

Balanced Multi-Channel Data Collection in Wireless Sensor Networks

Zahra Adelani¹, Ghasem Mirjalily², and Milad HajMirzaei³

¹Department of Computer Engineering, Yazd Science and Research Branch, Islamic Azad University, Iran

²Department of Electrical Engineering, Yazd University, Iran

³School of Electrical and Computer Engineering, Shiraz University International Division, Iran

Abstract: Data collection is an essential task in Wireless Sensor Networks (WSNs). In data collection process, the sensor nodes transmit their readings to a common base station called Sink. To avoid a collision, it is necessary to use the appropriate scheduling algorithms for data transmission. On the other hand, multi-channel design is considered as a promising technique to reduce network interference and latency of data collection. This technique allows parallel transmissions on different frequency channels, thus time latency will be reduced. In this paper, we present a new scheduling method for multi-channel WSNs called Balanced Multi Channel Data Collection (Balanced MC-DC) Algorithm. The proposed protocol is based on using both Non-Overlapping Channels (NOC) and Partially Overlapping Channels (POC). It uses a new approach that optimizes the processes of tree construction, channel allocation, transmission scheduling and balancing simultaneously. Extensive simulations confirm the superiority of the proposed algorithm over the existing algorithms in wireless sensor networks.

Keywords: Channel Assignment, Data Collection, Load Balancing, Multi-Channel, Scheduling, Wireless Sensor Networks.

1. Introduction

A Wireless Sensor Network (WSN) typically consists of a set of sensor nodes deployed in a geographical area. Each sensor measures physical parameters such as temperature, motion, etc. [1]. These sensor nodes are connected to each other through radio links and cooperate with each other to accomplish a common task named data collection. In data collection process, each node sends its recorded data to the sink. Sensors deployed in WSN are usually cheap, small, and most importantly battery-powered. According to these features, the implemented protocols should meet the objectives of saving energy and expanding the network lifetime.

One of the major sources for energy consumption in data collection is collisions between transmissions of multiple nodes. Therefore, in heavy traffic scenarios, using non contention-based Media Access Control (MAC) protocols, such as Time Division Multiple Access (TDMA), plays a vital role in increasing the network lifetime. These types of MAC eliminates collisions, idle listening and overhearing that are the major sources of energy consumption in wireless communication. Moreover, they solve the problem of hidden terminal by scheduling the interfering transmissions at different times, as well. In TDMA scheduling, time is slotted and the length of each slot is considered in a way a packet is completely transmitted or received.

In real time monitoring and control applications (e.g. forest fire detection, gas and oil leak detection, and battlefield surveillance), the latency of data collection is an important factor. In this way, the main problems are optimal design of

collection tree, channel assignment and scheduling process so that the sink collects data from all sensor nodes without collision and with minimum number of time slots. Here, data collection latency is defined as the number of time slots that is required for collecting data from all network nodes by the sink.

Many studies have been done on optimization of data collection process in WSNs. In most available works, for example [2] and [3], nodes use the same radio channel for data transmission. Using single channel causes long latency in collecting data and also reduces the performance of the sensor network. In recent years, multi-channel communication is considering as a technique that reduces the data collection latency in WSNs by increasing the number of parallel transmissions on distinct frequency channels.

Multi-channel approaches are classified into two categories. First approaches use Non-Overlapping Channels (NOC) [4-5], while second approaches use Partially Overlapping Channels (POC) [6-9]. For example, IEEE 802.11b/g standard defines 11 channels in 2.4 GHz ISM band. Each channel has a bandwidth of 22MHz and the distance between central frequencies of two adjacent channels is 5MHz. Thus, a channel is overlapped to several adjacent channels. According to this standard, if the separation of the channels is greater than four, the two channels are non-overlapping. Therefore, there are at most three NOCs (channels 1, 6 and 11). Due to the limited number of non-overlapping channels, interference will not be fully removed. The performance can be improved by using Partially Overlapping Channels.

In this paper, we propose a data collection algorithm for wireless sensor networks named Balanced Multi Channel Data Collection algorithm (Balanced MC-DC). We describe in brief our contributions in this paper as follows:

- The proposed data collection algorithm is multi-channel. It increases the spectrum utilization by simultaneous use of available NOCs and POCs.
- The key feature of the proposed algorithm is joint optimization of the data collection tree, channel assignment and scheduling.
- To improve the energy efficiency and to increase the lifetime of the network, we use load-balancing techniques to construct a load-balanced data collection tree.
- We evaluate the proposed algorithm with extensive simulations using the C#.Net programming language to study the performance of the Balanced MC-DC in terms of network lifetime and data collection latency. The results show the superiority of the Balanced MC-MLAS algorithm compared to the similar multi-channel

algorithms proposed by Wang and Jia for using POCs [8] and Ghods, et al. for using a combination of NOCs and POCs [10].

The rest of paper is organized as follows. We summarize the previous work in Section 2. Section 3 contains system model description and assumptions. In Sec. 4, the proposed data collection algorithm is introduced. In Sec. 5, we evaluate performance of the proposed algorithm by simulation. The paper concludes in Section 6.

2. Related Work

Improvement of the scheduling and routing algorithms for data collection in WSNs is an active research topic. Florens et al. [11-13], assuming the protocol interference model, proposed some optimized centralized scheduling algorithms for specific topologies such as linear, multi-linear and tree for both directional and omnidirectional antennas. In [14], authors proposed an energy-efficient distributed scheduling algorithm named Clu-DDAS based on a cluster-based aggregation tree. also they considered an adaptive strategy for updating the schedule to accommodate dynamic network topology. In [15], Neamatollahi et al. proposed a fuzzy-based hyper round policy (FHRP) to efficiently and flexibly schedule the clustering task. In FHRP, clustering is performed at the beginning of every Hyper Round (HR), which is composed of many rounds. Lai et al. [16] proposed a scheduling approach based on Virtual Node Expansion. In this approach, graph coloring is used to find the minimum scheduling length. Also in [17] and [18], some centralized scheduling algorithms (Node-based scheduling, Level-based scheduling and Congestion-based scheduling) are proposed using graph coloring.

Gandham et al. [19, 20] suggested a distributed scheduling algorithm for raw data collection. They formulized the problem as Integer Linear Programming (ILP) and proposed a time slot allocation scheme, as well. In [20], authors not only considered minimizing the length of scheduling, but also considered the memory restrictions of sensor nodes in their proposed algorithm. Incel et al. [21] showed that the data collection rate is often limited by routing topologies. In this way, they proposed using trees with specific properties to improve the data collection rate.

In the algorithms presented in [22], a parent node does not wait for receiving data from all of its children in a single frame before the transmission. This approach is executable for continuous and periodic monitoring applications which are stable for a long time. In [23], authors proposed an efficient packet scheduling technique for data merging in WSNs to reduce the number of transmissions. This technique is used to merge the data packets to the same destination in intermediate nodes. It appends the merged packets with received packets until the maximum packet size or maximum waiting time is reached. Real-time data packets are directly forwarded to the next node without merging. Incel et al. [4] proposed using different techniques to provide some successive improvements. For reducing the interference, they utilized using transmission power control. Also, they used multiple orthogonal frequency channels in order to activate more simultaneous transmissions.

In a WSN with low traffic, the lost energy in nodes due to switching the modes (sleep, awake, and idle) is non-negligible. In [24, 25], authors proposed centralized algorithms that minimize the number of switching between modes using TDMA-based scheduling. In [26], a MAC protocol with On-demand Convergecast Scheduling (OCS) principle is presented. It is a multi-hop, adaptive and centralized TDMA protocol, which supports convergecast applications. In this protocol, there is no need to collect data from all network nodes at the time of the event. Only the nodes in the event area transmit their data to the sink. This algorithm improves the network throughput, energy consumption and the scheduling length, but it is applicable only for event-driven applications.

Traditionally, WSNs have been deployed with a single sink. Due to the emergence of sophisticated applications, WSNs may require more than one sink. The main contribution of [27] is the development of a distributed data aggregation scheduling (DAS) algorithm for WSNs with two sinks. they also proposed a distributed energy-balancing algorithm to balance the energy consumption for the aggregators. then, they proved that their algorithm generates low latency, compared to an algorithm that have been developed for a WSN with a single sink.

The work in [7] is the first attempt to use POCs in multi-channel radios systematically. There is mentioned that how, in most cases, POCs causes significant progress in spectrum utilization and performance improvement. Cui et al. [28] proposed a greedy algorithm called Minimum Interference for Channel Allocation (MICA), to minimize the total interference in wireless networks. This algorithm has considered physical distance and separation of adjacent channels between two communicating nodes as interference metrics. In addition, authors proved that when two nodes are physically separated, they could be regarded as orthogonal even if they use adjacent channels.

In [8], for the first time, Wang et al. proposed a POC approach to minimize the length of scheduling and data collection latency in WSNs. The primary idea of this algorithm is scheduling a node in a time slot that is already used by other nodes, by proper selection of the parent and channel for that node. However, this method cannot use full capacity of the spectrum. Jin et al. [29] proposed a convergecast scheduling algorithm for industrial WSNs with multiple radio interfaces in order to reduce data collection latency. then, based on our proposed scheduling algorithm, they proposed two algorithms to minimize the number of radio interfaces under the temporality constraint of industrial production.

Recently, in [10], authors proposed Multi-channel Minimum Latency Aggregation Scheduling protocol (MC-MLAS) which is a TDMA-based MAC protocol to minimize the total latency in data collection. For first time, this method used both NOCs and POCs in order to minimize the length of scheduling and latency. Although this method benefits from the advantages of both types of channels in order to maximize the data collection rate, it does not consider the load balancing in data collection tree. Load balancing is an important factor in increasing the lifetime of the network.

In this paper, we introduce a modified version of the MC-MLAS algorithm named Balanced MC-DC. The proposed algorithm not only benefits from the advantages of using both

types of channel allocation (i.e., NOCs and POCs) to reduce the data collection latency, but also considers load balancing in tree construction to improve the energy efficiency and to increase the lifetime of WSNs.

Balanced MC-DC is a Multi-channel Time Division Multiple Access scheduling algorithm for data collection in WSNs. In this algorithm, we aim at minimizing the data collection latency by using both NOCs and POCs and increasing the network lifetime by balancing the routing tree. In following sections, we first describe the system model and assumptions, and then we show that using a combination of NOCs and POCs works better than each one separately. Finally, by considering this motivation and regarding the load balancing in tree construction, we describe the proposed algorithm.

3. System model description and assumption

In our model, wireless sensor network consists of a number of sensor nodes randomly distributed in a two-dimensional area. We consider the network as a graph $G(VUv_s, E)$

where V is the set of all sensor nodes, v_s is the sink node and E is the set of all communication links between nodes. We assume that all nodes are similar and have equal transmission power. In this work, we regarded protocol interference model, in which the transmission range of each node is a circle with radius r . Under these circumstances, if nodes u and v are tuned to channels i and j , respectively;

$|i - j|$ is defined as the channel separation of u and v . In addition, $(u, v) \in E$ means that u and v are located in each other's transmission range. Interference between two nodes depends on their channel separation and their distance. Interference range r' refers to the maximum distance from the transmitter that the transmitted signal causes interference on an unintended receiver.

For data collection, a tree with sink root v_s is constructed on the network graph $G(VUv_s, E)$. In raw data collection, each intermediate node relays its sensed data as well as the raw data received from the lower levels of the tree to its parent. In aggregated data collection, each node has exactly one packet to transmit, because before transmission, the node aggregates its sensed data along with the data received from the lower levels of the tree in the form of a packet. We assume that each link needs exactly one time slot to transmit the data and the capacities of all the links are the same. Since each node is equipped with a single half-duplex transceiver, in each slot, an intermediate node is only able to transmit or receive data, but cannot both transmit and receive. If necessary, the radio must switch between different channels. Here, for channel switching, we use the receiver-based channel allocation strategy.

In this paper, the main problem is how to jointly find routes to the sink and allocate different frequency channels and time slots to the links, in order to avoid collisions at the aim of minimizing the data collection latency and maximizing the network lifetime.

As shown in Figure 1, the 802.11b/g standard defines 11 channels. Each channel has a bandwidth of 22 MHz and separation between adjacent channels is 5 MHz. Based on this standard, if the separation of two channels is greater than

or equal to five, they are NOCs (here, channels 1, 6 and 11); otherwise, they are POCs. When employing NOCs to two links that are in interference range of each other, links can simultaneously transmit and receive only if they use different NOCs. Furthermore, each pair of transmitter and receiver should be set on the same channel. For example, see Figure 2 where two orthogonal channels C1 = channel 1 and C6 = channel 6 are used.

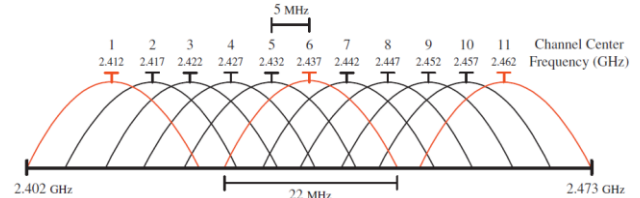


Figure 1. Frequency Channels in 802.11b/g Standard

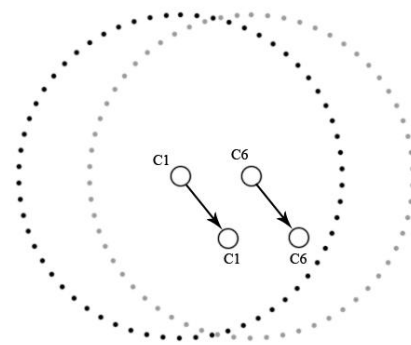


Figure 2. Parallel Transmissions using NOCs

In POCs, by increasing the channel separation, valid transmission and interference ranges are reduced, because only part of the transmitted signal is decodable in the receiver. If the separation of channels assigned to transmitter and receiver nodes is up to two, they can communicate with each other. For more separation, the power of the received signal is not acceptable, unless the distance between nodes be less than 0.5 Centimeters [26]. Also in parallel transmissions of two links, at least a channel separation of three is required. Valid transmission range is the maximum distance from the transmitter that the receiver is able to decode the transmitted signal correctly. The valid transmission range can be classified to three categories based on the channels separation [10]:

- Zero Transmission Distance (ZTD): is the valid transmission range for the channel separation of zero, i.e. when the transmitter and its related receiver use the same channel. ZTD depends on the network parameters. For example, the useful bandwidth of the received signal is 22MHz and ZTD is about 200m for minimum data rate in 802.11b/g standard.
- One Transmission Distance (OTD): is the valid transmission range for the channel separation of one, i.e. when the transmitter and its related receiver use adjacent channels. In this case, the useful bandwidth of the transmitted signal (B_1) is decreased in the receiver side in comparison to the zero channel separation (B_0). For example, the bandwidth is reduced from 22MHz to 17MHz in 802.11b/g. As a result, the received signal power (i.e. P_{R1}) is decreased proportional to the bandwidth reduction:

$$\frac{P_{R1}}{P_{R2}} = \frac{B_1}{B_0} \quad (1)$$

- Therefore, the SNR also decreases at the same rate. Moreover, in the spatial propagation model, received signal power is proportional to the inverse square of the distance from the transmitter. Therefore, in order to obtain the same SNR as the ZTD case, the distance between transmitter and receiver must be decreased to compensate the signal losses caused by channel separation as follows:

$$OTD = \sqrt{\frac{P_{R1}}{P_{R2}}} \times ZTD = \sqrt{\frac{B_1}{B_0}} \times ZTD \quad (2)$$

For example in 802.11b/g,

$$OTD = \sqrt{\frac{17}{22}} \times 200m = 175.8m.$$

- Two Transmission Distance (TTD): is the valid transmission range for the channel separation of two. Similar to the previous case, the useful bandwidth of the transmitted signal is decreased more in comparison to the one channel separation. For example, bandwidth is reduced from 22MHZ to 12MHZ in 802.11b/g. Similar to the previous case, we have:

$$TTD = \sqrt{\frac{B_2}{B_0}} \times ZTD \quad (3)$$

For example in 802.11b/g,

$$TTD = \sqrt{\frac{12}{22}} \times 200m = 147.7m.$$

As shown before, by increasing the channel separation between a pair of transmitter-receiver, the received signal level at the receiver is reduced. Thus, the valid interference range also is decreased proportional to the valid transmission range reduction. In this paper, for simplicity, we consider interference range as q times of the transmission range, which q is typically between 1 and 2.5.

4. Balanced MC-DC Algorithm

In this paper, we introduce a new heuristic algorithm named Balanced MC-DC at the aim of minimizing the scheduling length and maximizing the lifetime. The proposed algorithm optimizes the construction of routing tree, allocation of both NOCs and POCs to the links, and scheduling of the links simultaneously. In addition, in tree construction, we specify the parent-child relationships so that the tree is balanced. Balanced MC-DC algorithm starts from data collection sink and operates in a top-down manner, level-by-level, and node-by-node at each level. The level of each node is the number of hops where that node is far from the sink. For each level, the nodes of the level are prioritized to determine the order of their visit. Then, for each node, we specify its parent, its channel and its related time slot. Afterward, we go to the higher level and the same process is repeated. At the last, in order to implement the scheduling, we reverse the order of

time slots allocated to the nodes because children need sooner time slots than their parents [8].

In the following, we present some definitions required for description of the Balanced MC-DC algorithm in details:

i. Number of Interfering nodes (NoI): For each node u in the network graph $G(VUv_s, E)$, NoI is the number of nodes that will be interfered by transmission from node u :

$$NoI_u = \|\{v | v \in V, d(u, v) \leq r'(u)\}\| \quad (4)$$

where $r'(u)$ is the interference range of node u .

ii. Feasible Parent Set (FP-Set): If s' is the set of nodes that have been visited by the algorithm so far and $level_u$ is the level of node u , for each node u with valid transmission range $r(u)$, the set of possible parents are defined as below:

$$FP-Set_u = \{v | v \in s', d(u, v) \leq r(u), level_v = level_u - 1\} \quad (5)$$

Except the sink node that does not have a parent and nodes at first level that the sink is their only possible parent, $FP-Set$ can be calculated level by level for all nodes in the network.

iii. Feasible Channel Set (FC-Set): If P_u is a possible parent of node u in the data collection tree, and C_{P_u} is its frequency channel, and C_u is the possible channel that node u can use to communicate with its parent P_u , the set of feasible channels of node u is the set of 2-tuples (C_{P_u}, C_u) that is defined as follows:

$$FC-Set_u = \{(C_{P_u}, C_u) | P_u \in FP-Set_u, |C_{P_u} - C_u| \leq 2\} \quad (6)$$

To calculate the number of 2-tuples for each possible parent P_u of node u , one of the three following conditions may occur in IEEE802.11b/g standard:

If the Euclidean distance between two nodes is less than the valid transmission range for a channel separation of two, the size of the possible channel set is five. In other words, if $d(u, P_u) \leq TTD(u)$, then $\|FC-Set_u\| = 5$. This is because the node can choose a channel with separation of 0, 1, or 2 from its parent's channel.

- In a similar manner, if $TTD < d(u, P_u) \leq OTD(u)$ then $\|FC-Set_u\| = 3$.
- If $OTD < d(u, P_u) \leq ZTD(u)$ then $\|FC-Set_u\| = 1$.

iv. Feasible Time Set (T-Set): If T_{P_u} is the time slot allocated to the parent P_u of node u and L is the last time slot allocated so far, the set of feasible time slots for each node u is defined as follows:

$$T-Set_u = \{T | T_{P_u} + 1 < T \leq L\} \quad (7)$$

To reduce the data collection latency as much as possible, we allocate minimum possible time slot number T from the set $T-Set_u$ to u such that transmission from u do not interfere with other nodes that have similar time slot. If such time slot

is not available in $T - Set_u$, a new time slot will be allocated to u and L will be increased.

v. Number of Rely Data packets (N_{RD}): In raw data collection method, each node u relays the data generated by itself and the data generated by its children to its parent. In a network graph $G(VUv_s, E)$, for each node u , the number of relay data packets (N_{RD}^u) is the number of data packets that will be transmitted by node u in each round.

vi. Remained Energy of Last Ancestor (RELA): Assume LA_u is the last ancestor of node u in the data collection tree that connects u to the sink indirectly. For LA_u define $N_{RD}^{LA_u}$, $E_I^{LA_u}$ and e as the number of relay data packets, the initial energy and the minimum consumed energy to transmit a data packet, respectively. $RELA_u$ is defined as the remaining energy of the last ancestor of node u as follows:

$$RELA_u = E_I^{LA_u} - N_{RD}^{LA_u} * e \quad (8)$$

In Balanced MC-DC, at each level, we prioritize nodes according to their features. In this way, we first consider the size of the $FC - Set$ as a metric to prioritize the nodes. Nodes with lower size of $FC - Set$ give higher priority, and therefore will be visited sooner than other nodes. By using this strategy, nodes with fewer choices gain opportunity to be configured first. If there is more than one node with the same size of $FC - Set$, the node with larger Euclidean distance from its parent gives higher priority. In Balanced MC-DC, simultaneous to the allocation of the priorities in each level, we assign the best parent, best channel and best time slot to each node by considering the load balancing in the construction of the tree.

In this approach, we allocate channel 6 (middle channel) to the sink. This maximizes the number of available channels (size of $FC - Set$) for the nodes located at level one. At level one, sink is the parent of all nodes; first, we select a node with higher priority and then we choose a feasible channel from $FC - Set$ with maximum separation from the sink. Since the sink is equipped with a single radio, there is no parallel transmission at level one. Therefore, we allocate different time slots to the nodes according to their priority. As a result, the last time slot is equal to the number of nodes at level one. The algorithm will then enter the next level.

At each level, first a node with highest priority is selected, we call it u . Now, for each feasible parent P_u and each 2-tuple (C_{P_u}, C_u) of $FC - Set_u$, we specify the minimum possible time slot number from $T - Set_u$. If simultaneous transmission on previous assigned time slots is not possible, the corresponding 2-tuple takes a new time slot and L increases one unit. Finally, we select the parent and the frequency channel that results minimum time slot number. This strategy reduces the data collection latency in Balanced MC-DC. In described process, Balanced MC-DC allocates higher priority to the POCs and the NOCs are used only when applying the POCs leads to the allocation of new time slot. This is because in NOCs, the receiver must be switched

to the transmitter channel in order to receive the transmitted data that causes the waste of energy and time.

In described process, if the same time slot is selected by more than one 2-tuples, the parent with the maximum $RELA$ will be selected. This strategy helps to have more balanced collection tree, at the aim of increasing network lifetime and reducing latency. If the value for more than one parent is the same, the parent that causes minimum NoI_u will be selected to reduce the effects of interference. In this case, node u chooses a channel with maximum separation from C_{P_u} . This process repeats level-by-level from top to down and node-by-node at each level. At final, data collection latency will be L time slots.

As mentioned before, the proposed algorithm allocates time slots to nodes in a top-down approach; therefore, a parent has a time slot less than its children's time slots (see Equation 7). In practice, children should have a time slot less than their parent's time slot; therefore, we must reverse the order of time slots. In this way, we use following equation for each node u :

$$T_u^{new} = L - T_u^{old} + 1 \quad (9)$$

So far, we obtained the scheduling for aggregated data collection because we assigned only one time slot to each node. In some critical applications, there is not any correlation between data sensed by different nodes. In these cases, each packet generated by the nodes must be delivered to the sink without aggregation with other data packets. To extend the proposed algorithm to be applicable in raw data collection, at final step, for each time slot T , the maximum number of relay data packets (N_{RD}^{Max}) for the nodes that are allocated to slot T is calculated. Then, we extend the time slot T to N_{RD}^{Max} time slots. This process will be repeated for all time slots.

Pseudo-code of the Balanced MC-DC algorithm is presented in Algorithm 1. The algorithm ends when all nodes in the network $G(VUv_s, E)$ are considered from level 0 to the network radius (MaxLevel). At that time, each node u would be aware of its parent (P_u), its channel (C_u) and its assigned time slots sets (T_u).

Algorithm 1- Balanced MC-DC Algorithm

```

1: % Main Algorithm
2:  $C_{v_s} = 6; T_{v_s} = 0; L = 0;$ 
3: for (k = 1 to Maxlevel)
4:   { N=(set of nodes in level k)
5:     if (k=1 then)
6:       { while (N!=0)
7:         { Choose node  $u \in N$  with max  $d(u, v_s)$ 
8:           Choose  $(6, i) \in FC - Set_u$  with max  $|6 - i|$ 
9:            $P_u = v_s; C_u = i; T_u = L + 1;$ 
10:           $L = L + 1;$ 
11:           $N = N - \{u\};$ 
12:        }
13:      }
14:    else
15:      { while(N!=0)

```

```

16:      { Choose node  $u \in N$  with  $\min \|FC - Set_u\|$ 
and then  $\max d(u, v_s)$ 
17:      for  $(\forall((C_p, i) \in POCs) \in FC - Set_u)$ 
18:      { Search the earliest available time slot  $T_{POC}$ 
19:      }
20:      If  $(T_{POC} \in T - Set_u)$  then
21:      { Choose  $(C_p, i)$  POCs with earliest time
 $T_{POC}$  and then  $\max RELA$  and then  $\min NOI$  and then
max separation
22:       $P_u = p; C_u = i; T_u = T_{poc};$ 
23:       $N = N - \{u\};$ 
24:      }
25:      else
26:      { for  $(\forall((C_p, i) \in NOCs) \in FC - Set_u)$ 
27:      { Search the earliest available time slot  $T_{OC}$ 
28:      }
29:      If  $(T_{OC} \in T - Set_u)$  then
30:      { Choose  $(C_p, i) \in NOCs$  with earliest time
 $T_{NOC}$  and then  $\min NOI$ 
30:       $P_u = p; C_u = i; T_u = T_{NOC};$ 
31:       $N = N - \{u\};$ 
32:      }
33:      else
34:      { Choose  $(C_p, i)$  POCs with earliest time
 $T_{POC}$  and then  $\max RELA$  and then  $\min NOI$  and then
max separation
35:       $P_u = p; C_u = i; T_u = L + 1;$ 
36:       $L = L + 1;$ 
37:       $N = N - \{u\};$ 
38:      }
39:      }
40:      }
41:      }
42:      }
43:      % Reverse the Scheduling
44:      for  $(\forall u \in V)$ 
45:      {  $T_u = L - T_u + 1;$ 
46:      }
47:      % Raw Data Collection
48:       $T = 1;$ 
49:      While ( T is less than or equal to L) do
50:      { M=(all nodes that assigned to time slot T)
51:      choose  $\max N_{RD}$  in M
52:      for  $(\forall u \in M)$ 
53:      {  $T_u = \{T + 1, \dots, T + N_{RD}^{Max}\}$ 
54:      }
55:       $L = L + N_{RD}^{Max};$ 
56:       $T = T + N_{RD}^{Max};$ 
57:      for (all other nodes V not visited so far)
58:      {  $T_v = T_v + N_{RD}^{Max};$ 
59:      }
60:      }

```

5. Performance Evaluation

In this section, we evaluate the performance of the Balanced MC-DC algorithm in terms of following metrics:

- **Latency:** that is defined as the number of time slots required to collect data from all sensor nodes by the sink in a round.
- **Standard deviation of the remaining energy of hot spot nodes:** In a WSN, nodes located at first level are called hot spots, because they usually consume energy more than other nodes. If E_{R_i} is the remaining energy of i th hot spot (located at level 1), $\overline{E_R}$ is the average remaining energy of the hot spot nodes and n is the number of hot spots, then standard deviation of the remaining energy of hot spot nodes is defined as follows:

$$SD_{HSspots} = \sqrt{\sum_{i=1}^{i=n} (E_{R_i} - \overline{E_R})^2} \quad (10)$$

$SD_{HSspots}$ represents the balance in energy consumption. It is clear that the small value of this metric is desirable.

- **Network lifetime:** that is defined as the number of rounds until the first node dies.

We compare the results with MC-MLAS protocol [10] and the POC-based approach introduced in [8], which is known as one of the best data collection algorithms in multichannel sensor networks. For simulation, we used the Microsoft Visual Studio software environment and C#.Net programming language.

In each scenario, we randomly distribute a number of sensor nodes in a square area to generate a connected network topology. The network nodes are equipped with a single radio. The number of available frequency channels is 11 as shown in Figure 1. All sensor nodes have the same transmission ranges: ZTD=200 m, OTD=175.8 m, and TTD=174.7 m (Refer to Section III). The initial energy of each sensor node is considered 5000 units and the energy consumption per data packet for sensing, processing and transmission is assumed one unit. Each numerical result presented in this section is average of 10 different runs.

First, we consider a square area of 1000×1000 m². Here, the interference range of each node is assumed 1.5 times of the transmission range ($r' = 1.5r$).

Figure 3 shows the latency of data collection for three different approaches and for different number of nodes. From this figure, it is clear that the latency in POCs-based approach is more because Balanced MC-DC and MC-MLAS approaches use both NOCs and POCs and therefore they provide more potential parallel transmissions compared to POCs-based approach. According to Figure 3, data collection latency in Balanced MC-DC algorithm is much less than MC-MLAS approach. That is because the proposed algorithm considers the load balancing in constructing the data collection tree. This means that the proposed algorithm tries to balance the number of relayed data packets (N_{RD}) for nodes that are at the same level; therefore, the data collection latency is reduced considerably.

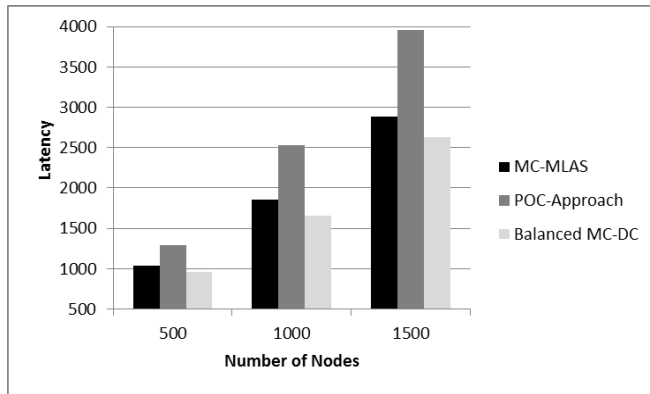


Figure 3. Comparison of the data collection latency in three different approaches

Another interesting point is that by increasing the number of nodes (network density), obviously the interference increases, so the data collection latency increases. In addition, in more dense networks, our algorithm in comparison to MC-MLAS decreases the latency more, because Balanced MC-DC algorithm provides load balancing and so performs better in dense networks. In other words, proposed algorithm is more scalable compared to MC-MLAS.

In addition, by considering load-balancing concept in Balanced MC-DC algorithm, nodes that are at the same level consume the same amount of energy. This means that data transmission from source nodes to sink will be distributed evenly among different branches of the data collection tree. Without load balancing, more data transmission on a branch leads to sooner energy loss at high-level nodes of that branch (nodes close to the sink) and this result less network lifetime. In Figure 4, standard deviation of the remaining energy of hot spot nodes ($SD_{HSspots}$) is depicted. As can be seen, in Balanced MC-DC algorithm, standard deviation of the remaining energy of hot spots, which represents the balance in energy consumption, is less than the other two algorithms. In summary, our proposed method provides less latency than the POCs-based and the MC-MLAS approaches (on average 31.15% and 6.6%, respectively). Also, the standard deviation of the remaining energy of hot spots in the Balanced MC-DC is less than MC-MLAS (on average 18.79%).

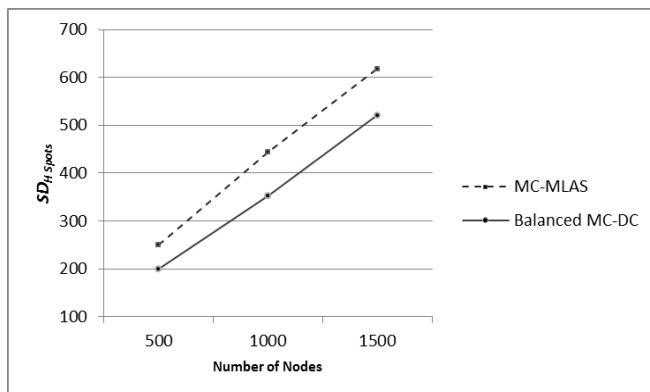
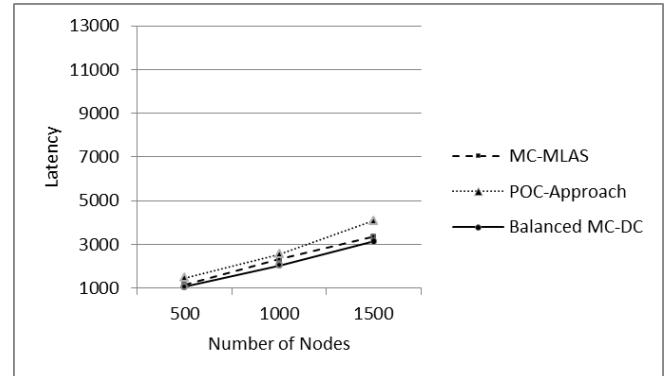


Figure 4. Comparing the standard deviation of remaining energy of Hot Spot nodes

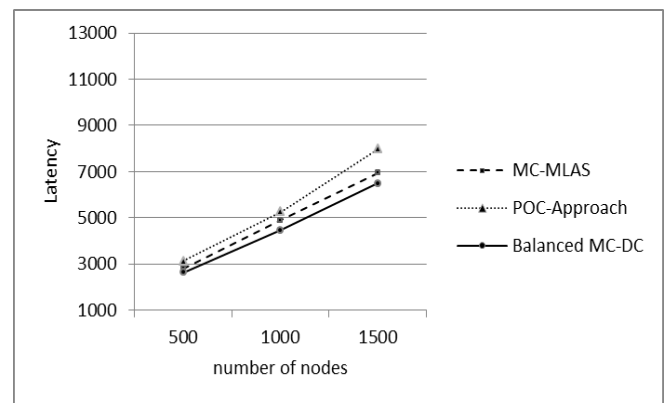
This increases the lifetime of the wireless sensor network. The reason for the superiority of Balanced MC-DC to MC-

MLAS is considering the load balancing in the construction of the data collection tree, as mention before.

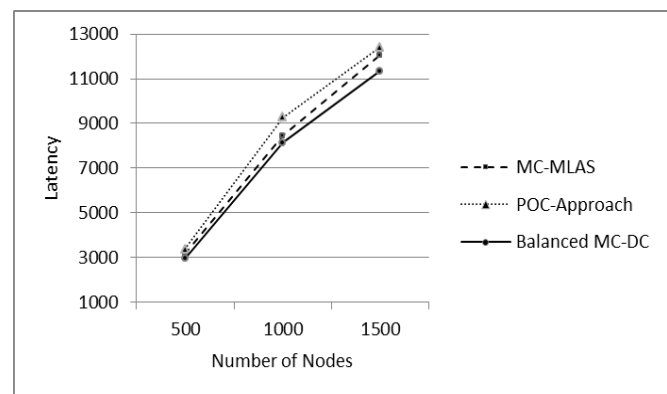
In the following, we examine the impact of the network size on the efficiency of the algorithms. In the rest, the interference range is considered twice the transmission range. In Figure 5, data collection latency in a square network of size 1000, 2000 and 3000 meters is presented.



(a) Network Size 1000 × 1000



(b) Network Size 2000 × 2000



(c) Network Size 3000 × 3000

Figure 5. Comparison of data collection latency for three approaches with different network sizes

Here, with increasing network size, the distance between nodes is increased. Increasing the distance between nodes leads to reducing the number of nodes in each level and, therefore, it will increase the number of levels (depth) of the tree. Thus, as can be seen, by increasing the network size, data collection latency in the three mentioned approaches increases. However, the latency of the proposed method is always less than the other two methods.

Our next assessment is studying the efficiency of Balanced MC-DC in a 31 × 31 grid topology. Here, the sink is located

on the right upper corner and we put sensor nodes on the cross points of the grid. So, overall 960 sensor nodes are available in the network. In Figure 6, comparison of the Balanced MC-DC and MC-MLAS approaches in terms of data collection latency is demonstrated. In each scenario, we considered three different sizes (25 m × 25 m, 50 m × 50 m and 75 m × 75 m).

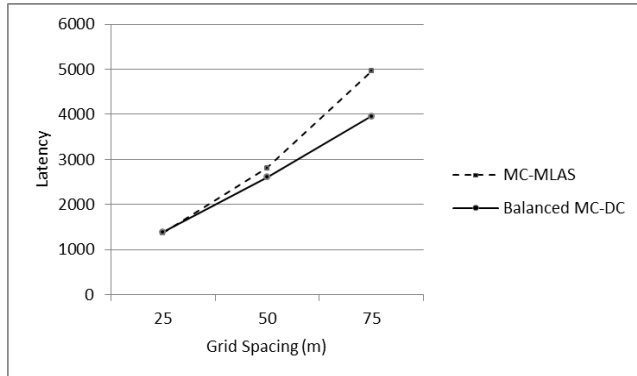


Figure 6. Comparison of data collection latency for different approaches in the grid topology

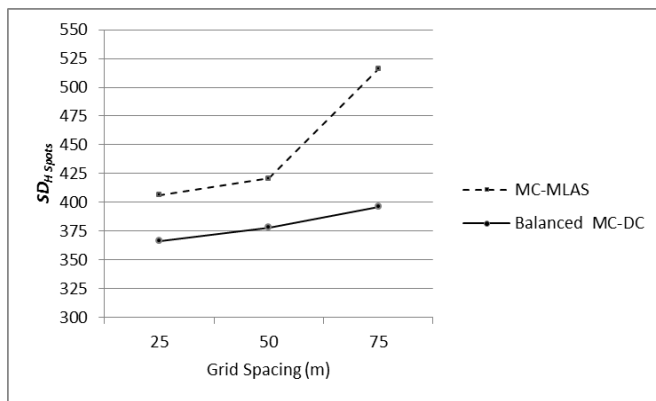


Figure 7. Comparison of the standard deviation of remaining energy of grid topology

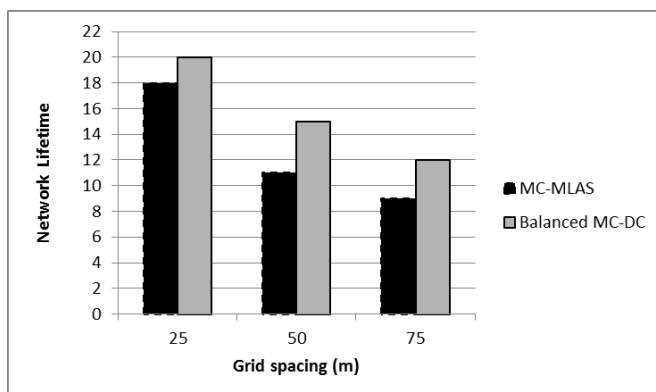


Figure 8. Comparison of Network Life Time based on FND criterion

Figure 6 shows that the proposed approach has lower latency compared to MC-MLAS approach (on average 13.20%). Moreover, as we can see in Figure 6, with the increase of the distance between nodes, the number of tree levels increases, which increases the data collection latency.

In Figure 7, we have shown that the standard deviation of the remained energy in Balanced MC-DC algorithm improved 14.36% compared to MC-MLAS algorithm. This increases the lifetime of the wireless sensor network significantly.

Figure 8 shows the network lifetime based on First Node Die (FND) criterion for Balanced MC-DC and MC-MLAS. These results indicate an increase in the lifetime of the network in the proposed algorithm.

6. Conclusions

Energy efficiency and data collection latency are two significant problems in WSNs. In this paper, we proposed a multi-channel data collection protocol based on TDMA called Balanced MC-DC, to reduce latency in raw data collection and to increase the lifetime of the network. It benefits from the advantages of both POCs and NOCs to maximize data collection rate in continuous monitoring applications. The key feature of the proposed algorithm is simultaneous running of collection tree construction, channel allocation and scheduling. In this algorithm, by considering the load balancing during the creation of data collection tree, we increase the network lifetime.

Simulation results show that the Balanced MC-DC significantly outperforms other well-known multi-channel data collection scheduling approaches, especially in dense WSNs. We compared Balanced MC-DC with MC-MLAS and POCs based approaches. The evaluation was done in networks with different conditions and different topologies in order to compare the data collection latency and network lifetime. In this way, we compared Balanced MC-DC with MC-MLAS in terms of standard deviation of the remained energy. As the simulation results show, our proposed approach had better performance in all cases.

For future works, the Balanced MC-DC algorithm can be extended based on the aspects that are not considered in this paper. For example, implementation of the distributed Balanced MC-DC, use of other interference models, use of multi-radio sensor nodes and implementation of algorithm on real test-beds can be as some of the major issues for future research.

References

- [1] F. Akyildiz, Su. Weilian, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.
- [2] S. C. Huang, P. J. Wan, C.T. Vu, Y. Li, and F. Yao, "Nearly constant approximation for data aggregation scheduling in wireless sensor networks," in *Proc. of IEEE INFOCOM Conf.*, pp. 366-372, 2007.
- [3] X. Xu, X.Y. Li, X. Mao, S. Tang, and S. Wang, "A delay-efficient algorithm for data aggregation in multihop wireless sensor network," *IEEE Trans. Parallel and Distributed Syst.*, vol. 22, no. 8, pp. 163-175, 2010.
- [4] O. D. Incel, and B. Krishnamachari, "Enhancing the data collection rate of tree-based aggregation in wireless sensor networks," in *Proc. of Sensor, Mesh and Ad Hoc Communications and Networks Conf.*, pp. 569 – 577, 2008.
- [5] Ghosh, O. D. Incel, V. A. Kumar, and B. Krishnamachari, "Multi-channel scheduling algorithms for fast aggregated convergecast in sensor networks," in *Proc. of Mobile Adhoc and Sensor Systems Conf.*, pp. 362-372, 2009.
- [6] A. Mishra, V. Shrivastava, S. Banerjee, and W. Arbaugh, "Partially overlapped channels not considered harmful," *SIGMETRICS Performance Evaluation Review*, vol. 34, no. 1, pp. 63–74, 2006.
- [7] A. Mishra, S. Banerjee, and W. Arbaugh, "Weighted Coloring Based Channel Assignment for WLANs," *ACM SIGMOBILE*

- Mobile Computing and Comm., vol. 9, no. 3, pp. 19-31, 2005.
- [8] B. Wang, and X. Jia, "Reducing data aggregation latency by using partially overlapped channels in sensor networks," in Proc. of IEEE CLOBECOM Conf., pp. 1-6, 2009.
- [9] Y. Ding, Y. Huang, G. Zeng, and L. Xiao, "Using Partially Overlapping Channels to Improve Throughput in Wireless Mesh Networks," IEEE Trans. Mobile Comp., vol. 11, no. 11, pp. 1720-1733, 2011.
- [10] F. Ghods, H. Yousefi, A. M. A. Hemmatyar, and A. Movaghar, "MC-MLAS: Multi-channel Minimum Latency Aggregation Scheduling in Wireless Sensor Networks," Computer Networks, vol. 57, no. 18, pp. 3812-3825, 2013.
- [11] C. Florens, M. Franceschetti, and R. McEliece, "Lower bounds on data collection time in sensory network," IEEE Journal on Selected Areas in Communications, vol. 22, no. 6, pp. 1110-1120, 2004.
- [12] C. Florens, and R. McEliece, "Scheduling algorithms for wireless ad-hoc sensor networks," in Proc. of IEEE GLOBECOM Conf., vol. 1, pp. 6-10, 2002.
- [13] C. Florens, and R. McEliece, "Packets distribution algorithms for sensor networks," in Proc. of IEEE INFOCOM Conf., vol. 2, pp. 1063-1072, 2003.
- [14] L. Guo, Y. Li, and Z. Cai, "Minimum-Latency Aggregation Scheduling in Wireless Sensor Network," J Comb Optim, vol.31, no. 1, pp. 279-310, 2016.
- [15] P. Neamatollahi, M. Naghibzadeh, and S. Abrishami, "Fuzzy-Based Clustering-Task Scheduling for Lifetime Enhancement in Wireless Sensor Networks," IEEE Sensors Journal, vol. 17, no. 20, pp. 6837 - 6844, 2017.
- [16] N. Lai, C. King, and C. Lin, "On maximizing the throughput of convergecast in wireless sensor networks," in Proc. of Int. Conf. on Advances in Grid and Pervasive Computing, pp. 396-408, 2008.
- [17] S. Ergen, and P. Varaiya, "TDMA scheduling algorithms for wireless sensor networks," Wireless Networks, vol. 16, no. 4, pp. 985-997, 2010.
- [18] V. Zibakalam, "A New TDMA Scheduling Algorithm for Data Collection over Tree-Based Routing in Wireless Sensor Networks," ISRN Sensor Networks, pp. 1-7, 2012.
- [19] S. Gandham, Y. Zhang, and Q. Huang, "Distributed minimal time convergecast scheduling in wireless sensor networks," in Proc. IEEE Int. Conf. on Distributed Computing Systems, pp. 50, 2006.
- [20] S. Gandham, Y. Zhang, & Q. Huang, "Distributed time-optimal scheduling for convergecast in wireless sensor networks," Computer Networks, vol. 52, no. 3, pp. 610-629, Feb. 2008.
- [21] D. O. Incel, A. Ghosh, and B. Krishnamachari, and K. Chintalapudi, "Fast data collection in treebased wireless sensor networks," IEEE Trans. Mobile Comp., vol. 11, no. 1, pp. 86-99, 2012.
- [22] A. Ghosh, O. D. Incel, V. A. Kumar, and B. Krishnamachari, "Multi-channel scheduling algorithms for fast aggregated convergecast in sensor networks," in Proc. of IEEE Int. Conf. on MASS, pp. 362-372, 2009.
- [23] V. Akila, T. Sheela, and G. Adiline Macruga, "Efficient packet scheduling technique for data merging in wireless sensor networks," China Communications, vol. 14, no. 4, pp. 35 - 46, 2017.
- [24] J. Ma, W. Lou, Y. Wu, X. Li, and G. Chen, "Energy Efficient TDMA Sleep Scheduling in Wireless Sensor Networks," in Proc. of IEEE INFOCOM Conf., pp. 630-638, 2009.
- [25] B. Zeng, Y. Dong, J. He, and D. Lu, "An energy-efficient TDMA scheduling for data collection in wireless sensor networks," in Proc. of Int. Conf. on Communications in China (ICCC), pp. 633-638, 2013.
- [26] A. Y. Barnawi, "Adaptive TDMA Slot Assignment Using Request Aggregation in Wireless Sensor Networks," Procedia Computer Science, vol. 10, pp. 78-85, 2012.
- [27] S. Saginbekov, and A. Jhumka, "Many-to-many data aggregation scheduling in wireless sensor networks with two sinks," Computer Networks, vol. 123, pp. 184-199, 2017.
- [28] Y. Cui, W. Li, and X. Cheng, "Partially overlapping channel assignment based on node orthogonality for 802.11 wireless networks," in Proc. of IEEE INFOCOM Conf., pp. 361-365, 2011.
- [29] X. Jin, H. Xu, C. Xia, J. Wang, and P. Zeng, "Convergecast scheduling and cost optimization for industrial wireless sensor networks with multiple radio interfaces," Wireless Networks, pp. 1-15, 2016.