An Optimal Design Testing Assignment for Wireless Sensor Networks

Alexander Robert Kutzke[1], Andréa Weber[1], Stefano Chessa, member IEEE[2]

[1] Federal University of Paraná – Dept. Informatics
Centro Politécnico, Jdim. Américas, 81531-990
Curitiba, Pr, Brazil

{alexander,andrea}@inf.ufpr.br

[2]Institute for Science and Technology of Information
Via G. Moruzzi, 1, 56124, Pisa, Italy
and Computer Science Department - University of Pisa
Largo Pontecorvo 3, 56127 Pisa, Italy

stefano.chessa@isti.cnr.it

Abstract. *We consider the problem of defining an energy optimal* testing assignment *for the identification of node malfunctions in Wireless Sensor Networks (WSNs). In* system-level diagnosis*, a testing or connection assignment is a set of mutual tests among the n units of a system S. If t is the number of faulty units in the system and (i) $n \geq 2t+1$; (ii) each unit is tested by at least t other units, the system is said to be t-diagnosable. A t-diagnosable system is optimal if $n = 2t+1$. In this work we prove that, given a set of $2t+1$ sensors, our* Optimal Design Testing Assignment *(ODTA) approach generates a connection assignment with minimum cost from the point of view of the energy required by the sensors to sustain it. Furthermore, we present simulation experiments that compare the present approach with our preliminary results.*

1. Introduction

Wireless Sensor Networks (WSNs) [Baronti et al. 2007] are a specific kind of ad hoc networks with sensing, processing and communication capabilities subject to stringent resource constraints. Applications of WSNs include medical diagnosis, infrastructure monitoring, agriculture and environmental sensing, between many others [Chen et al. 2010, Ko et al. 2010a, Shuman et al. 2010, Corke et al. 2010, Ko et al. 2010b, Hughes et al. 2011].

Since WSNs are often deployed for unattended operations, they need to run dependable applications. As in our previous works [Weber et al. 2010, Weber et al. 2011], we concern with an energy-aware testing approach for identifying node malfunctions in a WSN. Specifically, we introduce an optimal *testing assignment*, based on the concepts of optimal systems and optimal designs of the PMC model [Preparata et al. 1967].

PMC stands for the initials of Preparata, Metze and Chien, which proposed the first *system-level* diagnosis model. In the PMC model a system is considered as a set of units that are able to execute tests among themselves. The outcome of a test performed by a fault-free unit is assumed to be reliable (it is 0 if the tested unit is fault-free and it is 1 if the tested unit is faulty), while it is completely unreliable if the testing unit is faulty.

The system is represented as a directed graph where vertices represent the system units, and a testing link exists between units u_i and u_j if there is a communication link between them and unit u_i tests unit u_j. The set of test outcomes, known as the *syndrome*, is decoded by a centralized system supervisor using a suitable diagnosis algorithm [Dahbura and Masson 1984, Caruso et al. 2007].

In the PMC model a system is diagnosable provided the number of faulty units does not exceed a certain threshold. Furthermore, the so called diagnosability of the system is subject to some topological properties of the diagnostic graph. More specifically, a system of n units is said to be t-diagnosable if the number of faulty units does not exceed t. In addition, the following conditions must hold: $(c1)$ the number n of units in the system must be greater than or equal to $2t + 1$, and $(c2)$ a unit must be tested by at least t other units [Preparata et al. 1967].

The conditions $(c1)$ and $(c2)$ above are necessary and sufficient for t-diagnosability provided there are not reciprocal tests, i.e, no two units test each other. Furthermore, in the context of system-level diagnosis, a system S, consisting of n units, is defined as optimal if $n = 2t + 1$, where t is the number of faulty units in S and condition $(c2)$ also holds.

The strategy introduced here, called *Optimal Design Testing Assignment* (ODTA) relies on the non-reciprocity of tests and aims to reduce the number of sensors that participate to the diagnosis to a minimum. Thus the testing graph generated by the strategy ODTA always uses the optimal number of $2t + 1$ sensors.

An optimal system is defined by an *optimal design*, i.e., a set of edges, or tests, which makes the system optimal. In general, there are several optimal designs for a given system. Preparata et al. define a set or family of optimal designs; the so called $D_{\delta t}$ family [Preparata et al. 1967], based on which the strategy ODTA builds the testing graph, will be presented later in this work.

The rest of this work is organized as follows: Section 2 presents related work. The diagnosis model is presented in Section 3. Section 4 defines the energy model. The *Optimal Design Testing Assignment* is introduced in Section 5. In Section 6, simulation results are presented in comparison with our previous results. Section 7 concludes the paper.

2. Related Work

The work of Corke et al. [Corke et al. 2010] is concerned with outdoor applications of wireless sensor networks involving natural environment or agriculture like microclimate monitoring for farms and rain forests, water-quality monitoring and cattle monitoring and control. Nevertheless, the work also addresses the challenges faced by the authors to ensure the reliability of the deployed sensors and networks.

In [Santi and Chessa 2001], Chessa and Santi present a comparison based testing strategy in which the diagnosis model exploits the one-to-many communication paradigm typical of ad-hoc networks. Both hard and soft faults are considered and the diagnosis is based upon comparison of the results generated by testing tasks assigned to pairs of units with a common neighbor.

In [Weber et al. 2010] the problem of determining a connection assignment of the sensors in a WSN is considered. Two strategies are shown, one for the scenario in which reciprocal tests among sensors are possible and other for the scenario in which there are no reciprocal tests. In the strategy without reciprocal tests, the region R is partitioned into four quadrants of equal size. Sensors present in one quadrant test and are tested by sensors in other quadrants, thus avoiding the reciprocity.

In [Weber et al. 2011] an evolution of the previous approach is presented, in which a reduction of the number of tests is required for diagnosis. Both approaches are compared in that work and will also be used as a parameter for comparison to the strategy introduced in the present paper. More details will be given in Section 6.

In [Zhang et al. 2008], the authors propose a comparison-based fault locating arithmetic for multi-source network cluster nodes. The approach is based on layer-built topology structure and one-to-many communication mode. In [Choi et al. 2009] the authors present a distributed adaptive scheme for detecting faults in WSN where each node makes a local decision based on comparisons between neighbors, along with the dissemination of the decision to them. Time redundancy is used to enhance the accuracy of detection and tolerate transient faults in sensing and communication.

In [Chessa and Santi 2002] the authors propose an energy-efficient fault diagnosis protocol for wireless sensor networks. This protocol, called *WSNDiag*, is capable of diagnosing *crash* faults. The diagnosis works on demand and the protocol is capable of correct diagnosis in a system with up to t faulty units, where $t < k(G)$ and $k(G)$ stands for the connectivity of the system. In [Taghikhaki and Sharifi 2008] an energy-efficient distibuted approach improves network lifetime by detecting data faults locally in cluster heads. The sensors that belong to the same cluster share and compare their readings. From these comparisons, the cluster head verifies which sensors present more fault indications to find the set of possible faulty sensors in the cluster. In [Venkataraman et al. 2008], another energy-efficient cluster-based approach avoids performance degradation aiming at detecting in advance the failures that may cause connectivity loss.

Some works study topological properties of networks. In [Penrose 1999] a formal proof is presented for the minimum degree a network must have in order to be k-connected with high probability provided the number n of the nodes in the network is big enough. In [Xue and Kumar 2004], the authors show how many neighbors the nodes of a network with n randomly placed nodes should be connected to in order to the overall network to be connected. The problem of determining the critical transmitting range (CTR) for connectivity in mobile ad hoc networks is studied in [Santi 2005].

3. Diagnosis Model

In the PMC model [Preparata et al. 1967] a test (v_i, v_j) consists of a set of input stimuli that are produced by the testing unit v_i and sent to the tested unit v_j. In turn, v_j produces a test result that is sent back to v_i. Finally v_i compares the output produced by v_j with the expected output and it produces the test result that is a binary outcome: it is 0 if the two results match (and then the test succeeds), and it is 1 otherwise (i.e. the test fails). Thus in the PMC model the execution of the test requires a bidirectional link between v_i and v_j. However, in WSNs the tests may also be executed in presence of unidirectional communication links [Santi and Chessa 2001]: a sensor may start a self-test on a predefined set of stimuli and it may send the output to another sensor that compares it with the expected results. For this, it is sufficient only that the tested sensor be able to send its output to the tester.

In this paper we assume that links are bidirectional (an assumption that can easily enforced by disregarding unidirectional communication links), but we consider only unidirectional tests, that is, even if there is a bidirectional communication link between two sensors v_i and v_j, an unidirectional testing link between them means that either v_i tests v_j or v_j tests v_i. Thus, the diagnosability can be derived by conditions ($c1$) and ($c2$) described in Section 1 above.

In this work we are not concerned with the nature of the sensor faults, but with the network topology properties necessary to build an optimal connection-assignment. As in our previous works [Weber et al. 2010, Weber et al. 2011], in ODTA's diagnosis model we assume that sensors are deployed in a sensing area with uniform distribution. Furthermore, each sensor knows its geographical coordinates and the topology of the WSN is known to the sink. As in [Weber et al. 2011], the sink is responsible for generating the connection assignment and to inform it to the sensors.

The WSN is defined as the system graph $G = (V, E)$ where each vertex v in V represents a sensor and an edge $(v_i, v_j) \in E$ if either v_i is within the transmission range of v_j or v_j is within the transmission range of v_i or both. Eventhough PMC's model assumes bidirectional links and a fully-connected graph, this assumption may be weakened, provided all links that may be necessary to the tests exist. Note that this assumption can be enforced by properly calibrating the sensor's transmission ranges.

4. Energy Model

In order to estimate the energy consumption of a given testing assignment, we consider the one-slope model [Patwari et al. 2003], a widely used propagation model in wireless communications. This model assumes a linear dependence between the path loss (dB) and the logarithm of the distance d between the transmitter and the receiver, as expressed in 1:

$$L(d)\ |_{dB} = l_0 + 10\alpha \log_{10}(d) \tag{1}$$

where l_0 is the path loss at a reference distance of 1 meter, and α is the power decay index (also called path loss exponent). In general, to ensure a communication between a transmitter t and a receiver r placed at distance d from each other it is necessary that the packet sent by t reaches r with a power level higher than the sensitivity of the receiver. In other words, letting E_t be the transmission power of the transmitter, E_r the power of the signal at the receiver (where E_r depends on E_t and the distance d) and E_m be the sensitivity of the receiver, must be $E_r > E_m$.

Knowing that $L(d)$ is equal to the difference in decibels of the power of the signal at the transmitter and the power of the signal when it reaches the receiver [Rappaport 2001], by Equation 1, we have:

$$10 \log_{10}\left(\frac{E_t}{E_r}\right) = l_0 + 10\alpha \log_{10} d \tag{2}$$

Thus:

$$E_r = \frac{E_t}{10^{\left(\frac{l_0}{10} + \alpha \log_{10} d\right)}} \tag{3}$$

Adding $E_r > E_m$ in Equation 3 we obtain that the minimum of transmission power E_t on the transmitter that ensures that the package reaches the receiver with the power level required is:

$$E_t = E_m 10^{\left(\frac{l_0}{10} + \alpha \log_{10} d\right)} \tag{4}$$

Therefore, Equation 4 depends on the distance d, on the sensitivity of the receiver and on the parameters l_0 and α. Typical values for these parameters are used for the simulations [Rappaport 2001] (in this work we use $l_0 = 10$ and $\alpha = 3$), while E_m depends on the sensors' hardware. Developing Equation 4 we have:

$$E_t = E_m 10^{\frac{l_0}{10}} d^{\alpha} \tag{5}$$

It is shown from Equation 5 that the energy transmitted grows polynomially with the distance d, with an exponent equal to α. Thus the energy costs of the testing strategies presented in this work are based exclusively on the geographical distance between sensors. So for $l_0 = 10$ and $\alpha = 3$, we have:

$$E_t = E_m 10 d^3 \tag{6}$$

Because E_m depends on the hardware used, and in general is a small value, the energy expended is defined most as the cube of the distance between source and destination. Thus, we establish that an energy unit, or e.u., is equal to $10E_m$. Thus, we define that the energy expended by a transmission of d meters is equal to d^3 energy units. Knowing that the test consumption will be a xd^3 e.u., where x defines the number of messages needed for a test, we use only the value d^3 for comparisons purposes between testing strategies in Section 6 of this work.

5. Optimal Design Testing Assignment

Let us assume that a set T (with cardinality t) of sensors in a WSN generate alarms that are received by the sink. In turn, the sink defines a connection assignment in order to diagnose the region of the network where the alarms were generated. Then, the sink informs each selected sensor to perform the tests in which it is involved as tester in the testing assignment.

If the WSN is defined as the graph G in Section 3, the connection assignment is the testing graph $D = (V_D, E_D)$, where $V_D \subset V$, $E_D \subset E$ and an edge $(v_i, v_j) \in E_D$ if and only if v_i tests v_j. We define n as the cardinality of V_D. It should also be noted that the set $T \subset V_D$.

In order for the diagnostic graph D to be t-diagnosable, the contitions $(c1)$ (i.e. that $n \geq 2t + 1$) and $(c2)$ (i.e. the indegree of each vertex in D is at least t) should be met [Preparata et al. 1967].

The strategy presented here, an *Optimal Design Test Assignment* (ODTA) aims at reducing the number of sensors that participate to the diagnosis to a minimum, i.e., the testing graph made by the strategy ODTA always uses $2t + 1$ sensors.

The algorithm that defines ODTA's diagnostic graph D is based on the concepts of optimal systems and of *optimal designs* of the PMC model [Preparata et al. 1967]. In the context of system-level diagnosis, a system S, consisting of n units, is defined as optimal if $n = 2t + 1$, where t is the number of faulty units in S, and each unit of S is tested by exactly t other units. An optimal system is defined by an *optimal design*, i.e., a set of edges, or tests, which makes S optimal. In general, there are several *optimal designs* for a system S. Preparata et al. [Preparata et al. 1967] define $D_{\delta t}$ as a set, or family, of *optimal* designs. A system S belongs to a *design*, or graph, $D_{\delta t}$ when a test (u_i, u_j) exists in S if and only if $(j - i) \bmod n = (\delta m) \bmod n$ with m assuming the values $1, 2, \ldots, t$.

Preparata et al. prove that if a system S employs a graph D_{1t}, then S is t-diagnosable [Preparata et al. 1967]. The cyclical characteristic of the graphs of type D_{1t} produces t-diagnosable testing graphs without reciprocal tests. In their work, Preparata et al. also prove that a graph $D_{\delta t}$ generates t-diagnosable assignments if δ and t are relatively prime [Preparata et al. 1967]. For this proof, it is demonstrated that the D_{1t} graphs are isomorphic to the graph $D_{\delta t}$ when δ and t are relatively prime.

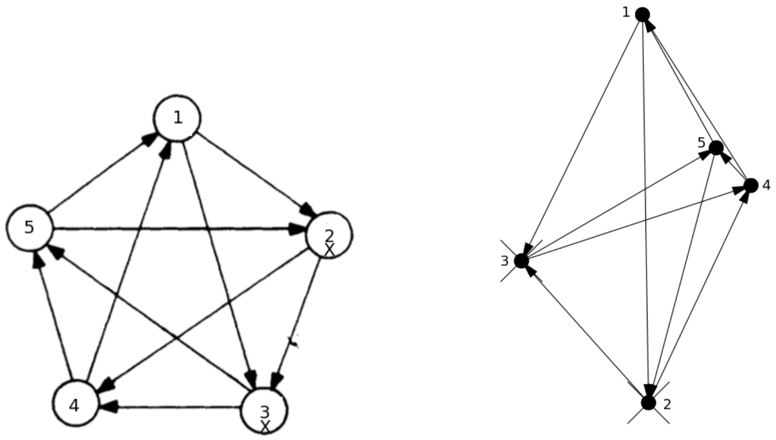
The ODTA strategy aims to define the set of edges that creates a graph D_{1t} for a system composed of the units of V_D. For this assume that the set V_D has already been defined and that $n = 2t + 1$. To define a testing graph D belonging to a family of graphs D_{1t}, each sensor present in V_D receives an unique numeric identifier i, where $i = 0 \ldots 2t$. This identifier is assigned randomly by the *sink*, creating an overlay network over the sensors of V_D and the edges that must exist in the testing graph according to the definition of the D_{1t} design. Please observe that the transmission ranges of the sensors that belong to V_D may be tunned in order to ensure that all necessary edges in E_D exist.

The graph D is then built as follows: each sensor will test the t next sensors following the increasing order of their identifiers. Thus, a sensor v_i, where i is its identifier in the overlay network, tests the sensors $v_{(i+1) \bmod n}, \ldots, v_{(i+t) \bmod n}$ on the overlay network. Therefore, each sensor will test t sensors and will be tested by t other sensors. The total energy cost for the strategy ODTA is defined as:

$$C_T(D) = \sum C_{i,j} | \forall (v_i, v_j) \in E_D \tag{7}$$

where $C_{i,j}$ is the energy cost spent by the sensors v_i and v_j when the sensor v_i executes a test over the sensor v_j. The cost of a test executed by sensor v_i on sensor v_j at a distance d from each other is computed as the sum of the energy spent by v_j to send the output sequence of a self-test to v_i and by v_i to receive such outcome.

The nonexistence of reciprocal tests is provided by the fact that a sensor v_i tests its t next sensors and is tested by its t previous sensors in the overlay network. This is true for all sensors, which, in a cyclic manner, avoid reciprocal tests. Figure 1(a) shows an example of the testing graph represented on the overlay network. Figure 1(b) has an example of the same tests over the sensor network graph. Please note that the two graphs are isomorphic in relation to each other.



(a) Testing graph generated by strategy ODTA, represented over the overlay network, with $n = 5$ and $t = 2$.

(b) Testing graph generated by strategy ODTA, represented over the sensor network, with $n = 5$ and $t = 2$.

Figure 1. Different representations of a testing graph

Although the assignment of identifiers have no relation to the geographical position of each sensor in V_D, the strategy ODTA ensures that, given an initial set V_D, the test assignment with the lowest total energy cost for the set V_D is generated.

Theorem 1. *Given a set V_D of sensors with cardinality $2t + 1$, the ODTA strategy generates a testing graph D whose $C_T(D)$ is minimum for the given V_D.*

Proof. As far as we assume that by enlarging the transmission ranges of the sensors we can obtain any edge E_D that is necessary for a test in V_D, the assignment of the identifiers of the $2t + 1$ sensors of V_D may be totally randomic and no matter which is the identifier of each sensor, the total cost $C_T(D)$ will always be the same because in PMC's optimal designs each sensor v_i has a relation to each one of the other $2t$ sensors either as a tested unit or as a tester unit. If another set of identifiers or another set of relations between the sensors is chosen (provided conditions (c1) and (c2) are guaranteed) the cost is the same, once $C_{i,j} = C_{j,i}$. □

Given that the present algorithm generates the test assignment with lowest total cost to V_D, the choice of sensors that form the set V_D is responsible for lowering the limit of the total energy consumption used in the diagnostic process.

The high number of possible combinations for the choice of the $t + 1$ sensors not belonging to T in order to ensure that $C_T(D)$ is minimal, suggests that the problem is NP-complete. This proof is not shown here, and is considered to be a future work.

Thus, an heuristic for the choice of the sensors is used. The heuristic ensures the choice of a set of sensors where most of them are next to each other. The heuristic is based on the definition of R, the smallest rectangular area that includes all sensors in T. The center of R, R_c, is used as a parameter for the choice of the sensors. The chosen sensors are the $t + 1$ sensors not belonging to T which are geographically closest to the point R_c.

Since the set V_D has $2t + 1$ sensors, the heuristic ensures that, at least, most of the sensors ($t + 1$) are the geographically closest to each other. Thus, the cost of most of the edges, or tests, in D is also minimal. However, it is not possible to ensure that the total cost is minimal.

The best case occurs when sensors in T are all close to each other. For the best case, the heuristic will present a set V_D of sensors that are close to R_c and also the closest to the sensors of T, thus minimizing the costs.

The worst case occurs when sensors of T are far from each other. For the worst case, the heuristic will ensure that most of the sensors of V_D will be close to each other. Considering real applications and fault-free sensors, it is unlikely that the sensors in T are located far from each other, once a monitored phenomenon tends to occur in a neighborhood of sensors. Thus, the worst case is not expected in most real applications.

6. Evaluation

The results presented in this paper were obtained through simulation. The simulator, programmed in C++, generates a geographic deployment of sensors based on a probabilistic distribution. The cartesian coordinates of the sensors are used to calculate the euclidian distances between them. The costs of the connection assignments are proportional to the euclidian distances between nodes. The testing strategies TAWR [Weber et al. 2010], EETA [Weber et al. 2011] and ODTA are implemented.

TAWR stands for *Testing Assignment Without Reciprocal* tests and is our preliminary work, in which an approach is proposed to generate a connection assignment that is t-diagnosable. In that work we were not concerned with energy awareness, so the strategy involved a great number of tests.

The EETA strategy (*Energy-Efficient Testing Assingment*) was proposed in [Weber et al. 2011] as our first energy-aware approach. It was built over our previous work and used the method of dividing the rectangular region that comprises the sensors belonging to T into quadrants. The center R_c of the region is used in order to choose the nearest sensors not belonging to T that take part on the diagnosis. The testing procedure ensures that sensors in one quadrant tests sensors in other quadrant in a cyclic manner so that reciprocal tests are avoided. Furthermore, an heuristic is used to exchange sensors with higher test costs for sensor with smaller costs. The number of the sensors used in the diagnosis procedure is fixed in $4t$, where t is the diagnosability of the system.

The ODTA strategy improves the gains obtained by EETA by diminishing further the number of sensors used for the establishmnet of a connection assignment.

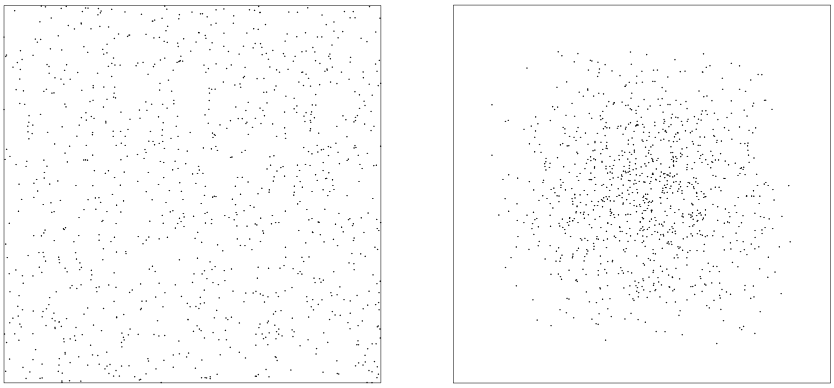
6.1. Simulation Environment

For the generation of the networks, the simulator receives a set of parameters, among which there are: (1) the network size, which defines the size (in meters) of the sensing field; (2) the number of sensors deployed in the sensing field; (3) the number of sensors that report alarms, i.e., t; (4) the statistical distribution used to geographically position the sensors over the sensing field; and (5) the *alarm region factor* (ARF), which defines the size of the region which sensors T are chosen from.

More specifically, the *alarm region* is defined as a square region positioned randomly on the network. Only the set of sensors positioned within the region are candidates to be chosen as sensors of T. The alarm region factor, or ARF, is a number that indicates the proportion of the network's size which define the size of the alarm region. So, with ARF equal to 0.1, a square region with 10% of the network size will be randomly defined for the selection of the sensors of T. The value of ARF has a direct effect on the spatial positioning of the sensors in T.

Two statistical distributions are supported by the simulator: the uniform distribution and the triangular distribution [Evans et al. 2000]. The uniform distribution allows the network to be composed of sensors distributed in an homogeneous way over the sensing field. In turn, the triangular distribution enables the generation of networks with a higher concentration of sensors at some specific point.

Figure 2 shows an example of the geographical positioning of 1024 sensors generated by (a) the uniform distribution and (b) the triangular distribution. Sensors are represented by dots distributed over a sensing field with $512mX512m$,



(a) Uniform distribution. (b) Triangular distribution.

Figure 2. Example of the positioning of sensors.

The experiments run aim at evaluating and comparing the testing approaches mentioned in the beginning of the Section. The properties evaluated in each experiment, for different values of t, are:

- the total energy cost of the test assignment D ($C_T(D)$);
- the average energy cost of the test assignment D ($C_A(D)$);
- the number of sensors in the testing graph;
- the impact of different values of ARF.

Different sets of simulations were performed; for each one, the four properties listed above were evaluated. In each simulation, 100 different networks were generated and the average values of the properties were obtained.

In all simulations, networks of size $512mX512m$, composed by 512 and 1024 sensors were considered. Simulations were run using the following values of t: 1, 3, 5, 8, 10, 12 and 15, and with the following values for ARF: 0.01, 0.1, 0.5 and 1.

Table 1 lists the set of simulations parameters.

Parameter	Value
t	1,3,5,8,10,12,15
ARF	0,01, 0,1, 0,5, 1
Distribution	Uniform, Triangular
Network Size	$512mX512m$
Number of sensors	512, 1024
Number of generated network by experiment	100

Table 1. Simulations parameters.

6.2. Results

This section presents the results obtained. At first, results for networks generated with uniform and triangular distributions are presented and compared. Then, the behavior exhibited by the testing strategies for tests with different values of ARF are described.

6.2.1. Experiments with Uniform and Triangular Distributions

In this subsection, two sets of experiments were carried out, one for a sensing field generated using the uniform distribution and the other for a sensing field generated using the triangular distribution.

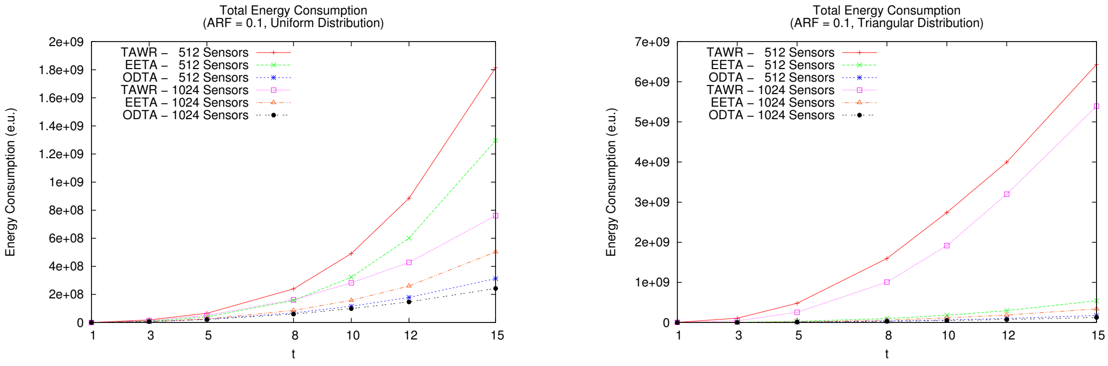
The sensing fields have the size of $512mX512m$. The ARF value is set to 0.1; so in these cases the sensors of T are chosen from a square region with $51.2mX51.2m$. Simulations with 512 and 1024 sensors were performed.

Total Energy Cost Figure 3(a) shows the average of the total energy cost ($C_T(D)$) obtained by the three strategies for networks with 512 and 1024 sensors and different values of t. The results show that, in all strategies, the total energy consumption increases proportionally to the value of t for both 512 and 1024 sensors. Furthermore, all strategies have their total energy costs reduced in denser networks because of the proximity between the sensors. As expected, the strategy ODTA, which uses the minimum number of sensors for diagnosis, shows the smallest values for $C_T(D)$.

In the experiments with triangular distribution, as shown in Figure 3(b), both EETA and ODTA have reductions in total consumption, compared with results obtained with uniform distribution due to the high concentration of sensors in one area. Furthermore, ODTA presents the smallest values of energy consumption.

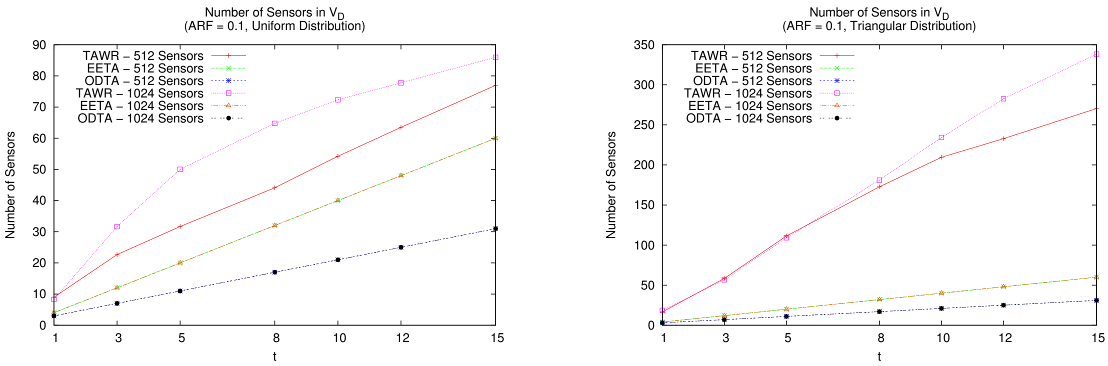
The strategy TAWR presents higher values than the other strategies due to the high concentration of sensors in a point of the network, which causes the selection of a large number of sensors for the region R in that strategy.

Number of Sensors Used Figure 4 shows the average number of sensors used by the strategies. As shown in [Weber et al. 2011], the number of sensors used for EETA is $4t$ while for ODTA it is $2t + 1$ for either distribution. Clearly, the strategy TAWR uses a larger number of sensors in both scenarios.



(a) Uniform distribution. (b) Triangular distribution.

Figure 3. Total energy consumption for each strategy, in networks with 512 and 1024 sensors.



(a) Uniform distribution. (b) Triangular distribution.

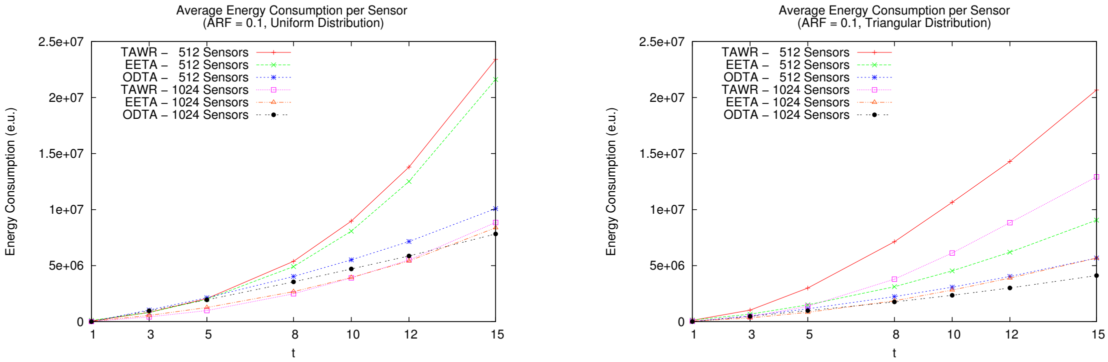
Figure 4. Number of sensors used for each strategy, in networks with 512 and 1024 sensors.

Average Energy Cost per Sensor Figure 5 shows the average energy costs ($C_A(D)$) obtained by the three approaches. For experiments using the uniform distribution, and 512 sensors, TAWR and EETA have higher values due to the smaller number of sensors in the network and, thus, the higher distance between them. On the other hand, for denser networks, TAWR uses more sensors. Thus, despite getting a total cost higher, it obtains, in some cases, an average consumption smaller than the obtained by the other strategies.

For ODTA, the lower number of sensors used, the possible farther position of the sensors in T and the heuristic for the choice of sensors makes the average energy costs per sensor higher in some cases. For experiments using the triangular distribution, the high density favors the strategy ODTA that uses very few tests and sensors close to each other.

6.2.2. Experiments Varying the Alarm Region Factor

In this subsection we present a comparison of the behavior of the testing approaches for different values of ARF. The alarm region factor influences the distance between the sensors of T. For all experiments shown below simulations with 1024 sensors and uniform distribution were carried out.



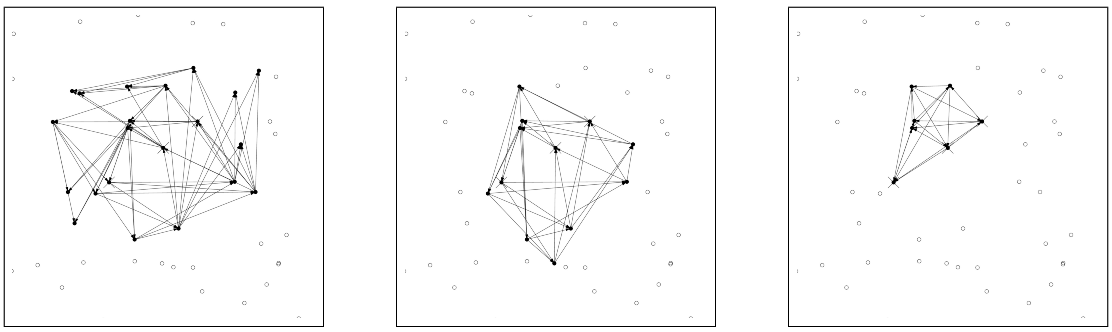
(a) Uniform distribution. (b) Triangular distribution.

Figure 5. Average energy consumption for each strategy, in networks with 512 and 1024 sensors.

Total Energy Cost Not surprisingly, in all strategies, the total consumption grows with increasing values of ARF. For EETA and ODTA, higher ARF causes the use of "longer" tests (of higher cost). For $ARF = 1$, EETA has lower values than ODTA, despite using a larger number of sensors. For ODTA, with sensors of T far from each other, tests with higher energy costs are realized. TAWR's energy consumption is much higher than that of EETA and ODTA. The graphic is not shown here due to lack of space.

Number of Sensors Used Strategies EETA and ODTA use fixed numbers of sensors, $4t$ and $2t + 1$, respectively. In strategy TAWR, for higher sizes of the region R, more sensor are added to V_D, leading to a high number of sensors used.

Figure 6 compares the final testing graph generated by each strategy for the same case, where $t = 3$. In the Figure, the sensors are represented by dots and tests are represented by edges. It is clear that the strategy ODTA presents a significant reduction on the number of used sensors,



(a) TAWR (b) EETA (c) ODTA

Figure 6. Comparison of the testing graph generated by each strategy for the same case, with $t = 3$.

7. Conclusion

In this work we were concerned with the problem of defining an energy optimal testing assignment for the diagnosis of a WSN. The testing assignment proposed was based on the concept of optimal designs of the PMC model, more specifically, on the so called set $D_{\delta t}$ of designs. The *Optimal Design Testing Assignment* (ODTA) approach always uses the optimal number of $2t + 1$ sensors for diagnosis, where t is the diagnosability of the system.

The strategy proposed was compared to two previous approaches, the *Testing Assignment Without Reciprocal* tests (TAWR) and the *Energy-Efficient Testing Assignment* (EETA). The strategy ODTA presented the smallest energy costs for the great majority of the cases. Furthermore, the smallest number of sensors used increases the network lifetime. Nevertheless, EETA presents lower energy consumption in cases where the sensing region being monitored is very large. In these cases, a set with more low-cost tests (EETA) generates lower total costs than few tests with high energy consumption (ODTA). Although EETA has lower consumptions in these cases, ODTA presents better performance in general, once the occurrence of alarms in sensors far apart is not expected in most real applications. Future work includes investigating the fairness of the approach in relation to the sensors used in the diagnosis process and studying the efficiency or intractability of the problem of the choice of the sensors that take part in V_D. The comparison of ODTA to other approaches may also be cited as future work.

References

Baronti, P., Pillai, P., Chook, V. W. C., Chessa, S., Gotta, A., and Hu, Y. F. (2007). Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards. *Computer Communications*, 30(2):1655–1695.

Caruso, A., Chessa, S., and Maestrini, P. (2007). Worst-case diagnosis completeness in regular graphs under the pmc model. *IEEE Transactions on Computers*, 56(7):917–924.

Chen, G., Hanson, S., Blaauw, D., and Sylvester, D. (2010). Circuit design advances for wireless sensing applications. In *Proceedings of the IEEE*, volume 98, pages 1808–1827. IEEE.

Chessa, S. and Santi, P. (2002). Crash faults identification in wireless sensor networks. *Computer Communications*, 25(14):1273–1282.

Choi, J.-Y., Yim, S.-J., Huh, Y. J., and Choi, Y. H. (2009). A distributed adaptive scheme for detecting faults in wireless sensor networks. *WSEAS Transactions on Communications*, 8(2):269–278.

Corke, P., Wark, T., Jurdak, R., Hu, W., Valencia, P., and Moore, D. (2010). Environmental wireless sensor networks. In *Proceedings of the IEEE*, volume 98, pages 1903–1917. IEEE.

Dahbura, A. T. and Masson, G. M. (1984). An $0(n^{2.5})$ fault identification algorithm for diagnosable systems. *IEEE Transactions on Computers*, 100(6):486–492.

Evans, M., Hastings, N., and Peacock, B. (2000). *Statistical Distributions*, chapter 40 - Triangular distribution, pages 187–188. New York: Wiley, 3^{rd} edition.

Hughes, D., Ueyama, J., and et al. (2011). A middleware platform to support river monitoring using wireless sensor networks. *Journal of the Brazilian Computer Society*, 17(2):85–102.

Ko, P., Lu, C., Srivastava, M. B., Stankovic, J. A., Terzis, A., and Welsh, M. (2010a). Wireless sensor networks for healthcare. In *Proceedings of the IEEE*, volume 98, pages 1947–1960. IEEE.

Ko, T., Hyman, J., and et al. (2010b). Embedded imagers: Detecting, localizing, and recognizing objects and events in natural habitats. In *Proceedings of the IEEE*, volume 98, pages 1934–1946. IEEE.

Patwari, N., Hero III, A. O., Perkins, M., Correal, N. S., and O'dea, R. (2003). Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*, 51(8):2137–2148.

Penrose, M. D. (1999). On k-connectivity for a geometric random graph. *Random structures and Algorithms*, 15(2):145–164.

Preparata, F. P., Metze, G., and Chien, R. T. (1967). On the connection assignment problem of diagnosable systems. *IEEE Transactions on Computers Electronic*, 16(12):848–854.

Rappaport, T. S. (2001). *Wireless Communications: Principles and Practice*. Prentice Hall, 2^{nd} edition.

Santi, P. (2005). The critical transmitting range for connectivity in mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 4(3):310–317.

Santi, P. and Chessa, S. (2001). Comparison-based system level fault diagnosis in ad hoc networks. In *Proceedings of the 20th IEEE Symposium on Reliable Distributed Systems*, SRDS '01, pages 257–266.

Shuman, D. I., Nayyar, A., and et al. (2010). Measurement scheduling for soil moisture sensing: From physical models to optimal control. In *Proceedings of the IEEE*, volume 98, pages 1918–1933. IEEE.

Taghikhaki, Z. and Sharifi, M. (2008). A trust-based distributed data fault detection algorithm for wireless sensor networks. In *11th International Conference on Computer and Information Technology*, ICCIT 2008, pages 1–6.

Venkataraman, G., Emmanuel, S., and Thambipillai, S. (2008). Energy-efficient cluster-based scheme for failure management in sensor networks. *IET Communications*, 2(4):528–537.

Weber, A., Kutzke, A. R., and Chessa, S. (2010). Diagnosability evaluation for a system-level diagnosis algorithm for wireless sensor networks. In *IEEE Symposium on Computers and Communications*, ISCC '10, pages 241–244.

Weber, A., Kutzke, A. R., and Chessa, S. (2011). Energy-aware test connection assignment for the diagnosis of a wireless sensor network. In *Proeedings of 5th Latin American Symposium on Dependable Computing*, LADC '11. A ser publicado.

Xue, F. and Kumar, P. R. (2004). The number of neighbors needed for connectivity of wireless networks. *Wireless Networks*, 10(2):169–181.

Zhang, J., Jing, B., and Sun, Y. (2008). Fault locating arithmetic for multi-source network cluster nodes based on comparison. In *7th World Congress on Intelligent Control and Automation*, WCICA 2008, pages 7567 –7571.