication cost when selecting CHs
- The authors in Meghanathan *et al.* (2010), Eriksson *et al.* (2008) used mobile sinks to obtain potential energy savings for the sensors during data dissemination in wireless sensor networks. Each sink node is assigned a particular cluster of sensors to monitor and collect data. A sink node moves to the vicinity of the sensor nodes (within a few hops) to collect data. The collected data is exchanged with peer mobile sinks and can also be transferred to a control center through multi-hop sink-to-sink data propagation
- The authors in Bandyopadhyay and Coyle (2003) proposed a distributed, randomized clustering algorithm to organize the sensors in a wireless sensor network into clusters. They extended this algorithm to generate a hierarchy of cluster heads and observe that the energy savings increase with the number of levels in the hierarchy
- The authors in Buttyan and Holczer (2009) presented simple protocol suitable for both locations based and topology based clustering. Also, they proposed a useful extension to this basic protocol. They showed how to fine tune its parameters such that the amount of information leaked by the protocol about the identity of the cluster heads is minimized
- The authors in Lin *et al.* (2007a) considered cluster-based architecture and provided distributed clustering algorithms for mobile sensor nodes which minimize the energy dissipation for data-gathering in a wireless mobile sensor network
- The authors in Imran *et al.* (2010) proposed a gossip based protocol that consumes little resources. The proposed scheme aimed to keep the routing table size R as low as possible
- The authors in Lin *et al.* (2007b) proposed priority random linear codes, to maintain measurement data in different priorities, such that critical data have a higher opportunity to survive node failures than data of less importance
- The authors in Yu *et al.* (2009) proposed optimal data storage (ODS) algorithms that can produce global optimal data storage position in linear, grid, and mesh network topologies. To reduce the computation of ODS in the mesh network topology, they presented a near-optimal data storage (NDS) algorithm, which is an approximation algorithm and can obtain a local optimal position
- The authors in Kokalj-Filipovic *et al.* (2009) studied decentralized, Fountain and network-coding based strategies to allow for a reduced delay collection by a data collector who accesses the network at a random position and random time. Data dissemination is performed by a set of relays which form a circular route to exchange source packets

Fig. 13: Wireless sensor devices are deployed in Minna field for gas monitoring, temperature monitoring and crowd sensing. Approximately 50.000 camp tents are located in Mina to accommodate 3-5 million people for 4-8 days during pilgrimage, according to 2010 KSA statistics

PRACTICAL ASPECTS

The proposed algorithms are suitable to be applied on the American-made camp tents in Minna and Arafat fields located in the east south of Makkah, KSA for monitoring and measuring certain phenomenon such as temperature, gases, pollution and crowd estimation. Approximately 50.000 camp tents are located in Minna to accommodate 3-5 million people for 4-8 days during pilgrimage, according to 2010 KSA statistics. Hence, a safety system is needed to protect the camp tents against fires and pollution. Wireless sensor devices can be scattered in Minna field to gather and collect the required data to be monitored at the base station to take the safety precautions at emergency cases as shown in Fig. 13. Also, The Wireless sensor devices can detect the crowded areas and inform the base station about the non-crowded areas to be exploited.

CONCLUSION

This study presented two algorithms for data dissemination and collection in wireless sensor networks. Given n sensing nodes with limited buffers. The study demonstrated schemes to disseminate sensed data throughout the network with less computational overhead. The proposed algorithms did not assume any pre-known of routing tables or nodes locations. In addition, the time factor T increases the network life time, as it can be selected to be suitable for the intended applications and minimizing energy consumption. Our future work will develop accurate practical algorithms to optimize energy consumptions in the sensor network.

ACKNOWLEDGMENTS

This study is funded by a grant from the Long-Term National Plan for Science, Technology and Innovation (LT- NPSTI), No. 11-INF1702-10, the King Abdulaziz City for Science and Technology (KACST), Kingdom of Saudi Arabia. We thank the Science and Technology Unit at Umm Al-Qura University for their continued logistics support.

REFERENCES

Aly, S A., Z. Kong and E. Soljanin, 2008. Raptor codes based distributed storage algorithms for wireless sensor networks. Proceedings of the IEEE International Symposium on Information Theory, July 6-11, 2008, Toronto, Canada, pp: 2051-2055.

Aly, S.A., A. Ali-Eldin and H.V. Poor, 2011. A distributed data collection algorithm for wireless sensor networks with persistent storage nodes. Proceedings of the 4th IFIP International Conference on New Technologies, Mobility and Security, February 7-10, 2011, Paris, France, pp: 1-5.

Aly, S.A., M. Youssef, H.S. Darwish and M. Zidan, 2009. Distributed flooding-based storage algorithms for large-scale wireless sensor networks. Proceedings of the IEEE International Conference on Communications, June 14-18, 2009, Dresden, Germany, pp: 1-5.

Bandyopadhyay, S. and E.J. Coyle, 2003. An energy efficient hierarchical clustering algorithm for wireless sensor networks. Proceedings of the 22nd Annual Joint Conference of the Computer and Communications, Volume 3, March 30-April 3, 2003, San Franciso, CA., USA., pp: 1713-1723.

Buttyan, L. and T. Holczer, 2009. Private cluster head election in wireless sensor networks. Proceedings of the 6th International Conference on Mobile Adhoc and Sensor Systems, October 12-15, 2009, Macau, China, pp: 1048-1053.

Dimakis, A.G., V. Prabhakaran and K. Ramchandran, 2005. Ubiquitous access to distributed data in large-scale sensor networks through decentralized erasure codes. Proceedings of the IEEE 4th International Symposium on Information Processing in Sensor Networks, April 25-27, 2005, Los Angeles, CA., USA., pp: 111-117.

Dimakis, A.G., V. Prabhakaran and K. Ramchandran, 2006. Distributed fountain codes for networked storage. Proceedings of the International Conference on Acoustics, Speech and Signal Processing, May 14-19, 2006, Toulouse, France.

Dimakis, A.G., P.B. Godfrey, Y. Wu, M.J. Wainwright and K. Ramchandran, 2010. Network coding for distributed storage systems. IEEE Trans. Inform. Theory, 56: 4539-4551.

Eriksson, J., L. Girod, B. Hull, R. Newton, S. Madden and H. Balakrishnan, 2008. The pothole patrol: Using a mobile sensor network for road surface monitoring. Proceedings of the 6th International Conference on Mobile Systems, Applications and Services, June 17-20, 2008, Breckenridge, Colorado, pp: 29-39.

Handy, M.J., M. Haase and D. Timmermann, 2002. Low energy adaptive clustering hierarchy with deterministic cluster-head selection. Proceedings of the 4th International Workshop on Mobile and Wireless Communications Network, September 9-11, 2002, Stockholm, Sweden, pp: 368-372.

Heinzelman, W.B., A.P. Chandrakasan and H. Balakrishnan, 2002. An application specific protocol architecture for wireless microsensor networks. IEEE Trans. Wireless Commun., 1: 660-670.

Imran, N., S. Khan and I. Rao, 2010. A trustworthy and well-organized data disseminating scheme for ad-hoc WSNS. Int. J. Comput. Networks Commun., 2: 170-181.

Kamra, A., V. Misra, J. Feldman and D. Rubenstein, 2006. Growth codes: Maximizing sensor network data persistence. Proceedings of the Conference on Applications, Technologies, Architectures and Protocols for Computer Communications, September 11-15, 2006, Pisa, Italy, pp: 255-266.

Kokalj-Filipovic, S., P. Spasojevic and E. Soljanin, 2009. Doped fountain coding for minimum delay data collection in circular networks. IEEE J. Sel. Areas Commun., 27: 673-684.

Kokalj-Filipovic, S., P. Spasojevic, R. Yates and E. Soljanin, 2008. Decentralized fountain codes for minimum-delay data collection. Proceedings of the 42nd Annual Conference on Information Sciences and Systems, March 19-21, 2008, Princeton, NJ., USA., pp: 545-550.

Kokalj-Filipovic, S., R. Yates and P. Spasojevic, 2007. Infrastructures for data dissemination and in-network storage in location-unaware wireless sensor networks. Technical Report, WINLAB.

Kong, Z., S.A. Aly and E. Soljanin, 2010. Decentralized coding algorithms for distributed storage in wireless sensor networks. IEEE J. Sel. Areas Commun., 28: 261-267.

Lin, Y., B. Li and B. Liang, 2007a. Differentiated data persistence with priority random linear code. Proceedings of the International Conference on Distributed Computing Systems, June 25-27, 2007, Toronto, Canada, pp: 47.

Lin, Y., B. Liang and B. Li, 2007b. Data persistence in large-scale sensor networks with decentralized fountain codes. Proceedings of the 26th Annual Joint Conference on Computer and Communication, May 6-12, 2007, Anchorage, AK., USA., pp: 1658-1666.

Liu, C.M., C.H Lee and L.C. Wang, 2007. Distributed clustering algorithms for data-gathering in wireless mobile sensor networks. J. Parallel Distrib. Comput., 67: 1187-1200.

Meghanathan, N., S.K. Sharma and G.W. Skelton, 2010. On energy efficient data dissemination in wireless sensor networks using mobile sinks. J. Theor. Applied Inform. Technol., 19: 79-91.

Younis, O. and S. Fahmy, 2004. HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans. Mobile Comput., 3: 366-379.

Yu, Z., B. Xiao and S. Zhou, 2009. Achieving optimal data storage position in wireless sensor networks. Comput. Commun., 33: 92-102.