Pers Ubiquit Comput (2013) 17:1025–1034
DOI 10.1007/s00779-012-0549-7

ORIGINAL ARTICLE

# Towards augmenting federated wireless sensor networks in forestry applications

**Fadi M. Al-Turjman · Hossam Hassanein · Sharief Oteafy · Waleed Alsalih**

Received: 20 November 2011 / Accepted: 15 March 2012 / Published online: 11 May 2012


**Abstract** Environmental Monitoring (EM) has witnessed significant improvements in recent years due to the great utility of wireless sensor networks (WSNs). Nevertheless, due to harsh operational conditions in such applications, WSNs often suffer large-scale damage in which nodes fail concurrently and the network gets partitioned into disjoint sectors. Thus, reestablishing connectivity between the sectors, via their remaining functional nodes, is of utmost importance in EM, especially in forestry. In this regard, considerable work has been proposed in the literature tackling this problem by deploying Relay Nodes (RNs) aimed at reestablishing connectivity. Although finding the minimum relay count and positions is NP-Hard, efficient heuristic approaches have been anticipated. However, the majority of these approaches ignore the surrounding environment characteristics and the infinite 3-dimensional (3-D) search space that significantly degrades network performance in practice. Therefore, we propose a 3-D grid-based deployment for RNs in which the relays are efficiently placed on grid vertices. We present a novel approach, named fixing augmented network damage intelligently, based on a minimum spanning tree construction to re-connect the disjointed WSN sectors. The performance of the proposed approach is validated and assessed through extensive simulations, and comparisons with two main stream approaches are presented. Our protocol outperforms the related work in terms of the average relay node count and distribution, the scalability of the federated WSNs in large-scale applications, and the robustness of the topologies formed.

F. M. Al-Turjman (✉) · H. Hassanein · S. Oteafy
School of Computing, Queen's University,
Kingston, ON K7L 3N6, Canada
e-mail: fadi@cs.queensu.ca

H. Hassanein
e-mail: hossam@cs.queensu.ca

S. Oteafy
e-mail: oteafy@cs.queensu.ca

H. Hassanein · W. Alsalih
Department of Computer Science, King Saud University,
P.O. Box 51178, Riyadh 11543, Saudi Arabia
e-mail: wsalih@ksu.edu.sa

**Keywords** Wireless sensor network · Sparse connectivity · Relay placement · Grid deployment · Environmental applications

**Abbreviations**

| | |
|---|---|
| WSN | Wireless sensor network |
| RNs | Relay nodes |
| EM | Environmental monitoring |
| FADI | Fixing augmented network damage intelligently |
| MEMS | Micro-electromechanical systems |
| MST | Minimum spanning tree |
| BS | Base station |
| GUPS | Grid unit potential set |
| MGUPS | Maximal GUPS |
| TtP | Time to partition |
| ND | Node density |
| PoF | Probability of failure |
| SwMSP | Steiner with minimum steiner points |
| MSTA | Minimum spanning tree approach |
| CC | Connected component |

## 1 Introduction

Many factors are attributed to the wide use of wireless sensor networks (WSNs) in today's applications and across different environments. Improvements in micro-electro-mechanical systems (MEMS), transceiver hardware, sensing platforms and energy harvesting have all aided the design of more efficient WSNs. As such, WSNs are now deployed in many capacities over various domains, most notably in Environmental Monitoring (EM). Generally identified as harsh environments for WSNs, they extend to cover natural disasters, such as volcanoes, floods [2] and fires in forests [1].

As these scenarios encompass many harsh physical factors, they are more prone to failures. The scope of failure does not affect nodes in singularity, but often significant sectors of the deployed WSN; causing sizable partitioning in the underlying topology. Since multiple networks are often deployed, at large, to serve multiple applications, it is imperative to maintain connectivity between them to achieve the global goal of efficient and real-time monitoring of that environment. This entails maintained connectivity even under high probabilities of failure. We refer to failures at the node level as $P_{\mathrm{NF}}$ and at link levels as $P_{\mathrm{LF}}$ to caliber network operation. Understanding the performance of the global network under these failures, and the resulting connectivity measures, dictate the effectiveness of the WSN witnessing federation.

Establishing, or often re-establishing, connectivity has been approached in multiple ways in the literature of WSNs. Mainstream approaches include deploying relay nodes (RNs) to establish (often multiple) paths in the network [3, 13] as a whole, or utilizing mobile nodes that are able to re-connect partitions by moving into a median location [4, 15]. A solution scenario of the former is seen in [3] whereby a minimum threshold for relay nodes is established to deploy in a disconnected static WSN to regain connectivity. Another example of the latter [4] adopt mobile nodes that would re-locate to establish $k$-connectivity properties, as required by the application. This is also optimized by determining a minimum on the number of nodes that need to re-position to establish this metric.

The previous approaches could be utilized in benevolent environments, where probabilities of failure are under control, or witness a pattern that we can model. Nevertheless, in forestry EM scenarios, reestablishing connectivity between federated networks entails more hindrances. Dominantly, the irregularity of communication regions of such networks, not adhering to the regularly assumed disk-shape [14], dictate a hard-to-model partitioning problem; hence, making the establishment of re-connection zones a major issue. Challenges from significant distances between sectors, which might reach further than twice the communication range of a RN, added to the cost of RNs, are among the dominant hindrances.

Our contribution solves the problem of RN placement by dissecting the problem to polynomial-time operations; to leverage the intractable issues with the inherently complex re-connection problem. We adopt a two-fold approach to this problem, namely Fixing Augmented network Damage Intelligently (FADI). The base protocol adopts a grid-based approach to dissect the search space into a set of finite points, where in RNs would be deployed to reestablish connectivity. The second fold uses the derived sets of points to determine the optimal assignment of RNs to these points; herein minimizing the number of relay nodes while maintaining connectivity. Moreover, the $P_{\mathrm{NF}}$ and $P_{\mathrm{LF}}$ metrics form constraints on the optimization problem.

The remainder of this paper is organized as follows: Sect. 2 outlines the background to this problem, and the related work carried out in re-connecting federated WSNs. Then, Sect. 3 details the system level assumptions and parameters, all invoked by the problem definition. The proposed approach (FADI) is presented in detail in Sect. 4, with elaborated explanations and full pseudocode for all the underlying algorithms in this approach. This is followed by performance evaluation in comparison with two dominant contenders in this domain in Sect. 5. The paper is concluded in Sect. 6 with directions of future work.

## 2 Background and related work

Failures in WSNs is a common phenomenon that requires low-cost and real-time maintenance schemes. One of the most common failures is loss of links, which hinders network communication, sometimes resulting in complete network partitioning. In networks where co-processing takes place, especially when information fusion is utilized, network partitioning could be detrimental to its operation. WSNs can be federated either by employing mobile nodes in the originally deployed network, or by populating a few relay nodes based on the network damage size. In this paper, we focus on the latter approach due to its cost efficiency and applicability in outdoor large-scale forestry environments. In [7], Lloyd and Xue opt to deploy the fewest RNs such that each sensor node is connected to at least one RN, and the inter-RN network is strongly linked by forming a Minimum Spanning Tree (MST), and employing a Geometric Disk Cover algorithm. While in [8], the authors solve a Steiner tree problem to deploy the fewest RNs. Although the Steiner tree approach may guarantee the best network topology, it may not encompass the minimum number of relay nodes, as shown in Sect. 4.

Unlike [7] and [8], Xu et al. [9] study a random RN deployment that considers network connectivity for the longest WSN operational time. The authors proposed an efficient WSN deployment that maximizes network lifetime when RNs communicate directly with the Base

Station (BS). In this study, it was established that different energy consumption rates at different distances from the BS render uniform RN deployment a poor candidate for network lifetime extension. Alternatively, a weighted random deployment approach is proposed. In this random deployment, the density of RNs deployment increases as the distance to the BS increases; thus distant RNs can split their traffic among themselves. This in turn extends the average RN lifetime while maintaining a connected WSN.

In Ref. [10], a distributed recovery algorithm is developed to address specific connectivity degree requirements. The contribution is identifying the minimal set of relay nodes that should be repositioned in order to reestablish a particular level of connectivity. Nevertheless, these references (i.e., [9] and [10]) do not minimize the relay count, which may not be cost effective in forestry applications.

In contrast, considering both connectivity and relay count was the goal of [12]. In it, Lee and Younis focus on designing an optimized approach for federating disjointed WSN segments (sectors) by populating the least number of relays. The deployment area is modeled as a grid with equal-sized cells. The optimization problem is then mapped to selecting the fewest count of cells to populate relay nodes such that all sectors are connected.

Unlike [12], in this paper, FADI considers the relay count and neighborhood degree in a different model. It derives positions of highest potential in establishing connectivity between the disjointed functional nodes. This in turn renders a more time efficient approach than those generated by [6] and [8]; and unlike [9] and [10], FADI addresses the network federation problem without violating WSN cost-effectiveness.

## 3 System model

The system is inherently designed as an augmentation approach for connectivity reestablishment in forestry applications. As such, it is important to rigorously define the parameters of the system according to which our model operates. This section presents the problem definition in terms of how it would relate to a general forestry application, and the framework of the required solution. An elaborate explanation follows to highlight the network parameters which our system operates upon, the governing communication metrics, and the foundation for grid-based approach adopted in the model.

### 3.1 Problem definition

Consider a WSN that underwent partitioning into multiple sectors, often the result of a physical phenomenon or cascading failure in a given region. Without loss of generality, we assume that each sector is represented by a node, which we call a functional node. This is justified by locating the nearest node to the border of the damaged region, which is connected to the rest of its sector. As such, reestablishing connectivity with that node will re-introduce a path to every other node in that sector. Thus, our problem is formulated as follows:

> Given $n$ functional nodes, determine the minimum number of relay nodes required to establish intra-functional-node connectivity, and their positions respective to the network at large.

An optimal solution to this problem can be derived from the Minimum Spanning Tree (MST) algorithm, especially beneficial due to its dominantly polynomial-time solution (e.g., using Kruskal's algorithm). Forming a MST involves spanning the functional nodes with minimum weight edges; where the weight reflects the number of relays required to establish an edge between two functional nodes. We note that edge weight in this research reflects the cost of establishing an edge between the disjointed functional nodes. In forestry applications relay nodes are the most dominant expense in network hardware.

### 3.2 Communication model

Evidently, establishing the metrics and bounds of communication is important for accurate resemblance of environmental monitoring applications, which we target in this research. Unlike traditional scenarios of flat-earth communication and simplistic paradigms, many efforts have been invested in identifying the dominant factors of communication hindrance in long-range and outdoor treed environments like forests. It is imperative that signal power faces significant decay as it travels for longer distances, yet that is not straight forward to model in our scenario domain. As per the constraints of our problem formulation, we have adopted the communication model presented in [5], referred to as the log-normal shadowing model, since it accounts for irregular communication range scenarios. Thus, we represent the signal level at distance $d$ from a given transmitter as:

$$P_r = K_0 - 10\gamma \log(d) - \mu d \tag{1}$$

which follows a log-normal distribution centered around the average power value at that point. Here $K_0$ is a constant incurred at transmission (of transceiver electronics), which is derived from the mean heights of Tx and Rx. Having $d$ as the Euclidean distance (in 3-D space) between the transmitter and receiver, and $\gamma$ as the path loss exponent, we adopt $\mu$ as a normally distributed random variable (r.v.) with zero mean and variance, that is, $\mu \sim N(0, \sigma^2)$. Since the received signal could be quantified using Pr, we devise a lower threshold on the signal level to deem communication successful. Denoting it as $P_{\min}$ over distance $d$ (between transmitter and receiver), we denote the probability of successful communication as:

$$P_c = P(P_r(d) \geq P_{\min}) \tag{2}$$

which could be presented, after substitution from (1) and algebraic manipulations, as:

$$P_c(d, \mu) = Ke^{-d\mu} \tag{3}$$

where $K_0 = 10\log(K)$. This equation emphasizes the important role of surrounding factors in the environment, in signal attenuation due to obstacles and terrain properties; not simply the direct relationship with distance. Thus, we formalize the connectivity (symmetric communication) between two nodes in the network as:

**Definition 1** A "probabilistic connection" exists between two nodes, of distance $d$ apart, if for a given threshold parameter $\tau$ we assert that $P_c(d, \mu) \geq \tau$, where $0 \leq \tau \leq 1$.

### 3.3 Network model

We regard the network as a heterogeneous WSN with two tiers. The first is formed by functional nodes and the second is a layer of RNs, which establish long-range communication across the network and to the sink. Nevertheless, where partitioning happens and when the topology dictates long-range communications, optimal placement for RNs is adopted to federate the network. It is also important to note that the involvement of MAC protocols, and the impact of network partitioning and federation on them, is of great significance. However, recent advancements in adopting efficient MAC protocols for real-time environments, as that presented by Egea-López et al. [11], render this topic beyond the scope of our research.

Since network lifetime is an aggregation of that of its nodes, we elaborate on what terminates lifetime in this model. It is important to note that in forestry applications in general, $P_{LF}$ is quite elevated. As such, a fully operable node, with significant energy reservoir, may become useless to the network when its communication capabilities are jeopardized. As such, simply measuring the remaining pool of energy at nodes is not a significant indicator. Instead, the most applicable and realistic measure of lifetime would take into consideration the connectivity of its nodes. Hence, we formally define network lifetime as the duration before a partition occurs.

### 3.4 Grid model

As the search space for possible locations for RNs is inherently intractable, the task of identifying the best candidate points is imperative yet non-intuitive. Therefore, the protocol presented in this paper identifies the potential positions for RNs as an initial phase in the solution. These positions are used in placing relays which establish the MST based on a ranked approach, to achieve an optimal RN deployment scheme. Our approach here is adopting a 3-dimensional grid that uniformly dissects the region covered by the network—into virtual cubes—and reduces the infinite search space to a discrete and finite set. Accordingly, the intersection points of these grid lines (cube corners) are referred to as grid unit centers. Eventually, RNs will only be positioned at a defined set of grid unit centers. As such, we caliber the potential of a grid unit center as a candidate for RN placement according to the nodes able to communicate to that point; based on the communication metrics outlined in Sect. 3.2. Formally, we caliber the nodal coverage of a grid center as:

**Definition 2** Given a grid unit center $c$ in the deployment space, a functional node $x$ is said to cover $c$ if and only if $P_c(x, y) \geq \tau$. Where $y$ is a RN placed at $c$.

Thus we are able to quantify the potential of a grid center, in terms of its connectivity, as an aggregation of the nodes covered by it, more formally:

**Definition 3** Connectivity potential of a grid center $c$ is proportional to the sum of functional nodes covering $c$.

Hence, choosing the most representative parameters for the communication model of the network, we are able to assign a set of potential nodes for each grid unit center, which cover it, and use that as a ranking scheme for optimal locations for RN redeployment, formally:

**Definition 4** The Grid Unit Potential Set (GUPS) of a grid center $c$ holds all the functional nodes covering $c$. The degree $D = |\text{GUPS}|$. It is Maximal GUPS (MGUPS) if there exists no set $\theta$ s.t. $|\text{MGUPS}| \subseteq \theta$. The subset of functional nodes coordinates connected to $c$ is denoted $S(c)$.

## 4 Fixing augmented network damage intelligently (FADI): the approach

The proposed approach, namely **F**ixing **A**ugmented network **D**amage **I**ntelligently (FADI), is presented in this section. Given a network facing significant dissection/partitioning, we underline the procedural approach to efficiently locate positions for RN placement and detail the scheme for re-connecting it. Since the system is based on interchangeably operating procedures, the protocols are elaborately explained and their formal algorithms are presented. Figure 1 depicts our approach in light of the algorithms presented below; an example of a WSN that has undergone partitioning and the federation steps is highlighted.

**Algorithm 1:** General operation of FADI

1. **Function FADI** ($F$: Initial Set of nodes to construct FN)
2. **Input:**
3. A set $F$ of the functional nodes' coordinates.
4. **Output:**
5. A set $GC$ of the intelligently selected
6. grid centers for RNs placement.
7. **begin**
8. **while** WSN is partitioned **do**
9. $C$ = **FindGridUnitsPotential**($F$);
10. $M$ = **FindMGUPs**($F$, $C$);
11. $GC$ = **IMST** ($F$, $M$);
12. **endwhile**
13. **end**

**Algorithm 2:** Test grid unit potential for connectivity repair

1. **Function FindGridUnitsPotential**($F$)
2. **Input:**
3. A set $F$ of the functional nodes' coordinates.
4. **Output:**
5. A set $C = \{C_i \mid \forall\, i \in$ grid units centers$\}$.
6. **Begin**
7. **foreach** grid unit center $i$ **do**
8. $C_i$:=$\emptyset$; //the set of covered functional nodes by center $i$.
9. **foreach** functional node $j$ **do**
10. Compute $P_c(i, j)$;
11. **If** $P_c(i, j) \geq \tau$
12. $C_i$:= $j \cup C_i$; //Add the coordinates of $j$ to $C_i$
13. **endif**
14. **endfor**
15. **endfor**
16. **End**

**Algorithm 3:** Testing whether a grid unit set $\boldsymbol{C_i}$ is maximal or not

1. **Function Maximal**($C_i$, $C$)
2. **Input:**
3. A set $C_i$ of functional nodes covering a specific grid
4. unit center $i$ and a set $C$ of all non-empty $C_i$ sets.
5. **Output:**
6. True if $C_i$ is MGUPS and False otherwise.
7. **Begin**
8. Search for a set $C'$ such that $C_i \subseteq C'$.
9. **If** $C' := \emptyset$ **do**
10. **return** True;
11. **else**
12. **return** False;
13. **endif**
14. **End**

**Algorithm 4:** Finding all *Maximal Grid Unit Potential Set* MGUPS

1. **Function FindMGUPs**($F$, $C$)
2. **Input:**
3. A set $F$ of the functional nodes' coordinates.
4. A set $C$ of all non-empty $C_i$ sets.
5. **Output:**
6. A set $M$ that contains one position from every MGUPS.
7. **Begin**
8. $M := \emptyset$;
9. **foreach** $C_i \in \{C\}$ **do**
10. **If Maximal**($C_i$, $C$) **do**
11. $M := \{i\} \cup M$;
12. **endif**
13. **endfor**
14. **End**

The first phase of the approach identifies the set of grid centers that would satisfy the optimality highlighted earlier. Then, the connectivity potential of each is determined. Following that the MGUPs are derived, and the MST approach is utilized to federate the network. Having the set of functional nodes to start with, we present the iterative approach as proposed in Algorithm 1. Moreover, this meta-algorithm invokes the procedures highlighted in Algorithms 2–5, serving all together to identify MGUPS and assigning RNs to them until connectivity is reestablished. The constraints, inputs and outputs are all detailed in their respecitve algorithms, as presented below.

Algorithm 1 provides the high-level pseudo code of FADI. The algorithm takes the list of nodes that are still functional as an input. The detailed pseudo-code of finding the grid centers potential for connectivity, the grid centers of the highest potential, and the set of grid centers used to place the relays constructing the MST are shown in Algorthims 2–5. Algorithm 2 associates each grid center with its connectivity potential. In lines 9–14 of Algorithm 2, FADI computes the probability of the grid unit center $i$ being connected with each functional node individually based on Eq. (3). This is repeated by lines 7–15 until all probabilities between the grid units' centers and all functional nodes are computed. Based on these probabilities, the set C is initialized in line 12. Algorithms 3 and 4 check for grid centers that have the highest potential for connectivity (see Definition 3), which we call MGUPs.

Algorithm 4 calls Algorithm 3 to test whether the set of functional nodes covering a specific grid unit center is maximal or not. In line 7 of Algorithm 3, FADI searches for any set (other than $C_i$) in $C$ that has the same functional nodes that cover the grid center $i$. If such set is found, Algorithm 3 returns false, otherwise it returns true, meaning that the set $C_i$ is maximal. After discovering the grid centers that have the highest potential for connectivity (i.e., the set $M$ in line 9 of Algorithm 1), FADI calls Algorithm 5 to construct the MST using these grid centers (or MGUPs). Line 10 of Algorithm 5 carries out a search for the closest two functional nodes. If the closet two nodes are not connected (i.e. $P_c \leq \tau$), it looks at line 12 for the minumum number of grid centers in which the relays have to be placed to connect these two nodes. After connecting the closest two functional nodes (i.e., a connected component), we iterativelly look for the next closest functional node that has to be connected to them. And, in lines 20–36 of Algorithm 5, FADI iterativley searches for the least grid centers to be used for placing the relay nodes that will connect the next closest node with the connected component.

**Algorithm 5:** Intelligent Minimum Spanning Tree (IMST)

```
Function IMST (F, M)
Input:
A set F of the functional nodes' coordinates.
A set M of the MGUPs.
Output:
A set GC of the intelligently selected grid centers for RNs placement.
begin
GC := ∅;
FN := ∅;
If Pc(i, j) ≤ τ
Search for the least grid centers c required to connect i and j, such that c ∈ M
and Pc(i,c) ≥ τ and Pc(c, j) ≥ τ;
  GC := c ∪ GC; //Add the coordinates of j to GC//
  FN := c ∪ FN;
  F := F – I – j;   //Set of Functional nodes//
endif
N = no. of remaining F nodes which are not in FN;
i = 0;
foreach  remaining node ni in F do
   Find coi;
   i = i + 1;
endfor
Let CO = { coi }
i = 0;
while i < N do
   S = Smallest coi in CO;
   GC := S ∪ GC;
   FN = FN ∪ S ∪ ni;
   F = F - ni;
   CO = CO - coi;
   i = i + 1;
   foreach  remaining node nj in F do
      Find coj;
         j = j + 1;
    endfor
endwhile
end
```

To demonstrate the optimality of our approach, we introduce the following definition.

**Definition 5** A finite set of positions $P$ is optimal iff, there exists an ideal[1] placement of the least RNs in which each relay is placed at a position in $P$.

Accordingly, we have to show that the set GC in Algorithm 1 is ideal and derived from the set of MGUPS. Thus,

**Lemma 1** *For every GUPS $\beta$, there exist a MGUPS $\alpha$ such that $S(\beta) \subseteq S(\alpha)$.*

*Proof* If $\beta$ is a MGUPS, we choose $\alpha$ to be $\beta$ itself. If $\beta$ is not a MGUPS then, by definition, there exists a GUPS $\alpha_1$ such that $S(\beta) \subseteq S(\alpha_1)$. If $\alpha_1$ is a MGUPS, we choose $\beta$ to be $\alpha_1$, and if $\alpha_1$ is not MGUPS then, by definition, there exists another GUPS $\alpha_2$ such that $C(\alpha_1) \subseteq C(\alpha_2)$. This process continues until a MGUPS $\alpha_x$ is found; we choose $\alpha$ to be $\alpha_x$. Thus, Lemma 1 holds. □

**Theorem 1** *A set P that contains one position from every MGUPS is optimal.*

*Proof* To prove this theorem, it is sufficient to show that for any arbitrary placement $Z$, we can construct an equivalent[2] placement $\bar{Z}$ in which every RN is placed at a position in $P$. To do so, assume that in $Z$, a RN $i$ is placed such that it is connected to a subset $J$ of functional nodes. It is obvious that there exists a GUPS $\beta$, such that $J \subseteq C(\beta)$. From Lemma 1, there exist a MGUPS $\alpha$ such that $C(\beta) \subseteq C(\alpha)$. In $\bar{Z}$, we place $i$ at the position in $P$ that belongs to $\alpha$, so that $i$ is placed at a position in $P$ and is still connected with all functional nodes in $J$. By repeating for all minimum number of RNs, we construct a placement $\bar{Z}$ which is equivalent to $Z$, and thus Theorem 1 holds. □

**Lemma 2** *The GC set found by FADI is unique and has the least D.*

*Proof* This can be proved by contradiction. Assume FADI can find a relay node placement $A$ which contains the positions of the least RNs required to establish edges between the disjointed functional nodes. For contradiction, assume $A$ is not unique. Then, there is another placement $B$ in which the same relays count is used. Let $e_1$ be an edge that is in $A$ but not in $B$. As $B$ forms an MST, $\{e_1\} \cup B$ must result in a cycle $C$ in the federated network graph. Then $B$ should include at least one edge $e_2$ that is not in $A$ and lies on $C$. Assume the weight of $e_1$ is less than that of $e_2$. Replace $e_2$ with $e_1$ in $B$ yields the spanning tree $\{e_1\} \cup B - \{e_2\}$ which has a smaller weight compared to $B$, thus a contradiction, as $B$ was assumed to be a MST, yet it is not. □

As for time complixity of the proposed approach, consider the following:

**Lemma 3** *Finding a MGUPS takes at most $(n - 1)$ step, where n is the functional nodes' count.*

*Proof* By referring to the proof of Lemma 1, it is clear that $|C(\alpha_x)| \leq n$, and $|C(\alpha)| < |C(\alpha_1)| < |C(\alpha_2)| < \ldots < |C(\alpha_x)| \leq n$; where $|C|$ is the cardinality of $C$. Consequently, the process of finding the MGUPS $\alpha_x$ takes a finite number of steps $\leq n - 1$. □

Thus, the following theorem holds.

**Theorem 2** *The run time complexity of FADI approach is $O(n^2 \log n)$, where n is the number of functional nodes.*

*Proof* Let $g$ be the total number of grid unit centers on the assumed grid model, which is constant and known in advance for a specific monitored site. Since the total functional nodes is equal to $n$, the time complexity of Algorithm 2 is $O(gn) = O(n)$. In Algorithm 3, we search

[1] Ideal in terms of connectivity degree.

[2] Equivalent in terms of connected functional nodes. In other words, the placement of a RN at position $i$, within the communication range of nodes $x$ and $y$, is equivalent to the placement of the same RN at position $j$ which is within the communication range of the nodes $x$, $y$ and $z$.

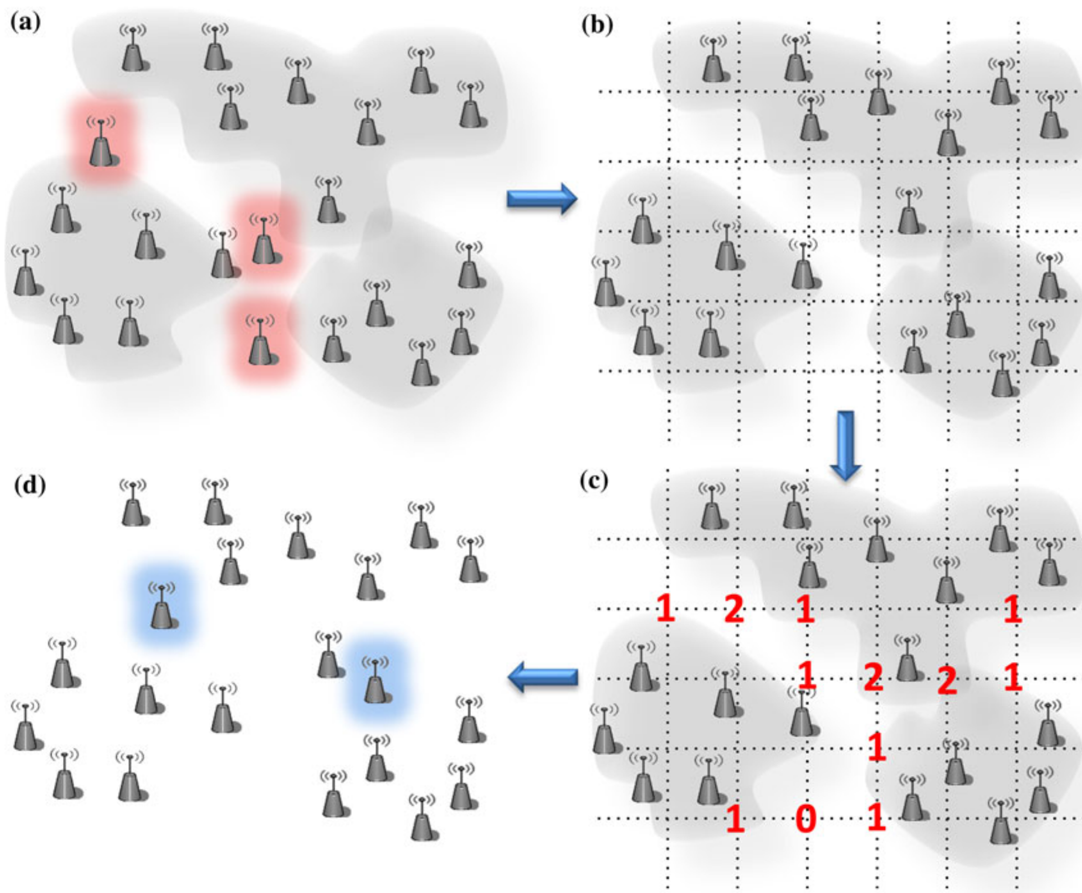


**Fig. 1** Depicts the operation of FADI. **a** In a WSN where some pivotal nodes are lost (*with glowing background*), federation is required to reestablish connectivity. **b** Grid construction takes place to test grid unit points for node coverage, shown in **c**, where the candidate MGUPS are chosen, resulting in the final deployment plan *highlighted* in **d** to restore minimum connectivity

for $C'$ such that $C_i \subseteq C'$, and this can be achieved in $O(\log(g)) = O(g)$. According to Lemma 3, Algorithm 4 will be executed in $O(n)$. As for Algorithm 5, the time complexity would be $O(n \log n)$ due to the nested loop in line 26. As Algorithm 1 iterates over Algorithms 2–5 $n$ times in the worst case of network damage where none of the functional nodes are connected, Algorithm 5 dominates the time complexity of the while loop in Algorithm 1, and thus, time complexity of Algorithm 1 is $n{*}O(n\log n) = O(n^2\log n)$. Thus, the FADI approach time complexity is $O(n^2\log n)$.□

## 5 Performance evaluation

Using MATLAB, we simulate randomly generated WSNs that have a graph topology and consist of varying number of partitioned functional nodes.[3] We simulate a realistic communication channel characteristics taken from experimental measurements in a densely treed environment [14].

### 5.1 Performance metrics and parameters

To evaluate our FADI approach, we tracked the following performance metrics:

- Average RNs degree ($D$): This is the number of functional nodes in the neighborhood of a RN. It reflects the federated network reliability under harsh forestry characteristics. It gives an indication for the federated WSN robustness. Where higher node degree yields stronger connectivity and enables better load balancing.
- Average RNs count ($Q_{RN}$): This represents the cost-effectiveness of the deployment approach and the main objective targeted by our approach.
- Recovering time (RT): The time required to federate the disjointed functional nodes and remove any partitioning.
- Time to Partition (TtP): The time span before the network experience a partition after being federated.

[3] Random in terms of functional nodes' count and positions.

Three main parameters are used in the performance evaluation:

- Number of functional nodes ($Q_{FN}$): This represents the complexity of the addressed problem.
- Node density (ND): This measures the federated network scalability in large-scale forestry applications.
- Probability of failure (PoF): is the probability of physical damage for the deployed node and the probability of communication link failure due to bad channel conditions, and uniformly affects any of the network nodes/links. We chose this parameter as it reflects the harshness of the monitored forest.

### 5.2 Baseline approaches

The performance of FADI is compared to the following two approaches: the first algorithm forms a minimum spanning tree without considering the intersections of the irregular communication ranges [6] and we call it Minimum Spanning Tree Approach (MSTA) and the second is for solving a Steiner tree problem with minimum number of Steiner points [8] and we call it Steiner with Minimum Steiner Points (SwMSP). The MSTA opts to establish an MST based on the Euclidean distance separating two functional nodes bearing in mind that the communication range is depending on the distance only. It first computes an MST for the given WSN partitions and then places RNs at the minimum number of grid vertices on the MST. The SwMSP approach places the least relays count to repair connectivity such that the maximum edge length in the Steiner tree is $\leq r$. SwMSP first combines functional nodes that can directly reach each other into one Connected Component (CC). The algorithm then identifies for every three CCs a vertex $x$ on the grid that is at most $r$ (m) away from the CC boundary nodes. A RN is placed at $x$ and these three CCs are merged into one CC. These steps are repeated until no partitioning in the network is found. In summary, both MSTA and SwMSP deployment strategies are used as baseline approaches due to their efficiency in linking WSNs partitions using the minimum relays count.

### 5.3 Simulation setup and results

The three deployment schemes, MSTA, SwMSP, and FADI, are executed on 500 randomly generated partitioned networks for statistically stable results. The average results hold confidence intervals of no more than 2 % of the average values at a 95 % confidence level. For each topology, we apply a random PoF, and performance metrics are computed accordingly. A linear congruential random number generator is employed. Dimensions of the deployment space 900*900*300 ($m^3$). Based on experimental measurements [14], we set our communication model variables as shown in Table 1.

**Table 1** Parameters of the simulated WSNs

| Parameter | Value |
|---|---|
| $\tau$ | 70 % |
| $r$ | 100 (m) |
| PoF | 35 % |
| $g_i^{RN}$ | 100 (byte/round) |
| $g_i^{SN}$ | 10 (byte/round) |
| $\gamma$ | 4.8 |

And $\mu$ to be a random variable that follows a log-normal distribution function with mean 0 and variance of 10. We assume a predefined fixed time schedule for traffic generation at the deployed WSN nodes. To simplify the presentation of results, all the transmission ranges of functional nodes and relays are assumed equal to 100 (m).

For varying number of disjointed functional nodes, Fig. 2 compares FADI approach with MSTA and SwMSP in terms of the total required relays. It shows how FADI outperforms the other approaches under different complexities of the targeted federation problem. Unlike the other approaches, the required relays count is slightly increasing when FADI approach is utilized as the total partitioned nodes are increasing.

This indicates more savings in cost which is very desirable in harsh environments targeted by large-scale forestry applications. Figure 3 depicts the efficiency of MST-based approaches in terms of time complexity with respect to other approaches such as the SwMSP. It is clear how an increment in the disjointed functional nodes leads to an exponential increment in the time required for recovering (federation) when a Steiner tree approach is

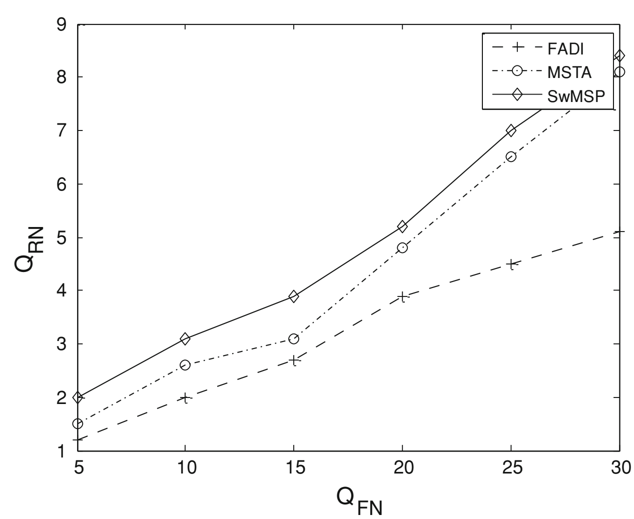


**Fig. 2** Functional nodes count versus the required relays count

utilized. This has a great draw back on forestry applications which are most often time sensitive.

Figure 4 justifies the optimality of FADI in terms of finding the least relays positions federating all partitioned functional nodes. Where the average relays degree achieved by FADI is much better than the average degree reached by MSTA and SwMSP with various counts for the disjointed functional nodes (i.e., different QFN values). This in turn provides a robust network topology structure under harsh operational conditions in the forest. In Fig. 5, we examined the practicality of our placement methodology under harsh operational conditions. FADI was able to generate a federated WSN that stays connected for long lifetime periods with respect to MSTA and SwMSP.

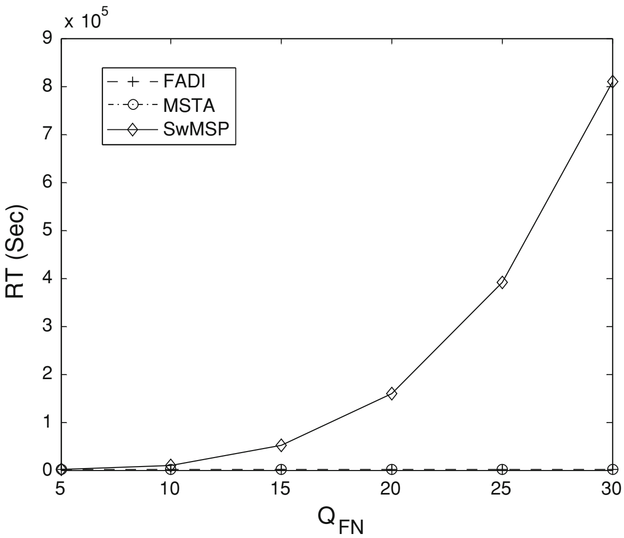


**Fig. 3** Functional nodes versus the required recovery time

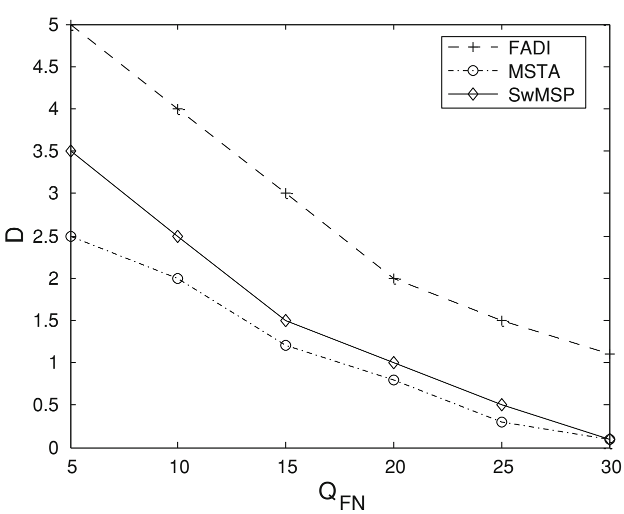


**Fig. 4** Functional nodes versus the relays' neighborhood degree

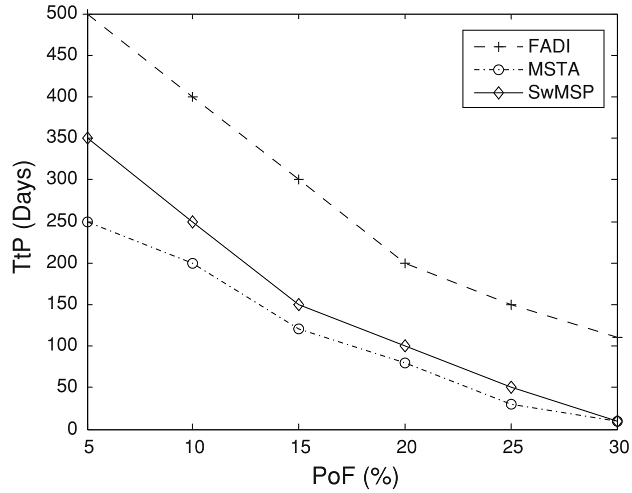


**Fig. 5** PoF versus the time to partition

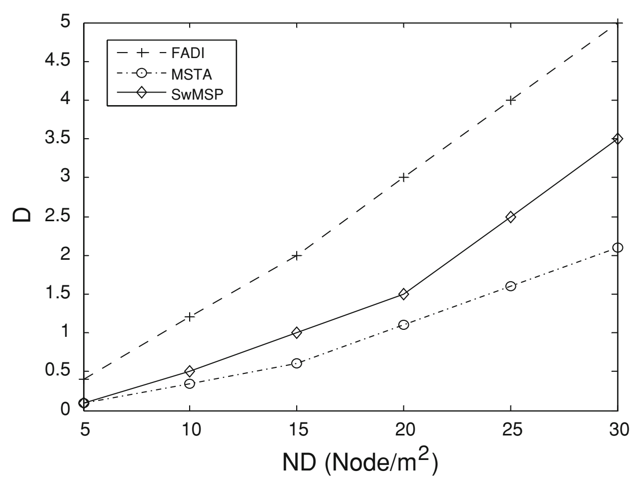


**Fig. 6** Node density versus the relays' neighborhood degree

It is also worth noting that FADI outperforms MSTA and SwMSP in terms of the relays degree, even in the presence of different node densities (ND) in the monitored site. This is shown in Fig. 6 and it points toward robust topologies in large-scale forestry applications where huge areas are targeted with a relatively small number of sensor nodes (i.e., very small ND values).

## 6 Conclusion

In this paper, we explored the problem of federating grid-based WSNs in forestry applications. A relay-based approach, called FADI, was presented using the minimum spanning tree algorithm. For practical solutions, varying probabilities of failures were considered, in addition to limiting the huge search space of the targeted deployment problem. The extensive simulation results, obtained under harsh operational conditions, indicated that the proposed approach can efficiently federate the disjointed WSNs. Moreover, the deployment approach presented in this paper can provide a tangible guide for network provisioning in large-scale environmental applications that require linking between vastly separated WSN sectors. Future work would investigate the deployment problem in further environment monitoring scenarios, where a subset of the relay nodes may have the mobility feature to repair connectivity and prolong the network lifetime. In addition, this work could extend to encapsulate a more inclusive framework for discovering nearby devices that may offer higher connectivity or act as intermittent relays, aiding in the federation scheme. The work presented by Gellersen et al. [16] facilitate a dynamic framework for discovering spatial relationships of devices with nearby (heterogeneous) devices.

**Acknowledgments** This research is funded by a grant from the Ontario Ministry of Economic Development and Innovation under the Ontario Research Fund-Research Excellence (ORF-RE) program. This research is also sponsored by the National Plan for Science and Technology at King Saud University, Project number: 11-INF1500-02.

## References

1. Younis M, Akkaya K (2008) Strategies and techniques for node placement in wireless sensor networks: a survey. Elsevier Ad Hoc Netw J 6(4):621–655
2. Hughes D, Greenwood P, Coulson G, Blair G, Pappenberger G, Smith F, Beven K (2006) An intelligent and adaptable flood monitoring and warning system. In: Proceedings of the UK E-Science All Hands Meeting (AHM'06), Nottingham, UK
3. Li N, Hou J (2005) Improving connectivity of wireless ad hoc networks. In: Proceedings of the IEEE International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous), San Diego, CA, pp 314–324
4. Abbasi A, Baroudi U, Younis M, Akkaya K (2009) C2AM: an algorithm for application-aware movement-assisted recovery in wireless sensor and actor networks. In: Proceedings of the ACM Wireless Communications and Mobile Computing Conference (IWCMC), Leipzig, pp 655–659
5. Rappaport T (2002) Wireless communications: principles and practice, 2nd edn. Prentice Hall, Upper Saddle River
6. Kim Y, Jeong S, Kim D, Lopez TS (2009) An efficient scheme of target classification and information fusion in wireless sensor networks. J Pers Ubiquitous Comput 13(7):499–508
7. Lloyd E, Xue G (2007) Relay node placement in wireless sensor networks. IEEE Trans Comput 56(1):134–138
8. Cheng X, Du D, Wang L, Xu B (2008) Relay sensor placement in wireless sensor networks. Wireless Netw 14(3):347–355
9. Xu K, Hassanein H, Takahara G, Wang Q (2010) Relay node deployment strategies in heterogeneous wireless sensor networks. IEEE Trans Mob Comput 9(2):145–159
10. Abbasi A, Younis M, Akkaya K (2009) Movement-assisted connectivity restoration in wireless sensor and actor networks. IEEE Trans Parallel Distrib Syst 20(9):1366–1379
11. Egea-López E, Vales-Alonso J, Martínez-Sala A, García-Haro J, Pavón-Mariño P, Bueno Delgado M (2006) A wireless sensor networks MAC protocol for real-time applications. J Pers Ubiquitous Comput 12(2):111–122
12. Lee S, Younis M (2010) Optimized relay placement to federate segments in wireless sensor networks. IEEE Trans Sel Areas Commun 28(5):742–752
13. Al-Turjman F, Hassanein H, Alsalih W, Ibnkahla M (2011) Optimized relay placement for wireless sensor networks federation in environmental applications. Wirel Commun Mob Comput 11(12):1677–1688
14. Al-Turjman F, Hassanein H, Ibnkahla M (2009) Connectivity optimization for wireless sensor networks applied to forest monitoring. In: Proceedings of the IEEE International Conference on Communications (ICC), Dresden, pp AHSN11.5.1–AHSN11.5.5
15. Alsalih W, Hassanein H, Akl S (2009) Routing to a mobile data collector on a predefined trajectory. In: Proceedings of the IEEE International Conference on Communications (ICC), Dresden, Germany, pp 1–5
16. Gellersen H, Fischer C, Guinard D, Gostner R, Kortuem G, Kray C, Rukzio E, Streng S (2009) Supporting device discovery and spontaneous interaction with spatial references. J Pers Ubiquitous Comput 13(4):255–264