Mobile sink based fault diagnosis scheme for wireless sensor networks

Article

Accepted Version

Chanak, P., Banerjee, I. and Sherratt, R. S. (2016) Mobile sink based fault diagnosis scheme for wireless sensor networks. Journal of Systems and Software, 119. pp. 45-57. ISSN 0164-1212 doi: https://doi.org/10.1016/j.jss.2016.05.041 Available at http://centaur.reading.ac.uk/65746/

Mobile Sink Based Fault Diagnosis Scheme for Wireless Sensor Networks

Prasenjit Chanak*[,a], Indrajit Banerjee[a], and R. Simon Sherratt[b]

[a] *Department of Information Technology, Indian Institute of Engineering Science and Technology, Shibpur, Howrah-711103, India*
[b] *Department of Biomedical Engineering, University of Reading, Reading, Berkshire, RG6 6AY, UK*

Abstract—Network diagnosis in Wireless Sensor Networks (WSNs) is a difficult task due to their improvisational nature, invisibility of internal running status, and particularly since the network structure can frequently change due to link failure. To solve this problem, we propose a Mobile Sink (MS) based distributed fault diagnosis algorithm for WSNs. An MS, or mobile fault detector is usually a mobile robot or vehicle equipped with a wireless transceiver that performs the task of a mobile base station while also diagnosing the hardware and software status of deployed network sensors. Our MS mobile fault detector moves through the network area polling each static sensor node to diagnose the hardware and software status of nearby sensor nodes using only single hop communication. Therefore, the fault detection accuracy and functionality of the network is significantly increased. In order to maintain an excellent Quality of Service (QoS), we employ an optimal fault diagnosis tour planning algorithm. In addition to saving energy and time, the tour planning algorithm excludes faulty sensor nodes from the next diagnosis tour. We demonstrate the effectiveness of the proposed algorithms through simulation and real life experimental results.

***Keywords* —Diagnostics, Infrastructure protection, Network monitoring, Wireless sensor networks.**

1. Introduction

Recently, Wireless Sensor Networks (WSNs) have risen as a practical solution for a variety of remote applications, such as battlefield surveillance, environmental monitoring, home security and automation, weather forecasting, medical and industrial monitoring, etc. (Banerjee et al., 2014; Wu et al., 2008; Chen et al. 2006; Lee and Choi, 2008 & Chanak et al. 2014). A WSN comprises of a set of smart sensing devices where each sensor node is equipped with limited memory, typically a low performance microcontroller, a power constrained transceiver and limited power availability.

*Corresponding author. Tel.:+91-9874040413, + 91 033 2668 2916 (fax), E-mail: prasenjit.chanak@gmail.com

In a monitoring field, sensor nodes are usually deployed without a preconfigured infrastructure. After deployment, sensor nodes form an ad-hoc network using a nearby node discovery process. In WSNs, the deployed sensor nodes are prone to various faults such as transceiver unit fault, sensor unit fault, processing unit or microcontroller unit fault and the power unit fault (Banerjee et al., 2014; Lee and Choi, 2008; Chen et al., 2012 ; Misra et al., 2014 & Bari et al., 2012). Faulty sensors produce erroneous data during the normal operation of the network which can reduce the number of available multihop paths in the network. Hence, faulty sensor nodes can potentially degrade the Quality of Service (QoS) of the WSN, since it is desirable to detect, locate, and ignore faulty sensor nodes during normal operation of the network.

Due to instability and uncertainty, fault diagnosis is very difficult in WSNs. However, network diagnosis is more crucial within a highly dynamic topology as the network structure of the WSN frequently changes due to environmental interference and uncertainty of the wireless medium. Therefore, the design of a Mobile Sink (MS) based distributed fault diagnosis scheme can effectively help network administrators monitor the network operation and maintain a wireless sensor network system.

In existing fault detection approaches, most of the energy available to a sensor node is consumed on two major tasks, *viz*. diagnosis status selection of deployed sensor nodes and localization of faulty nodes within the network. Fault diagnosis strategies depend on the network topology or the location of sensor nodes since fault diagnosis is an important factor that can directly impact the performance and lifetime of the network. In dynamic applications of WSNs, the network structure can frequently change due to rapid fault occurrences within the network. Existing static sink based fault diagnosis approaches lead to large numbers of messages sent over the network (both data and status) in order to adapt to the topological changes. Hence, the available energy of the sensor nodes in the network can be rapidly depleted. In addition, sensors close to the sink suffer from much more traffic being routed through them compared to sensors at the boundary of the network due to the need to route data and status packets from sensors that are far away from the sink (Lau et al., 2014 & Koushanfar et al., 2003). After these sensors fail, communication holes, or energy holes, are created near to the sink node and the network can then become unreliable or even disconnected. In some cases, nodes and link failures may potentially portion an entire network into several sub-networks, hence these sub-networks become disconnected from the rest. Then, a Base Station (BS), or network administrator can declare these sub-networks dead due to the lack of available health information and exclude these sub-network nodes from the main network, despite most of the sub-network nodes can still survive for a long period of time. Recently, interesting approaches have been used for MS data gathering that can collect data from deployed sensor nodes in an energy efficient manner (Zahhad et al., 2015 & Mi et al., 2015). These MS based data gathering strategies successfully collected data from different sub-networks in the network portion state. In these approaches, it was argued that MS based WSN management strategies are more effective in improving the performance of the network in dynamic environments. These works motivated us to propose an MS based fault diagnose strategy for WSNs.

To address the above problems, this paper proposes an MS based distributed fault diagnosis approach. In our approach, a mobile fault detector starts the fault diagnosis tour periodically from the BS, traverses the network, performs a diagnosis action on each static sensor node using single-hop communication and at the end of each fault diagnosis tour the mobile fault detector transports the entire network health information to the BS. During the fault detection tour, each deployed sensor node is directly diagnosed by the mobile fault detector and hence the current network structure is not affected by the fault diagnosis process. Moreover, if the fault detection tour is well planned, the mobile fault detector can accurately localize the abnormal nodes within in the network thus reducing the fault detection delay. This will give network administrators an up to date status of the network.

The major contributions of this paper can be summarized as follows:

1) We propose a Mobile Sink based fault detector to perform fault diagnosis in WSNs. It reduces the message overhead and is resilient to network topology changes during the fault diagnosis process.
2) We propose a hardware fault detection mechanism where each hardware component of the deployed sensor nodes is diagnosed by the mobile fault detector. Therefore, network administrators find the exact causes of any faults within the network. The proposed detector may also help to maintain the network.
3) We focus on the problem of minimizing the length of each fault diagnosis tour by excluding faulty sensor nodes from the WSN, improving QoS.
4) We carry out extensive simulations and real life experiments. The effectiveness of the proposed scheme is verified by comparing our method with other fault detection approaches in the literature. In addition, the real time applicability of the proposed scheme is confirmed by the real life experimental results.

The rest of the paper is organized as follows. Section 2 introduces the related work. Section 3 describes preliminaries and architecture of the proposed algorithm. The proposed fault diagnosis mechanism is presented in

Section 4. In Section 5, we discuss the shortest fault diagnosis tour planning mechanism. Section 6 gives the simulation results and discussion. Finally, Section 7 concludes this paper.

2. Related works

Existing fault detection approaches can be classified into two groups: a) centralized fault detection approaches (Lau et al., 2014), and b) distributed fault detection approaches (Lee and Choi, 2008; Chen et al., 2012 & Bari et al., 2012).

2.1 Centralized Fault Detection Approaches

In the existing centralized fault detection approaches, a centralized static sink makes diagnostic decisions by periodically injecting health requests or query messages to other nodes. Therefore, a large number of message exchanges are needed over the network for data and status exchange. This process also puts a significant level of traffic onto the network which itself depletes the energy of the deployed sensor nodes. Hence, the functionality of the network can be decreased due to the fault detection process itself.

Lau *et al*. (2014), proposed a centralized fault detection strategy for a WSN based on the Naïve Bayes framework. This approach explored end-to-end packet transmission delay to analyze the network status. The disadvantage of this approach was that it did not work in a dynamic environment where network topology frequently changes due to faulty nodes. It required a large time frame to diagnose the fault condition of the deployed sensor nodes in large scale WSNs and it also created a high volume of traffic through the central fault diagnosis node. Therefore, this approach is not suitable for large scale WSNs.

A management architecture based fault detection scheme for WSNs was proposed by Ruiz *et al*. (2014) where faulty sensor nodes were detected by the central manager. In this approach, a central manager with a global vision of the network was primarily responsible for diagnosis of sensor node failures within the network. This approach led to a large number of message exchanges over the network for status and data exchange which potentially reduces the lifetime of the network. This approach also puts a significant traffic load on the central node for large scale WSNs.

The taxonomy for classification of faults in the WSN and an on-line model based fault testing mechanism was introduced by Koushanfar *et al*. (2003). This fault detection system worked for a heterogeneous sensor network with an arbitrary type of fault model. In this centralized fault detection approach, the BS gathered all the sensor node information and conducted an on-line fault diagnosis process. This centralized fault detection approach suffered from large message overhead which potentially decreases the functionality of the network. Furthermore, this approach is not an efficient fault detection algorithm in terms of detection accuracy since it did not consider the dynamic changes of the network during the fault detection process.

A centralized fault detection approach “Sympathy” was studied by Ramanathan *et al.* (2005) and in this approach, a centralized sink node gathered data from the deployed sensor nodes and analyzed the gathered data through their “Sympathy” fault detection tool. Using this approach the fault diagnosis time and message overhead was very high because of the time taken to make a decision for each sensor node to send their data to the central sink node. This approach suffered from poor network lifetime and detection accuracy.

2.2 Distributed Fault Detection Approaches

In the distributed approaches in the literature, sensor nodes themselves make node failure decisions on the basis of results from their neighbor nodes and updates individual node status information to the sink or BS. Hence, distributed fault detection approaches can handle fault detection delay and traffic load problems that have been created during the centralized fault diagnosis process.

Lee and Choi (2008) proposed a distributed algorithm, termed FDWSN, for detecting and isolating faulty sensor nodes from a WSN. Faulty sensor nodes in FDWSN were detected based on local comparisons between the neighbor nodes. Each individual sensor made its own decisions based on the local comparison results. This approach reused faulty sensor nodes as communication nodes for data routing, but they are logically isolated from the network. This approach tolerated transient faults through time redundancy during the data exchange process. The main drawback of this distributed approach was that each sensor node collected data from their neighbor nodes multiple times. As a result, this approach consumed more energy compared to other distributed fault detection approaches. Furthermore, this approach did not consider transmission faults that occur during the fault diagnosis process.

A fault tolerant mechanism using out-of-band monitoring was introduced for WSNs by Chen *et al.* (2012) where separate nodes were placed within the network for monitoring other deployed nodes. Furthermore, an Integer Linear Programming (ILP) problem was formulated for small size networks and heuristic algorithms were used to place the monitoring nodes within the network. This approach added a message overhead on the network during the separate nodes placement and also suffered from high transmission and computation cost.

A two-tier architecture based fault detection scheme was studied by Bari *et al.* (2012), being similar to the proposed method by Chen et al. (2012), but factors concerned with load balancing were also considered with the fault detection problem. Interestingly, this approach was a distributed fault detection mechanism where relay nodes with high energy were used as Cluster Heads (CHs). The ILP problem based relay node placement and cluster formation strategy were analyzed and this approach also proposed a load balancing routing strategy for reducing the load of the deployed sensor nodes and the relay nodes.

Ding *et al*. (2005) proposed a localized fault detection approach for WSNs where each sensor node compared its own sensed data with the median of its neighbor nodes data in order to diagnose its own health status. The disadvantage of this approach is that, if all neighbors of the diagnostic sensor nodes are faulty, then a functioning diagnostic sensor node can detect itself as having a fault when a fault may not be present. Therefore, the fault detection performance of this approach is very poor.

A probabilistic fault diagnosis mechanism was introduced by You *et al*. (2011) which proposed a probabilistic fault diagnosis model for local and global performance analysis of a network. In their approach, the fault diagnosis mechanism was divided into four sessions: (a) a diagnosis session, (b) a testing session, (c) a comparison session, and (d) a dissemination session. This approach led to a large number of message exchanges over the network for multiple copies of data and status exchange which potentially decreases the lifetime of the network.

A majority voting based distributed fault diagnosis approach was studied by Jiang (2009) were each sensor node compared its own sensed data with its neighbors' sensed data. The data was either marked as 'likely fault free' or 'likely faulty' based on the neighbor nodes voting majority. On the basis of rigid criteria, then the likely fault free nodes were finally selected as fault free nodes. The fault free nodes were then used to identify other functioning and faulty sensor nodes from the remaining likely fault free or likely faulty nodes. In this approach, communication overhead was very high due to multiple message exchange between the neighbor nodes.

A three-sigma edit test based Distributed Soft Fault Detection (DSFD) approach was introduced by Panda and Khilar (2012). In DSFD, each sensor node shared their own sensed data with neighbor nodes in order to identify probable faults of its own and neighbor nodes using the three-sigma edit test. Then, probable fault status was shared with the neighbor nodes. For fault diagnosis, each sensor node compared its own sensed data with its neighbor node sensed data and fault decisions were made based on a threshold value. This approach detected faulty nodes within the network, but it did not identify the detailed hardware and software condition of the deployed sensor nodes. Therefore, this approach detected many non-faulty nodes as faulty nodes during the faulty diagnosis phase and decreased the performance of the network. In addition, it was unable to tolerate any communication link failure problems during the data exchange process.

It was seen from the literature that centralized fault detection approaches diagnose faulty sensor nodes more accurately compared to self-fault diagnosis processes because a central fault detector has a global vision of the network that can help compare sensed data more accurately with the other deployed nodes. However, in the central fault detection approaches, the performance of the network is very poor compared to the distributed fault detection approaches in terms of energy consumption, detection delay, network lifetime, etc. On the other hand, the centralized fault detection approaches are unable to diagnose node status when an entire network is portioned into several sub-networks. Existing static sink based network diagnosis approaches suffered from different problems such as poor detection accuracy, huge detection message overhead, etc.

In this paper, we propose a mobile sink based distributed fault diagnosis scheme for WSNs that can detect faulty sensor nodes in a distributed manner which also provides a centralized fault detection approach with improved accuracy. In addition, it also detects network status when a network is divided into several sub-networks, i.e. the mobile fault detector can tolerate topological changes during the network diagnosis process. The difference between existing approaches and our work is that our objective is to minimize the dynamic environment effect on the fault detection process. In addition we focus on energy consumption during the fault diagnosis process where mobile sink base fault diagnosis strategy minimizes energy consumption of the deployed network.

3. Preliminaries and Architecture

Let $S = \{s_1, s_2, s_3, ..., s_n\}$ be the set of static nodes that constitutes the considered WSN infrastructure considered in this paper, that are randomly deployed in a monitoring area. The network structure of WSNs can frequently change due to the presence of error-prone or faulty nodes; therefore prior information about the topological connection is hard to obtain for the packet transmission from a sensor node to sink node, particularly for the case of a MS. To solve this problem we propose a MS based fault detection mechanism where a mobile fault detector can visit the transmission range of every static sensor node, such that the hardware and software status can be diagnosed by single hop communication, i.e. without any relay. In our method, topological connections do not affect the diagnosis process. Before describing our MS based fault detection system, we will first define some terms that will be used throughout the paper.

While a mobile fault detector is moving, it can poll nearby sensor nodes one by one in order to diagnose their fault status. When a static sensor node receives diagnostic messages from the nearby mobile fault detector, the sensor node transmits its own health information to the mobile detector directly without any relay. We define the position where a mobile fault detector polls static sensor nodes for fault detection as diagnosis hub points. Diagnosis hub points are calculated depending on the deployed sensor node positions and when the mobile fault detector moves to a diagnosis hub point, it obtains the nearby sensor node status with the same transmission power, such that nodes that receive the diagnostic messages can directly upload health information status to the mobile fault detector using single-hop communication. After checking the health status of all the sensor nodes around the diagnosis hub point, the fault detector moves directly to the next diagnosis hub point in a fault diagnosis tour. Each fault diagnosis tour consists of a number of diagnostic hub points and the straight line segment connecting them.

For example, let denote a set of diagnostic hub points and *BS* be $D = \{h_1, h_2, h_3, ..., h_n\}$ the starting and ending point of each fault diagnosis tour at the BS. Then the fault diagnosis tour can be represented by the $BS \rightarrow h_1 \rightarrow h_2 \rightarrow .... \rightarrow h_n \rightarrow BS$ path. The problem of finding faulty sensor nodes under the each diagnosis hub point and an optimal fault diagnosis tour can be considered as the problem of determining the locations of diagnosis hub points and the order of the visit. Before a mobile fault detector starts a fault detection tour, it needs to identify the positions of all diagnosis hub points and which static sensor nodes can be diagnosed at each diagnosis hub point. We define the neighbor set of a diagnosis hub point in the plane as the set of sensor nodes, from this diagnosis hub point where a mobile fault detector can diagnose all static sensor nodes health status directly. Since the mobile detector can only diagnose at a diagnosis hub point, each sensor node must be in the neighbor set of at least one diagnosis hub point for verifying its health status, i.e. the union of neighbor sets of all diagnosis hub points must cover all static sensor nodes.

If the initial network connection pattern of the deployed WSN can be determined, or if the one hop neighbor set of each node is known, the neighbor node set of each diagnosis hub point can be found. To identify the connection pattern within the nodes, a triangle formation process has been conducted between the deployed nodes. After the sensor nodes are deployed, every sensor node broadcasts "*ADVERTISE*" messages with the same transmission power as the data transmissions. Fig. 1 details the format of the advertisement packet (*ADVERTISE)* and Table 1 explains the symbols in Fig.1.

AFI	SID	DID	XPS	YPD

AFI: Advertisement Frame Identification
SID: Source ID
DID: Destination ID

XPS: X Position of Source
YPD: Y Position of Source

Fig. 1. The advertisement packet (*ADVERTISE*) format.

Table 1
Explanation of symbols in Fig. 1

AFI	Identification field for the advertisement frame
SID	The MAC (Media Access Control) ID of source node
DID	The MAC (Media Access Control) ID of destination node
XPS	X determines geographic position of the source node
YPD	Y determines geographic position of the source node

Each sensor can decode the "*ADVERTISE*" messages and reply with an "*ACK*" message to acknowledge the neighbor nodes. In the acknowledgement message, the first field contains the identification for the acknowledgement frame. The second field is designated as the MAC address of the previous node that forwarded the advertisement frame. The third field refers to the MAC address of the original source node and last two fields of this frame contain geographic position of the forwarding node. Each node identifies all its single hop neighbor node set and then neighbor information is uploaded to the mobile fault detector when it polls for neighbor information. Initially the mobile fault detector visits all sensors and collects neighbor node information from the deployed sensor nodes. Then, the *BS* identifies the optimal number of diagnostic hub points and their location within the network according to the present network connection pattern by the triangle based hub point identification method. In addition, the *BS* computes the optimal fault diagnosis path between the optimal fault diagnosis hub points. A detailed description on the optimal number of diagnosis hub point selection and the tour planning algorithm is summarized in Section 5.

4. Fault Diagnosis Mechanism

This section considers the sink mobility based fault detection problem during the fault diagnosis tour.

Sensor nodes are primarily constructed of four major hardware components: a) sensor unit, b) transceiver unit, c) power unit, and d) processing or microcontroller unit. The mobile fault detector separately identifies each hardware unit status (hard fault) and software status (soft fault) of the deployed static sensor nodes. Detailed hardware and software health information of the deployed sensor nodes can help network administrators to recover or reuse the faulty sensor node and maintain the sensor network system effectively. Algorithm 1 depicts the pseudo code of the proposed fault detection algorithm and Table 2 describes the symbols in Algorithm 1.

Algorithm 1: Fault diagnosis

```
Input: $S=\{s_1,s_2,.....,s_n\}$ . $H=\{h_1,h_2,h_3,....h_m\}$
Output: Fault status of the sensor nodes.
/* Fault diagnosis process */
When mobile fault detection reaches to a diagnosis hub point.
for $i=1,.....,|m|$ do
    Mobile fault detector broadcasts DIAGNOSIS_REQ message within it single hop neighbor set.
    Set $\tau_{timer}$ timer.
    If static sensor node receives diagnosis message from the nearest fault detector, it prepares a HEAL_INFO
    massage including its current power condition, present sensor reading and then sends to the nearest detector.
/*Hardware fault diagnosis*/
    if $\tau_{timer} \le T_x^j$ then
        Receive HEAL_INFO message for $s_i$ sensor node.
        Transceiver unit or processing unit of node $s_i$ is good.
        if $P_i \le B_{wx}$ then
            Power unit fault has occurred in node $s_i$.
            Node $s_i$ is inserted into faulty node set ($S_F$).
        end if
    else
        Transceiver unit or processing unit of node $s_i$ is faulty.
        Node $s_i$ is inserted into faulty node set ($S_F$).
    end if
    if HEAL_INFO content SD then
        Sensor unit of node $s_i$ is likely non faulty.
    else
        Sensor unit of node $s_i$ is likely faulty.
    end if
end for
/* Software fault diagnosis*/
for $i=1,.....,|m|$ do
    if $\tau_{timer} \le T_x^j$ then
        Receive HEAL_INFO message for node $s_i$.
        Identify x's k nearest neighbors in the training data set.
            $D_x^2 = {}^{1}\!/_{k}\sum_{j=1}^{k} d_{xj}^2$
        if $D_x^2 \ge D_\beta^2$ then
            Software or sensor circuit of node $s_i$ is non faulty.
        else
            It is detected as a faulty.
            Faulty node is inserted into faulty node set ($S_F$).
        end if
    end if
    Current acting node set $A_{curr} = S \setminus S_F$ .
    Update current acting node set $A_{curr}$ to the BS.
end for
return
```

Table 2
Explanation of symbols in Algorithm 1.

DIAGNOSIS_REQ	Diagnostic message broadcasts by the mobile fault detector within the single hop neighbor node set
τ_{timer}	Timer is started by the mobile fault detection
H	Diagnostic hub points
T_x^j	Threshold value for mobile fault detector decision making.
S	Static node set
S_F	Faulty node set
s_i	Sensor node
HEAL_INFO	Health information message
SD	Current data reading
P_i	Remaining battery power of sensor node, s_i,

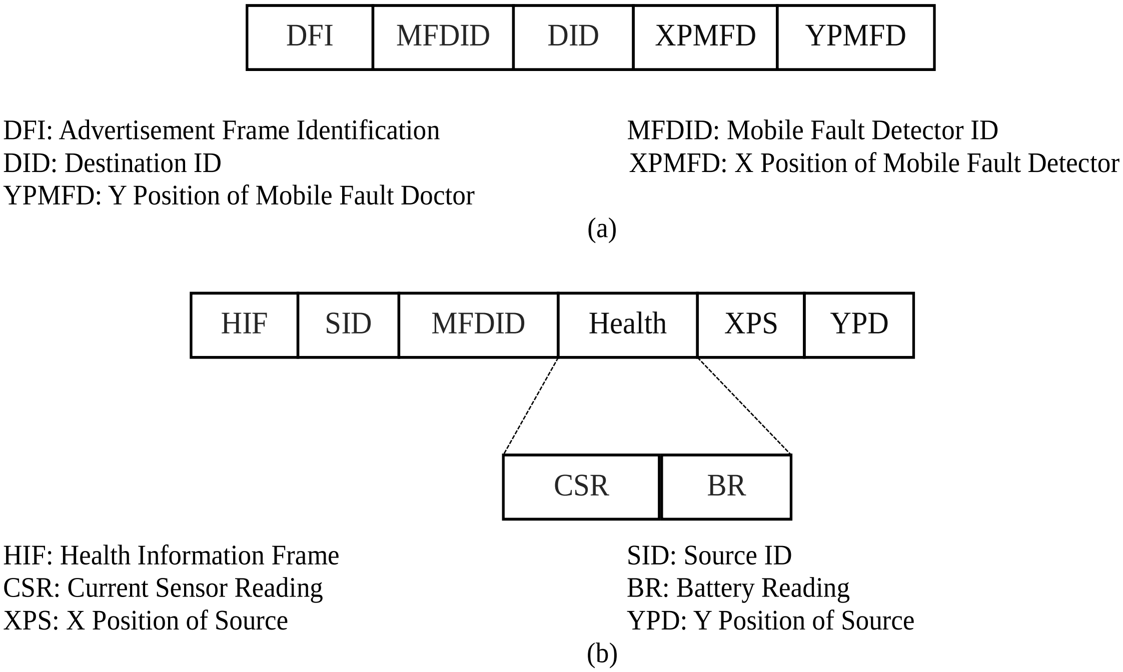
In the fault diagnosis phase, when a mobile fault detector reaches a diagnosis hub point, it broadcasts a diagnostic message “*DIAGNOSIS_REQ*” within the single hop neighbor node set and starts timer, τ_{timer} . Fig. 2.a is the format of a diagnostic message that is sent by the mobile fault detector and Table 3 explains the symbols Fig. 2. In the diagnostic message, the first field contains the identification for the diagnostic frame. The second field of this frame is designated to the MAC address of the mobile sink that is forwarding the diagnostic message. The last two fields of this frame contains the pair of coordinates x and y that determine the geographic position of the mobile fault detector. If a single hop neighbor node set receives the *DIAGNOSIS_REQ* message from the nearest mobile fault detector, each static node prepares a health information message “*HEAL_INFO*” and then sends the message to the nearest mobile fault detector. Fig 2.b is the format of a health information message that sent by the sensor node. In health information message, the first field of this frame contains identification for the health information frame. The second field refers to the MAC address of the actual node that is forwarding the health information message. The third field is designated to the MAC address of the mobile fault detector. The fourth field contains current sensor reading. The fifth field is designated as the battery reading and last two fields containing the pair of coordinates x and y that determines the geographic position of the node. If the mobile fault detector receives the *HEAL_INFO* message from a static sensor node (s_i) within time, τ_{timer} , the mobile fault detector makes a decision that the *transceiver unit* and the *processing unit* of the node s_i are functioning properly. The upper bound of τ_{timer} is dependent on the Round Trip Delay (*RTD*). It can be calculated as:

$$\tau_{timer} = \left\{\max\left\{RTD_{j,i}, \forall s_i \in Neg\ (d_j)\right\}\right\} \qquad (1)$$

where $RTD_{j,i}$ is the estimated transmission time delay between the diagnosis hub point j and its sounding neighbor node set $Neg\ (d_j)$. *RTD* can be calculated as:

$$RTD_{j,i} = 2 \times \left(\frac{dis\left(d_j, s_i\right)}{c} \right) + P_{ij} \qquad (2)$$

where c is the speed of light and P_i is the processing delay.



DFI: Advertisement Frame Identification
MFDID: Mobile Fault Detector ID
DID: Destination ID
XPMFD: X Position of Mobile Fault Detector
YPMFD: Y Position of Mobile Fault Doctor

(a)

HIF: Health Information Frame
SID: Source ID
CSR: Current Sensor Reading
BR: Battery Reading
XPS: X Position of Source
YPD: Y Position of Source

(b)

Fig. 2. (a) The diagnostic packet (DIAGNOSIS_REQ) format. (b) The format of health packet (HEAL_INFO) sent by sensor node.

Table 3
Explanation of symbols in Fig. 2

DFI	Identification for the diagnostic frame
HIF	Identification for the health information frame
SID	The ID of source node
MFDID	The ID of the mobile fault detector
DID	The ID of destination node
CSR	Current sensor reading
BR	Battery reading
XPMFD	x position of mobile fault detector
XPMFD	y position of the mobile fault detector
XPS	x position of source node

Initially, a *sensor unit* fault of a sensor node is likely to be detected by the mobile detector on the basis of the current data reading *(SD)* of the *sensor unit*. If a *HEAL_INFO* message contents the current *SD* value of node s_i, the mobile detector decides that the *sensor unit* of node s_i likely has a *non-fault* otherwise the *sensor unit* of the node is considered *likely faulty*. The *SD* value of deployed sensor nodes may vary depending on the applications, and in our experiments, we demonstrate our method using a hardware based temperature sensor. Therefore, during the experiment, *SD* was measured as 39^{o}C ambient by the deployed sensor nodes. However, a *sensor unit* fault is finally checked by the mobile fault detector during the software fault diagnosis process. The *power unit* fault of a sensor node is diagnosed by the mobile detector on the basis of the current battery reading of the sensor node. If the remaining battery power, P_i, of sensor node, s_i, is less than a threshold value B_{wx}, then the mobile fault detector makes a decision that the *power unit* of s_i is faulty.

In a WSN, if any software or sensor failure occurs within the deployed sensor node, the node can generate an incorrect data sample and respond to the mobile fault detector with this incorrect data sample. Therefore, this work uses the *k*-nearest neighbor (*kNN*) rule to diagnose a software fault in the deployed sensor nodes. The main concept of the *kNN* based fault detection system is that a normal data sample trajectory is similar to the trajectories of normal training data, therefore some deviation exists between the trajectory of a faulty sample and a normal one in the training data. A fault sample deviation or distance to the nearest neighboring training data samples must be greater than that of a normal data sample. This detection strategy is implemented by the average squared distance between a data sample to it *k* nearest neighbors in the normal training data samples. A threshold is determined based on the distribution of the normal training samples' distances to their *k* nearest neighboring training data samples. This threshold value is utilized to detect the unclassified sample or faulty nodes in a diagnosis hub point. If a sample distance to its nearest neighboring training samples is below the threshold, the sample is considered as normal, otherwise it is detected as faulty. In this MS based fault detection scheme, a threshold is determined based on the distribution of the collected data on a diagnosis hub point. This threshold can be utilized to detect the unclassified sensed data sample of the next nearest diagnosis hub point. If the sensed data sample distance to its nearest neighboring training data samples is below the threshold, the sample sensor node software is considered as normal. Otherwise, it is a faulty. For example, a training data set $X\left\{\infty x_1, x_2, \ldots\ldots., x_m\right\} \in R^{m\times d}, \left(i, j = 1,2,\ldots., m\right)$ where m is the number of sensor nodes covered by a single diagnostic hub point and d as the original dimension is covered by m sensors. The mobile fault detector discovers k number of nearest neighbor nodes for each data sample in the training data set. The squared distance of each sensed data sample x_i is defined by:

$$D_i^2 = {}^{1}\!/_{k} \sum_{j=1}^{k} d_{ij}^2 \tag{3}$$

where D_i^2 is the sum of the squared distance of the sample sensing data x_i to its k- nearest neighbors and d_{ij} denotes squared Euclidean distance from the sample sensing data x_i to its j^{th} nearest neighbor. The mobile fault detector determines the corresponding control limit of D_i^2 for fault detection of the sensor node in each diagnosis hub point. The threshold D_β^2 with a significant level β can be determined for a single diagnostic hub point and assumption that D_i^2 follows a non-central x^2 distribution.

5. Shortest Fault Diagnosis Tour Planning Scheme

This section considers the problem of finding the shortest fault diagnosis tour for the mobile fault detector.

In this proposed fault diagnosis strategy, each static sensor node belongs to a single hop diagnosis hub point. For simplicity, consider that the mobile fault detector moves at a fixed speed throughout the fault diagnosis tour and if a mobile fault detector moves through the shortest tour then the mobile fault detector can diagnose the maximum number of nodes in the shortest time since the BS or network administrator will have the most up-to-date network status.

After all the diagnosis hub points are obtained within the network, the shortest fault detection tour problem can be formalized as follows. Let $S = \left\{s_1, s_2, s_3, \ldots\ldots., s_n\right\}$ be the deployed node set and a set of diagnostic hub point as $H = \left\{h_1, h_2, h_3, \ldots\ldots., h_n\right\}$ and the neighbor set *nb(h$_i$)* of each candidate hub point $h_i\,(i = 1,\ 2, \ldots, n)$. Then the main objective is to find a set of diagnosis hub points and determine the sequence to visit them, such that every sensor in S

belongs to the neighbor set of at least one diagnostic hub point, with the total length of the line segments connecting all diagnosis hub points is minimized. Define a complete directed graph *G=(V, A)* and a non-negative cost c_{ij} with each arc $a_{ij} \in A$, where c_{ij} is equal to the cost of the distance between the diagnosis hub points h_i and h_j. The shortest detection tour problem can be formulated as:

$$\text{Minimize} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=1}^{n} c_{ij} x_{ijt} \tag{4}$$

subject to constraints:

$$\sum_{i} \sum_{j} x_{ijt} = 1 \text{ for all } t \tag{5}$$

$$\sum_{j} \sum_{t} x_{ijt} = 1 \text{ for all } i \tag{6}$$

$$\sum_{i} \sum_{t} x_{ijt} = 1 \text{ for all } j \tag{7}$$

$$\sum_{i} x_{ijt} = \sum_{k} x_{jkt+1} \text{ for all } j \text{ and } t \tag{8}$$

$$0 \le x_{ijt} \le 1$$

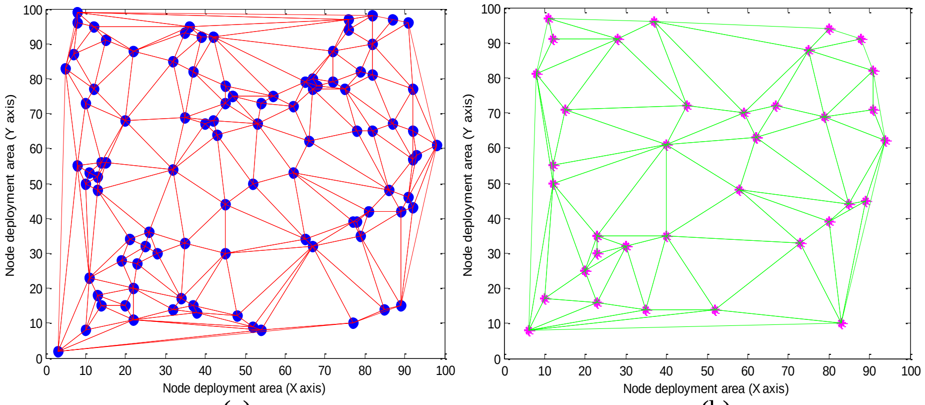
where

$$x_{ij} = \begin{cases} 1 & \text{if tour contains arc } a_{ij} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

In this formulation, the objective function (4) minimizes the total cost of the fault detection tour. Variable x_{ij} denotes whether arc a_{ij} of the diagnosis hub points h_i to h_j belongs to the optimal tour. Constraints (5) and (6) ensure all values of *t*, exactly one arc must be traversed by the mobile fault detector and there is just one other diagnosis hub which is being reached from it, at some time. Constraints (7) and (8) exclude other diagnosis hubs from being reached, at a time when a diagnosis hub is reached at time *t*, it must be left at time *t+1*, in order to exclude disconnected sub-tours that would otherwise meet all of the above constraints.

The proposed greedy algorithm selects an optimal subset of diagnosis hub points from the candidate diagnosis hub point set, each of which corresponds to a neighbor set of sensors. At each stage of the algorithm, a neighbor set of sensor nodes can be selected when its corresponding candidate diagnosis hub point is selected in the fault diagnosis tour. The algorithm will terminate after all sensor nodes are diagnosed by the mobile detector. Initially all the sensor nodes form a triangle with its single hop neighbor nodes and selects candidate diagnostic hub points at the mid-point of each triangle (Fig. 3.a). If the distance between the neighbor candidates' diagnostic hub points is less than the transmission distance, *r*, of a sensor node, neighbor candidate diagnostic hub points form another triangle between them and choose an optimal diagnostic hub point set (Fig. 3.b). Hence, an optimal path is computed between the optimal diagnosis points set (Fig. 4). A detailed description of the proposed greedy algorithm is described in algorithm 2.

Let H_{curr} contain all the diagnosis hub points, $H_{optimal}$ be the set of optimal candidate hub points, and T_{curr} contain the set of remaining uncovered non faulty sensor nodes at each stage of the algorithm. For each non faulty sensor node *t* in T_{curr}, *nb(t)* denote the neighbor set of *t*. Let cost, *nb(t)*, be the total communication cost of a neighbor node set *nb(t)*. Let $h_{cost} = cost\left(nb\left(t\right)\right)/k$, which denotes the average cost to cover all node in *nb(t)*. Choose a diagnostic hub point *h* with the h_{cost} value, add this diagnostic hub *h* of *nb(t)* into H_{curr}. Finally H_{curr} then contains all the diagnostic hub points. After obtaining all the diagnostic hub points, the proposed greedy algorithm chooses a subset of diagnostic hub points from the H_{curr} point set. For each candidate diagnosis hub point *h* in H_{curr}, let *nb(h)* denote the neighbor diagnosis hub point set of *h*. Let *cost {nb(h)}* be the cost of an uncovered neighbor hub point set *h*, which is equal to the shortest distance between *nb(h)* any covered neighbor diagnosis hub point set. Let *cost{r}* be the transmission distance of a sensor node. If $cost\left\{nb\left(h\right)\right\} \le cost\left\{r\right\}$, then choose an optimal hub point, insert the corresponding optimal hub point of *nb(h)* into $H_{optimal}$ and remove corresponding *nb(h)* from H_{curr}. The algorithm terminates when all H_{curr} are covered. Finally, $H_{optimal}$ contains diagnostic hub points in the fault diagnosis tour. After identification of the optimal diagnosis hub point set, the fault detection tour can be easily obtained by running any approximate algorithm for the traditional Travelling Salesman Problem (TSP).



(a) (b)

Fig. 3. Optimal number of diagnostic hub points. (a) Triangle formation between the deployed sensor nodes to find diagnostic hub points between them, (b) Triangle formation between the diagnostic hub points to select optimal hub points.

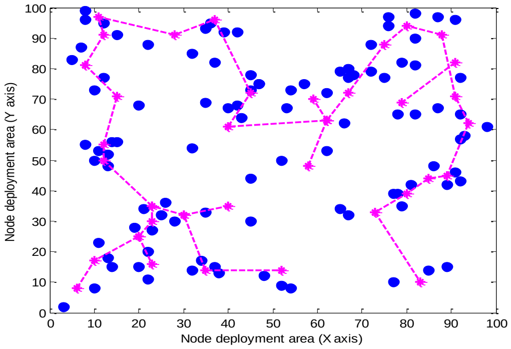


Fig. 4. Sink mobility path between the diagnosis hub points.

Algorithm 2: Optimal number of diagnosis hub point's selection

Input: $S = \{s_1, s_2, \ldots\ldots.., s_n\}, A_{curr}$

Output: Diagnostic hub point set $H_{optimal}$

```
/*Initialization Phase*/
Create an empty set $H_{curr}$
Create an empty set $H_{optimal}$
Create a set $T_{curr}$ containing all non-faulty sensor
while $T_{curr} \neq \Phi$
    Each node identifies $nb(t)$.
    Select diagnostic hub point $h$ at $h_{cost} = \frac{cost(nb(t))}{k}$.
    Add the corresponding diagnostic hub point of $nb(t)$ into $H_{curr}$.
end while
while $HR_{remain} \neq \Phi$
    Each $h$ finds it neighbour set.
    if $cost\ \{nb\ (h)\} \leq cost\{r\}$ then
        Merge to a single point and select optimal diagnostic hub point at mid-point.
        Add the corresponding optimal diagnostic hub point into $H_{optimal}$.
        Remove the corresponding diagnostic hub point of $nb(h)$ from $H_{curr}$.
    end if
end while
Find an approximate shortest fault diagnosis tour on optimal diagnosis hub points in $H_{optimal}$.
```

6. Performance Evaluation

6.1 Simulation Scenario Experiments

The performance of the proposed scheme was obtained through extensive simulations using MATLAB (version 7.5) (Niu et al., 2013 & Boudries et al., 2014). The sensor nodes are scattered in 2-D space and a mobile fault detector was used to diagnose the status of the deployed sensor nodes. Different artificial faults were injected into the network under various simulation settings. We used 2.246s as the fault diagnosis waiting time (τ_{timer}) threshold where the processing time of a microcontroller was 0.403s and ZigBee communication delay was within 0.72s.

Therefore, after reaching a diagnostic hub point, the mobile fault detector required 2.246s to diagnose the failure status of all sensors under a given hub point. The proposed mobile sink based fault diagnosis scheme was evaluated and compared with the existing well-known FDWSN (Lee and Choi, 2008), and DSFD (Panda and Khilar, 2012) distributed fault detection schemes in the literature in terms of Detection Accuracy (DA), False Alarm Rate (FAR), False Positive Alarm Rate (FPAR) and other important network parameters such as energy consumption and network lifetime. The common feature among them is the use of single hop neighbor information to diagnose the fault status of deployed sensor nodes. On the other hand, we also compared our proposed scheme with the existing PFDWSN (Lau et al., 2014) centralized fault detection scheme to verify fault detection accuracy of the proposed scheme. A common feature among them is the use of data comparison result to diagnose the fault status of the network. The fault detection performance of the proposed algorithm was evaluated using the DA, FAR, and FPAR metrics; where DA is the ratio of the number of detected faulty nodes as originally faulty and the total number of faulty nodes existing in the network, FAR is the ratio of the number of fault free nodes detected as faulty to the total number of fault free nodes present in the network, and FPAR is the ratio of the number of faulty nodes detected as fault free to the total number of faulty nodes existing in the network. Furthermore, the energy saving performance of the proposed algorithm and the network lifetime was evaluated where the initial energy of the each static sensor node is $E_0 = 0.5$ [Joules] (Banerjee et al., 2014 & Chanak et al., 2014). Other simulation parameters are listed in Table 4.

We assumed that each deployed sensor node in the network has a unique ID. Initial condition of the simulations, we assumed no collisions or retransmissions occur during wireless communications. In addition, we assumed that deployed sensor nodes can predict the length of the communication links using received signal strength. As sensor nodes, we used the TelosB motes which have an adjustable transmission range of R_{max}=100m. These modes are Zigbee/802.15.4 compliant so sensor nodes can both communicate with each other and also with MS. In this work, we used a simplified power model of radio communication as it is used in (Banerjee et al., 2014; Chanak et al. 2014; Lau et al., 2014). The energy consumption by s_i sensor node for single message transmission is represented as:

$$E_T(s_i) = (E_{elec} + \varepsilon_{amp} d^{\kappa})B$$

where E_{elec} is the basic energy consumption of sensor board to run the transmitter or receiver circuitry, ε_{amp} is the energy consumption of the amplifier, d is the distance between source and destination nodes, κ is the channel pathloss exponent of the antenna which is affected by the radio frequency environment and satisfies $2 \le \kappa \le 4$. B is the message size. On the other hand, the energy consumption at the receiver end is represented by:

$$E_R(s_j) = E_{elec}B$$

In our energy model, the noise and environmental factor are constant, only source node (s_i) can adjust its transmission power to make $E_T(s_i)$ reach a minimum value.

Table 4
Simulation parameters

Parameter	Value
Deployment area	100×100 m^2
Number of nodes	50 to 200
Data packet size	500 bits
Initial energy of each node (E_0)	0.5 Joules
Transmission power (e_{tx})	50 nJ/bit
Receiving power (e_{rx})	50 nJ/bit
Sensing range	5m
Control message size	100 bits
Sink mobility	16 m/s

6.2 Fault Diagnosis Accuracy

Fig. 5 shows the simulated DA of the proposed scheme in the presence of hardware-based faulty nodes. According to the fault probability previously presented (Lau et al., 2014 & Panda and Khilar, 2012), we also randomly introduced hardware faults in 5% of the deployed sensor nodes. When considering DA, the proposed scheme outperformed PFDWSN by a further 27%, likewise 20% for FDWSN and 16% for DSFD. Due to this proposed mobile fault detector based diagnosis strategy, the hardware of the deployed sensor nodes are interrogated by single hop communication thus causing a decrease in FPAR and an increase in DA.

We then randomly set 5% of the sensor nodes to have hardware faults and a further 5% of the sensor nodes to have software functions disabled. Among the 10% faulty nodes, we also set both hardware and software faults in 4% of the sensor nodes. In Fig. 6, it can be seen that the proposed scheme DA has improved by a further 35% compared to PFDWSN, with 21% for FDWSN and 18% for DSFD in the presence of the hardware and software faulty nodes over the network. The significant improvements seen are due to the proposed scheme diagnosing both hardware and software components separately.

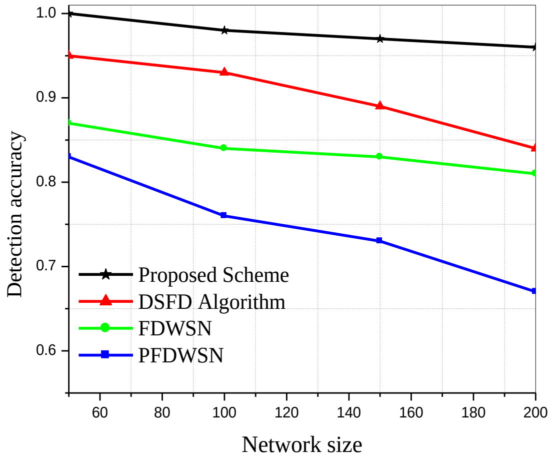


Fig. 5. Hardware failure detection accuracy.

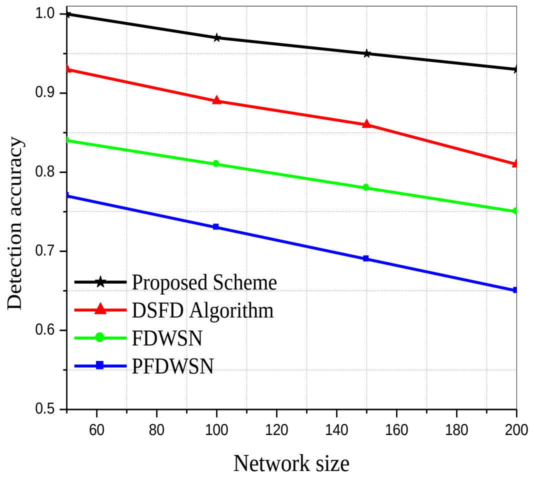


Fig. 6. Hardware and software failure detection accuracy.

6.3 False Alarm Rate Comparison

Fig. 7 compares the FAR metric between the PFDWSN, FDWSN, and DSFD algorithms and the proposed fault diagnosis scheme. The results show that the FAR of our proposed scheme is 54% less than PFDWSN, 48% less than FDWSN, and 42% of the DSFD algorithm in the presence of 5% of the deployed nodes having hardware faults.

Similarly, Fig. 8 shows the FAR in the presence of 5% hardware faults and a further 5% software faulty nodes where we set both hardware and software failures in 4% of the deployed sensor nodes. The results show that our proposed scheme also has an improved FAR compared to PFDWSN, FDWSN, and DSFD algorithms respectively. In our proposed scheme, the Fault Diagnosis Accuracy (FDA) is increased and the FAR decreased compared to the schemes in the literature because our approach uses the mobile fault detector to check each hardware circuit of the WSN nodes.

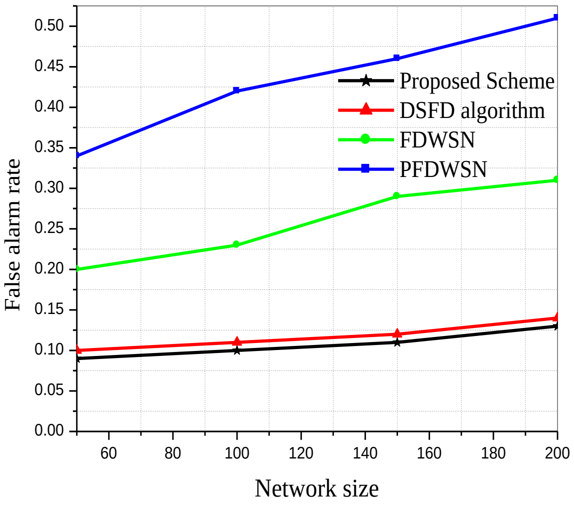


Fig. 7. Hardware failure false alarm rate.

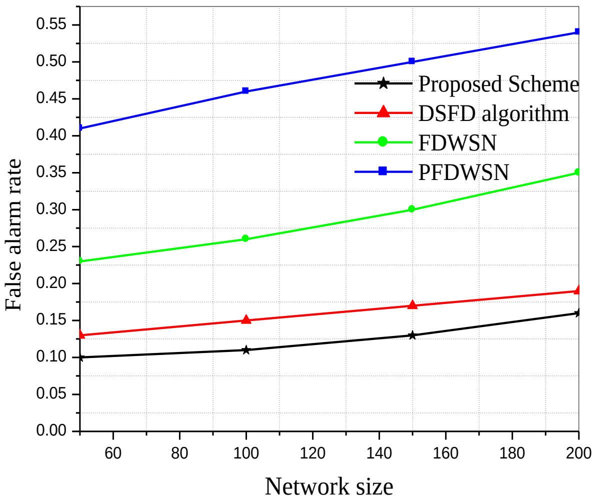


Fig. 8. Hardware and software failure false alarm rate.

6.4 False Positive Rate Comparison

The False Positive Rate (FPR) is the ratio of the number of faulty nodes that are incorrectly identified as fault free to the total number of faulty nodes present in the network. Fig. 9 and Fig. 10 show the FPR of our proposed scheme. Compared to the existing static sink based fault detection schemes (PFDWSN, FDWSN, and DSFD), the results are improved due to the single hop communication nature between the static sensor nodes and the mobile fault detector. Also, in the proposed scheme, FPR is decreased due to detailed hardware and software analysis performed during the fault diagnosis process itself.

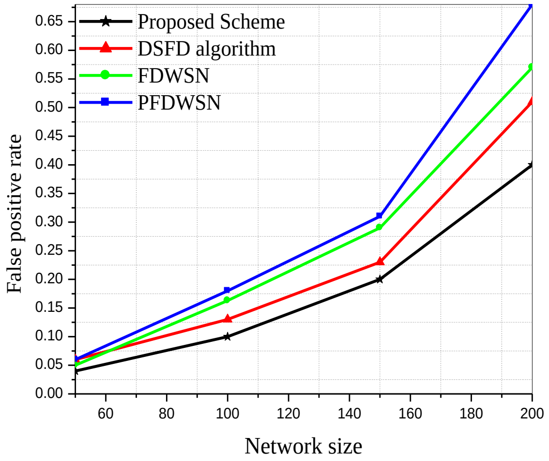


Fig. 9. Hardware failure false positive alarm rate.

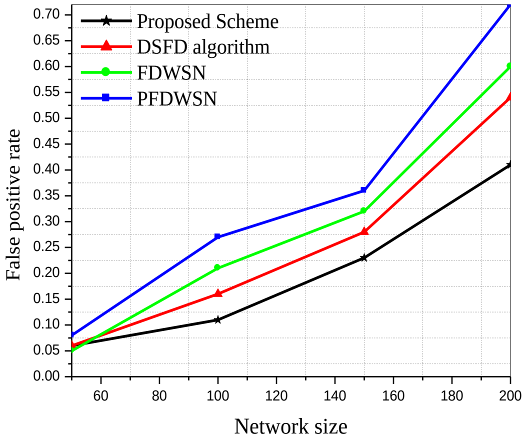


Fig. 10. Hardware and software failure false positive alarm rate.

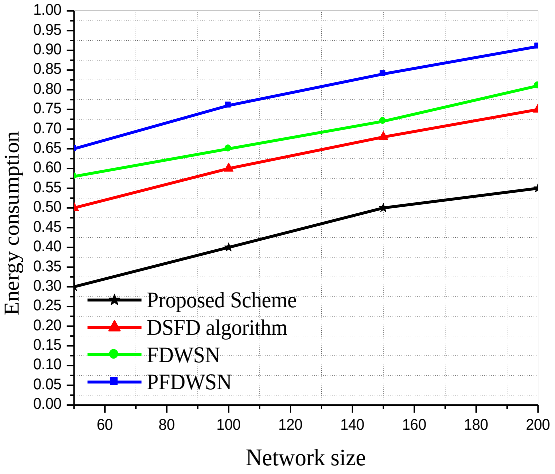


Fig. 11. Energy consumption rate during the fault diagnosis process.

6.5 Total Energy Consumption over the Network

The total energy consumption of the network depends on the total number of messages transmitted and received by the deployed static sensor nodes during the diagnosis process. Fig. 11 depicts the energy consumption of the schemes in the literature compared to our method as a function of the number of nodes. The total energy depletion of the proposed algorithm outperformed PFDWSN by a further 48%, likewise 42% for FDWSN and 40% for DSFD. The energy reduction is primarily due to the proposed scheme using single hop routing and having a reduced total number of messages within the network during the fault diagnosis process.

6.6 Network Lifetime

The network lifetime for all the approaches as a function of network size is shown in Fig. 12. The network lifetime of proposed fault diagnosis scheme has increased in our scheme by a further 37% compared to PFDWSN, 45% compared to FDWSN, and 70% compared DSFD algorithm respectively. The improvement in the network lifetime is due to the elimination of communication overhead because the mobile fault detector physically moves to each static sensor node and directly checks health status of the deployed sensor nodes by single hop communication, thereby reducing communication overhead within the proposed scheme.

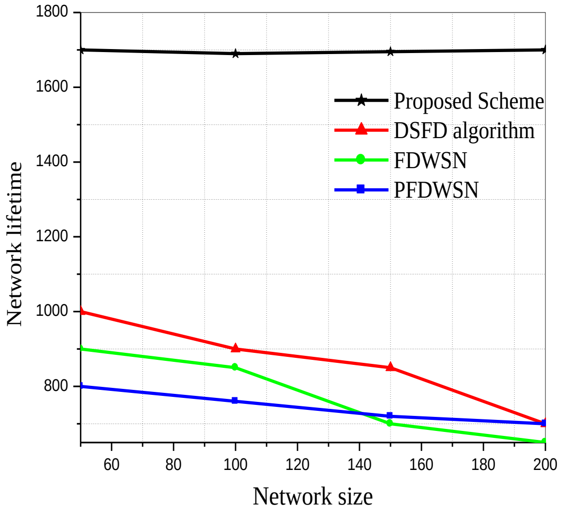


Fig. 12. Network lifetime as a function of network size (node number).

6.7 System Implementation and Testing

This section describes the hardware and software implementation of the proposed MS based fault diagnosis scheme tested in an outdoor scenario. As shown in Fig. 13, a wireless network of 8 fixed nodes was deployed outdoors on a grassy field at the Indian Institute of Engineering Science and Technology, India, for a total deployment area of approximately 15×8m^2. There was one coordinator which was connected with a laptop computer (acting as the BS). The proposed algorithm was implemented using the C language and deployed in mica2 motes running the TinyOS operating system. The detailed parameters and corresponding values used in real experiments are summarized in Table 5 and Table 6.

In the DSFD approach, the transmission range was simply assumed to be a disk-shaped area around the transceiver. Based on this assumption, the neighbor set of a sensor node consists of all the sensor nodes within the disk-shaped area around this node. However, due to the uncertainties of a wireless environment, such as signal fading, reflection from walls and obstacles, and interference, it is difficult to estimate the boundary of the transmission range without real measurements. Therefore, in practice, it is infeasible to obtain the neighbor set of an unknown deployed sensor node. Hence, in the hardware experiments we have not implemented the DSFD approach.

Initially, the mobile fault detector and BS discovered 8 diagnosis hub points. After optimization, the BS finally selects 5 optimal diagnosis hub points and selects a shortest path for fault status identification of the 8 deployed static sensor nodes. The mobile fault detector started its fault detection tour from the BS, traversed the network, and detected faulty sensor nodes while moving through the network. At the end of the fault diagnosis tour the mobile fault detector uploaded the health status of the 8 static sensor nodes to the BS with the help of the network coordinator. The detailed battery conditions of deployed sensor nodes with their initial charge condition are given in Table III.

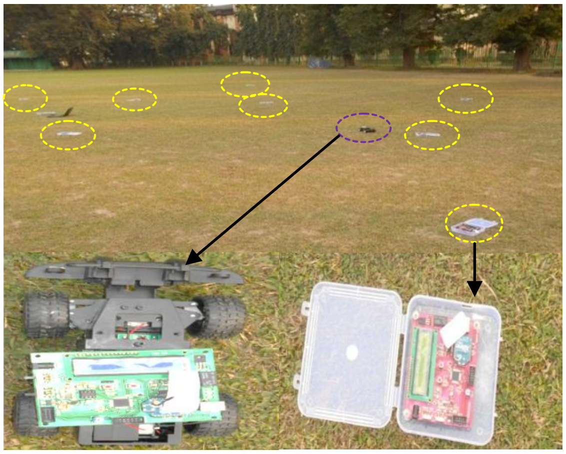


Fig. 13. Network topology in the outdoor test.

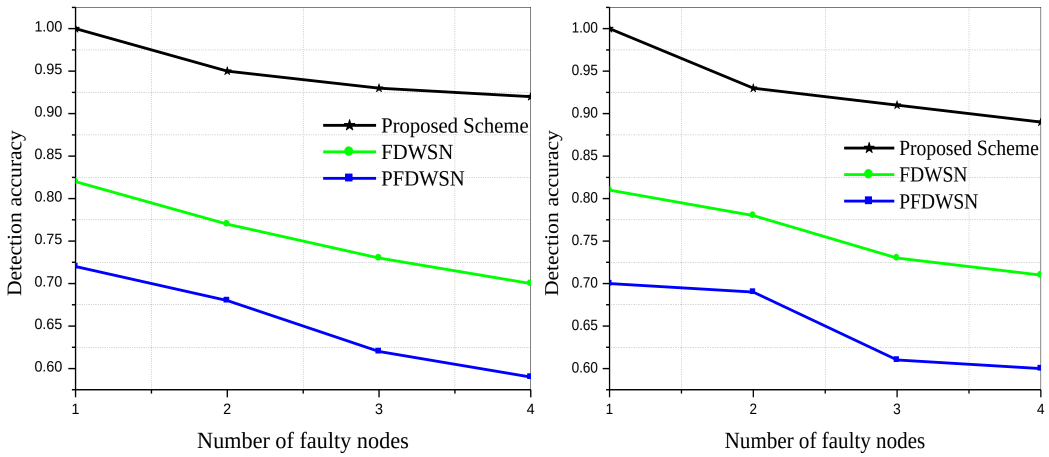
Table 5
Parameters Values Used in the Outdoor Testbed Experiment

Deployment area	20 m×40 m
Deployment distribution	Uniform
Distance between two adjacent nodes	10 m
Number of nodes	8
Mobile sink	1
Speed of the mobile sink	1.7 mph
Coordinator	1
Transmission power	4.5 dBm
Sending rate	1 packet/s
Packet size	120 bit

Table 6
Parameters Used in the Discharge Experiment

Initial discharge voltage of battery	1.603v
Test resistance	4Ω
Internal resistance of battery	34.9 mΩ
Discharge time	8h

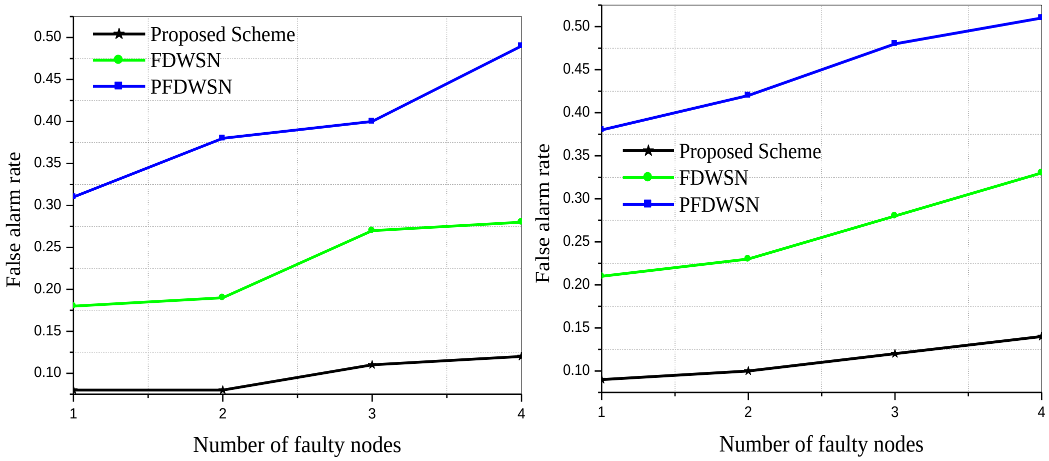
To test and verify the proposed MS based fault detection scheme, 4 sensor nodes at different positions in the implemented WSN were made faulty with differing hardware or software failures. A simple vehicle with a battery powered node was used as the mobile fault detector where it started the fault detection tour from the BS and reached the nearest diagnosis hub point with a relatively constant speed. Then, mobile fault detector polled each of the nearest nodes for their health information. Fig. 14.a and Fig.14.b show the FDA in the outdoor test. It can be seen that the FDA of the proposed scheme has been improved compared to the FDWSN and PFDWSN approaches. As described in the software simulation result, our proposed scheme shows an improved result because of the detailed hardware and software analysis stage.



(a) (b)

Fig. 14. Fault detection accuracy in the outdoor test: a) hardware failure, b) hardware and software failure.

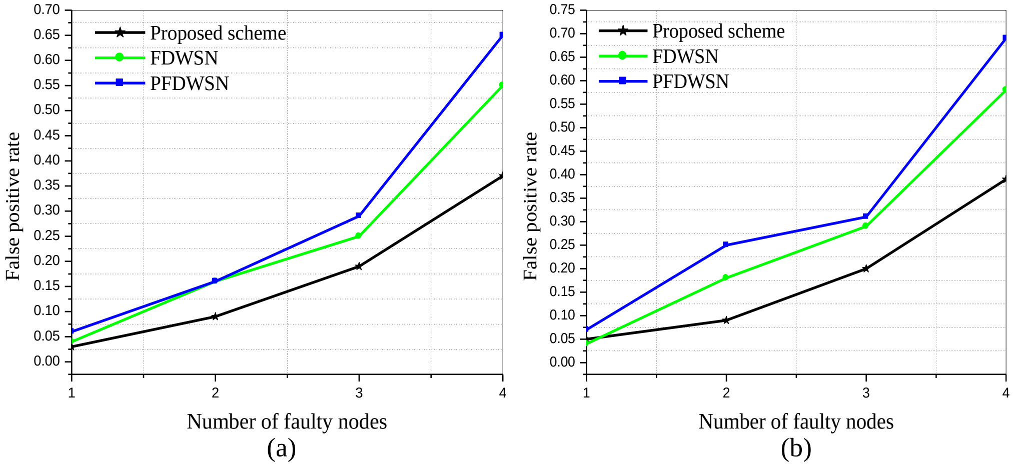
Fig. 15.a and Fig.15.b illustrate the FAR found in the practical outdoor tests. As predicted by the simulation tests, it can be seen that the FAR of the proposed scheme was increased when the number of faulty sensor nodes were increased within the network. From these results, it can be observed that our proposed scheme shows improved performance compared to the existing static sink based FDWSN, PFDWSN fault diagnosis approaches in the real hardware experiments.



(a) (b)

Fig. 15. False alarm rate: a) hardware failure, b) hardware and software failure.

Fig. 16.a and Fig. 16.b show that our proposed scheme has a reduced FPR compared to FDWSN and PFDWSN schemes, verifying the simulation results. From these figures, it is clear that the proposed scheme demonstrates desirable performance for faulty node identification.



(a)

(b)

Fig. 16. False positive rate: a) hardware failure, b) hardware and software failure.

Fig. 17 shows the average battery voltage on the deployed sensor nodes measured over time. As shown, the initial average voltage of each deployed sensor node was 3v and as expected the average battery voltage decreases over time. At the end of the outdoor experiments the measured energy efficiency of the proposed scheme was improved by a further 39% compared to FDWSN, and 43% compared to PFDWSN, which implies the more average residual energy is left.

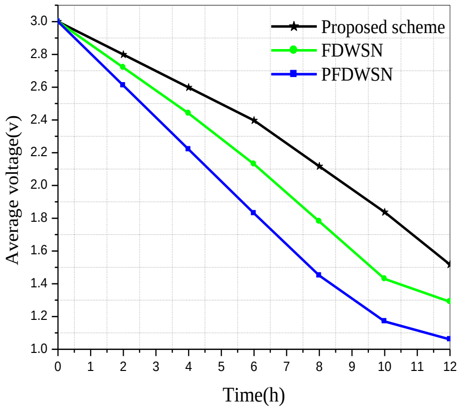


Fig. 17. Average battery voltage over time in the outdoor test.

7 Conclusions

In this paper, we have proposed a novel distributed mobile sink based fault diagnosis scheme for wireless sensor networks that outperforms the existing schemes in the literature because it only uses single hop communication. In our scheme, a mobile fault detector starts the fault diagnosis tour periodically from the BS, traverses the entire network, diagnoses fault status of the deployed sensor nodes, and then ends the fault diagnosis tour by returning to the BS to upload the network status information. The proposed mobile fault detector has been extensively simulated and also tested in an outdoor testbed. Simulation results show that the proposed scheme outperforms existing fault diagnosis algorithms in both simulated and actual network scenarios. Hence, the proposed fault diagnosis scheme improves the scalability and intrinsic problem of detecting faults in distributed homogeneous networks. By introducing the mobile fault detector, fault detection and diagnosis has become more flexible and adaptable under dynamic environments where the network structure of the wireless sensor network frequently changes due to the inherent self-organization. The proposed scheme can successfully detect faulty sensor nodes under severe environments. Abundant simulation results and a real testbed show that the proposed mobile sink based fault diagnosis algorithm makes significant improvements in all areas compared to other algorithms in the literature for fault detection accuracy, false alarm rate, false positive alarm rate, and energy. This work has clear real life applications, for example home automation, home monitoring, home healthcare, livestock monitoring and security.

References

Banerjee I., Chanak P., Samanta T., and Rahaman H., 2014. EFDR: effective fault detection and routing scheme for wireless sensor network. *J. Computers and Elec. Eng.* 40 (2), 291–306.

Bari A., Jaekel A., Jiang J., and Xu Y., 2012. Design of fault tolerant wireless sensor networks satisfying survivability and lifetime requirements. *J. Computer Com.* 35(3), 320–333.

Boudries A, Aliouat M, Siarry P., 2014. Detection and replacement of a failing node in the wireless sensors networks. *Computers and Electrical Engineering* 40 (2), 421–432.

Chanak P., Banerjee I., Wang J., and Sherratt R.S., 2014. Obstacle avoidance routing scheme through optimal sink movement for home monitoring and mobile robotic consumer devices. *IEEE Trans. Consumer Electronics* 60 (4), 596–606.

Chen J., Kher S., and Somani A., 2006. Distributed fault detection of wireless sensor networks. in *Proc. ACM Workshop Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks*, Los Angeles, CA, pp. 65-72.

Chen X, Kim Y.A., Wang B., Wei W., Shi Z. J., and Song Y., 2012. Fault-tolerant monitor placement for out-of-band wireless sensor network monitoring. *J. Ad Hoc Networks* 10(1), 62–74.

Ding M., Chen D., Xing K., and Cheng X., 2005. Localized fault-tolerant event boundary detection in sensor networks. In *Proc. Ann. Joint Conf. IEEE Computer and Communications Societies*, Miami, FL, pp. 902-913.

Gao S., Zhang H., and Das S. K., 2011. Efficient data collection in wireless sensor networks with path-constrained mobile sinks. *IEEE Trans. Mobile Computing* 10(4), 592–608.

Huang F. J., Chang Y.G., Chen H. G., 2015. A historical-beacon-aided localization algorithm for mobile sensor networks. *IEEE Trans. Mobile Computing* 14(6), 1109–1122.

Jiang P., 2009. A new method for node fault detection in wireless sensor networks. *J. Sensors* 9(2), 1282-1294.

Koushanfar F., Potkonjak M., and Vincentelli A.S., 2003. On-Line fault detection of sensor measurements. In *Proc. IEEE Conf. Sensors*, Toronto, Canada, pp. 974-979.

Lee M H, and Choi Y.H., 2008. Fault detection of wireless sensor networks. *J. Computer Com.* 31(14), 3469–3475.

Lee MH. and Choi Y. H., 2008. Localized Detection of Faults in Wireless Sensor Networks. In *Proc. 10th Int. Conf. on Advanced Communication Technology*, Gangwon-Do, Korea, pp. 637-641.

Lau B. C. P., Ma E. W. M., and Chow T. W. S., 2014. Probabilistic fault detector for Wireless Sensor Network. *Expert Systems with Applications* 41(8), 3703–3711.

Misra S., Kar P., Roy A., Obaidat S., 2014. Existence of dumb nodes in stationary wireless sensor network. *The Journal of System and Software* 91, 135-146

Mi Z., Yang Y., and Yang J. Y., 2015. Restoring connectivity of mobile robotic sensor networks while avoiding obstacles. *IEEE Sensors Journal* 15(8), 4640–4650.

Niu Shufen, Wang Caifen, Zhixuan Yu, Cao Suzhen., 2013. Lossy data aggregation integrity scheme in wireless sensor networks. *Computers and Electrical Engineering* 39(6), 1726-35.

Panda M. and Khilar M. P., 2012. Energy efficient soft fault detection algorithm in wireless sensor networks. In *Proc. IEEE Int. Conf. Parallel, Distributed and Grid Computing*, Himachal Pradesh, India, pp. 801-805.

Ramachandran G., Daniels W., Matthys N., Huygens C, Michiels S., Joosen W., Meneghello J., Lee K., Cañete E., Rodriguez D.M., and Hughes D., 2015. Measuring and modeling the energy cost of reconfiguration in sensor networks. *IEEE Sensors Journal* 15(6), 3381–3389.

Ramanathan N., Chang K., Kohler E., and Estrin D., 2005. Sympathy for the sensor network debugger. In *Proc. ACM Conf. Embedded Networked Sensor Systems*, San Diego, USA, pp. 255-267.

Ruiz B. L., Nogueira M.J., and Loureiro F.A.A., 2014. MANNA: management architecture for wireless sensor networks. *IEEE Communications Mag.* 41(2), 116–125.

Wu X., Chen G., and Das SK., 2008. Avoiding energy hole in wireless sensor network with nonuniform node distribution. *IEEE Trans. Parallel and Distributed Systems* 9 (5), 710–720.

You Z., Zhao X., Wan H., Hung N.N.W., Wang Y., and Gu M., 2011. A novel fault diagnosis mechanism for wireless sensor networks. *J. Math. and Computer Model* 54(1-2), 330–343.

Zahhad M., Ahmed S., Sabor N., and Sasaki S., 2015. Mobile sink based adaptive immune energy-efficient clustering protocol for improving the lifetime and stability period of wireless sensor networks. *IEEE Sensors Journal* 15(8), 4576–4586.

***Biography**

Biographies

Prasenjit Chanak received the B.Tech degree in information Technology for the Institute of Engineering and Technology (IET), Jaunpur, India, the M.Tech degree in information technology from Bengal Engineering and Science University (BESU), Shibpur, West Bengal, India, and the PhD. in Information Technology from Indian Institute of Engineering Science and Technology (IIEST), Shibpur, India. He is currently a research associate in the information technology department at the Indian Institute of Engineering Science and Technology (IIEST), Shibpur, India. He is the recipient of a senior research fellowship (SRF), CSIR, Govt. of India. His research interests are in the areas of wireless ad hoc and sensor networks and wireless and mobile systems.

Indrajit Banerjee received the Bachelor degree in Mechanical Engineering from Institute of Engineers, India, the Masters in Information Technology from Bengal Engineering and Science University, and the PhD. in Information Technology from Indian Institute of Engineering Science and Technology (IIEST), Shibpur, India. He is currently an assistant professor in the information technology department at the Indian Institute of Engineering Science and Technology (IIEST), Shibpur, India. His current interests are cellular automata, wireless ad hoc and sensor network, embedded systems and pervasive computing.

R. Simon Sherratt is a Professor of Biosensors in the Department of Biomedical Engineering, School of Bioscience, at the University of Reading, UK. He received his B.Eng. degree from Sheffield City Polytechnic, UK, and his M.Sc. and PhD from the University of Salford, UK, all in Electronic Engineering. His primary research topic is signal processing and communications in consumer electronic devices. He is an IEEE Fellow and an IET Fellow.

***Title page with author details**

Mobile Sink Based Fault Diagnosis Scheme for Wireless Sensor Networks

Prasenjit Chanak*,[a], Indrajit Banerjee[a], and R. Simon Sherratt[b]

[a]Department of Information Technology, Indian Institute of Engineering Science and Technology, Shibpur, Howrah-711103, India
[b]Department of Biomedical Engineering, University of Reading, Reading, Berkshire, RG6 6AY, UK

Email: prasenjit.chanak@gmail.com(P.Chanak), ibanerjee@it.iiests.ac.in(I.Banerjee), sherratt@ieee.org (R.S. Sherratt)