Information Technology and Quantitative Management, ITQM 2013


Carlos Velásquez-Villada, Yezid Donoso

Abstract

This paper proposes an optimization model for network management in multihop Wireless Sensor Networks (WSNs). Here, we develop a distributed, braided multipath algorithm to deliver the information from the information sources (targets) to the Base stations (sinks) giving the network the ability to adapt to changes or failures. The Base Stations are robust nodes with capabilities for positioning themselves and communicating outside the network, which grants them the benefit of knowing other Base Stations’ position in the area of interest. Targets are nodes that generate information and need a fixed amount of bandwidth to convey this information to a Base Station. Every element in the network is static. This paper shows the mathematical optimization model for the network’s energy minimization and resilience maximization. To increase network’s resilience, the devices within the network will try to create multiple paths from the beginning trying to reach at least one Base Station. The model is solved through a heuristic algorithm based on the nearest neighbor and minimum hop concepts and it is implemented in Java. Results show that the proposed algorithm finds multiple braided paths for most of the instances proposed and some backup nodes for nodes in the paths that can help the network to extend its lifetime operation and adapt to failures.

Keywords: WSN; optimization; multipath; metaheuristics; resilience

1. Introduction

Wireless Sensor Networks (WSNs) are small electronic devices data networks where each node has sensing and communicating capabilities. These networks are commonly used to retrieve information from natural phenomena (i.e., the targets). The issue addressed in this paper is how to provide network management and resilience in an energy efficient way for WSNs. Due to constant topology changes by energy depletion of nodes or hardware failures, paths can be destroyed and targets can get lost. The network should have the capability to adapt to changes in its environment, avoid losing targets and find new paths to carry the information from the targets to the cluster head or the base station. There are a lot of WSNs applications like target monitoring for civilian and military uses, map building, environmental sensing, health care monitoring and road traffic information systems. New applications are being developed like smarter cities based on ubiquitous sensing devices [1] and health monitoring inside a person’s body using nanosensors [2].

Research in WSNs has been carried out since the 70’s [3]. WSNs have strong energy constraints. The electronic improvements in the last decade have allowed the emergence of significant WSN applications. There has been a great interest in WSN routing [4], [5], with protocols like SPIN [6], Directed Diffusion [7], Gradient
Based Routing [8], Destination Sequenced Distance Vector [9], Optimized Link State Routing [10], Ad-Hoc On-Demand Distance Vector [11], among many others; and WSN resilience through multipath routing can be found in [12, 13, 14, 15]. There is also a previous work from Montoya, Velásquez and Donoso [16] that describes the target tracking problem with the energy efficiency goal in a formal mathematical model. This last work serves as the basis for the model presented in this one.

The paper is organized as follows. Section 2 presents the WSN’s model formal definition. Section 3 details the algorithms used for the initial path establishment and path reconfiguration. Section 4 summarizes the simulation setup and conditions to test the algorithms. Finally, Section 5 shows and discusses simulation results. Conclusions are written at the end.

2. Optimization Problem

WSNs are dynamic networks that need to be reorganized due to connections changes from target movements or node/link failures. In this work, the problem to solve is how can the network’s owner guarantee continuous connectivity and operation even when nodes are energy depleted or failures happen. The network model presented in this paper tries to increase (maximize) network resilience while at the same time it tries to reduce (minimize) the network management energy consumption. The model makes the following assumptions: Every node can act as a sensor node and as a communication relay node for multihop communications. There are special nodes called base stations (BS); BSs are nodes that have a significant amount of energy (more than three times the energy of a node), and know their position and the position of other BSs in the network. There is at least one target \( t \) in the area of interest. These targets have unknown behavior; the model does not take into account the knowledge it could gather from observations from the targets. Although the target behavior learning could be an advantage for the network managing; it is assumed that the nodes do not have positioning capabilities neither the resources to learn target behavior.

The communication model as well as the sensing model follows the well-known ideal disc model. The cost of sensing or transmitting information will be kept constant regardless of the distance between the elements communicating as long as the two elements are within the reach given from the disc model. The communication radius as well as the sensing radius will be parameters for the model. It is assumed that if two nodes can communicate they will incur in a constant cost for every bit transmitted, this cost will be a parameter for the model and it depends on the technology used for transmission, the transmitting node will incur in an \( \text{energy}_{tx} \) energy consumption (for broadcast transmissions, since they are in a wireless shared medium), while the receiving node will incur in a \( \text{energy}_{rx} \) energy consumption. It is also assumed that the working area is flat and without obstacles. And, every target needs certain amount of bandwidth to transmit its information to a BS; this bandwidth will be kept constant in the model and it will be a parameter for the model.

Table 1 presents the mathematical notation used for the mathematical model and Table 2 shows all the constant parameters for the model. The model starts with a network represented by a graph \( G = (N, E) \), where \( N \) is the nodes set and \( E \) is the edges set. The nodes set is composed by the targets, the sensors nodes and the BSs. The edges set is composed by the connections within all these elements. The graph \( G \) is converted in a new graph \( G' = (N', E') \) when something in the network changes. This process is performed iteratively to make decisions on path creation or maintenance as the network changes due to failures. \( P_t \) is the set of paths for every target \( t \).

<table>
<thead>
<tr>
<th>Set notation</th>
<th>Set definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>Nodes set, indexed by ( i, j )</td>
</tr>
<tr>
<td>( T )</td>
<td>Targets set, indexed by ( t, T \subset N )</td>
</tr>
<tr>
<td>( S )</td>
<td>Sensors set, indexed by ( i, j, S \subset N )</td>
</tr>
<tr>
<td>( B )</td>
<td>Base Stations set, indexed by ( b, B \subset N )</td>
</tr>
<tr>
<td>( E )</td>
<td>Edges set, indexed by ( (i, j), i, j \in N )</td>
</tr>
<tr>
<td>( P_t )</td>
<td>Paths set of element ( t ), indexed by ( p, t \in T )</td>
</tr>
</tbody>
</table>
Parameters shown in Table 2 are assumed to be constant for a given scenario. Initial energy level for every element (node and BS) is given by $bati$. The energy is measured in mAh. Link’s capacity is given by $Cap_{ij}$. It is assumed that the capacity is constant and it is the same for all links. $dist_{ij}$ is the distance between two elements $(i, j)$ in the network. $avail_{ij}$ is the availability of the edge $(i, j)$; it is supposed that all links have the same availability for the simulation. The information flow from the phenomenon of interest (target) is modeled through $ft$. The parameter $RC$ sets the radius limit up to which a link can be created. $RS$ parameter sets the radius limit for the sensing link. Finally, $PL$ is the maximum number of multipaths that can be established for every target.

The objective functions showed in equations (1) and (2) will minimize energy consumption in the network and maximize network availability. The network resilience will be achieved through the objective function (2) and enforced by constraints (5) and (6), which establish the limits for the multipaths. Decision variable $x_{ij}^{tp}$ is a binary variable (12) that states if a connection is used in the network. It will be 1 if the edge $(i, j)$ is in the path $p$ for target $t$ to a base station, and 0 otherwise.

$$\min \sum_{t \in T} \sum_{p \in P_t} \sum_{(i,j) \in E} (energy_{tx} + energy_{rx}) \times f_t \times x_{ij}^{tp}$$

(1)

$$\max 1 - \prod_{p \in P_t} \left( 1 - \prod_{t \in T, (i,j) \in E} avail_{ij} x_{ij}^{tp} \right)$$

(2)

Construction constraints

$$dist_{ij} x_{ij}^{tp} \leq RC \quad \forall (i, j) \in E, t \in T, p \in P_t$$

(3)

$$dist_{it} x_{it}^{tp} \leq RS \quad \forall t \in T, p \in P_t, i \in S$$

(4)

$$\sum_{p \in P_t, j \in S} x_{ij}^{tp} > 1 \quad \forall i = t, t \in T$$

(5)

$$\sum_{p \in P_t} x_{ij}^{tp} \leq PL \quad \forall (i, j) \in E, t \in T$$

(6)

Energy consumption constraints

$$\sum_{t \in T} \sum_{p \in P_t} energy_{tx} \times f_t \times x_{ij}^{tp} \leq bat_i \quad \forall i, j \in S$$

(7)

$$\sum_{t \in T} \sum_{p \in P_t} energy_{rx} \times f_t \times x_{ij}^{tp} \leq bat_j \quad \forall i, j \in S$$

(8)
Data flow and connectivity constraints

\[ \sum_{t \in T} \sum_{p \in P_t} f_{i} x_{ij}^{pt} \leq \text{Cap}_{ij} \quad \forall (i, j) \in E \]  

(9)

\[ \sum_{p \in P_t} \sum_{j \in S} f_{i} x_{ij}^{pt} = \sum_{p \in P_t} \sum_{j \in S} f_{j} x_{ji}^{pt} \quad \forall t \in T \]  

(10)

\[ \sum_{t \in T} \sum_{p \in P_t} \sum_{j \in S} x_{ij}^{pt} - \sum_{t \in T} \sum_{p \in P_t} \sum_{j \in S} x_{ji}^{pt} = 0 \quad \forall i \in S \]  

(11)

Integrality constraints

\[ x_{ij}^{pt} \in \{0, 1\} \quad \forall (i, j) \in E, p \in P_t, t \in T \]  

(12)

Construction constraints (3)-(6) determine the configuration of the network and the limits for the number of multipaths to be created. Energy consumption constraints (7) and (8) assure that the system does not use more energy than the one remaining in the nodes. Dataflow constraints (9)-(11) guarantee that the information flow from the targets is kept below the channels capacity and that it gets to a base station.

The BSs are positioned near the working area borders. BSs’ positions can be established from the beginning and it is assumed they should be near the borders since from there their coverage can be increased, and their exposition to threats can be minimized.

3. Proposed Algorithm

To solve the optimization problems stated in Section 2 several heuristic algorithms are used. Table 3 shows the nodes algorithm. It assumes that each node makes decisions with local information. Nodes do not know their own positions or other nodes’ position, they can know who their neighbors are and determine if they are in a path from a target to a Base Station. They also can keep a notion of time that does not need to be very precise. Table 4 shows the base stations algorithm. It assumes that BSs know their positions, can communicate and organize in a centralized way.

Every node is simulated by an independent agent that implements the Algorithm 1. The nodes initially wait for a target detection signal to start building a path. If a node detects a target, it sends its ID in a packet stating that it had acquired a target. Then, the first node to receive this packet will reply with its own ID stating that he just added the previous node as part of his first path. After that, this new node will broadcast its ID adding one hop to the target and another node will answer and replicate the message, creating the path from the target to the BS. When several paths get to BSs, the BSs will give them priority numbers and they will choose the path with the minimum number of hops to the target. The nodes who are part of the minimum hop path will receive a packet from the BS telling them about their new status and asking them to find a backup node between their neighbors; the backup node must be connected to the previous and next node in the path. All other nodes will enter in an energy saving mode, activating from time to time all functionalities to check for changes in the network and maybe reconfiguration packets.

When a Base Station receives a packet with the route to a target, it will check its own route to this target (initially infinite or some very large value) and will update its path if the received one has less hops than the stored one. If two or more Base Stations have received packets from a target, the centralized server will inform them about this and will choose the minimum hop path as the active path, all other paths will enter an energy saving mode and will be available as backup. When a path fails, the Base Station will try to activate some of the backup nodes, if this does not work, another Base Station will activate its path to the target. See Algorithm 2.
node start (initialization)
send ID
listen for neighbors’ IDs
if IDs received then
  | save them (IDs) in neighbors list
else
  | go OFF (no neighbors, disconnected node)
end
wait until a packet comes from a BS or a target
if packet received then
  | if packet from BS then
  |   | reinforce path
  |   | start looking for backup nodes
  |   | send packet to next node
  | else if packet from target then
  |   | add one hop to hop count
  |   | send message to next node on path
  | else if k time without packets then
  |   | send packet warning status change
  |   | enter energy-saving mode (cycling ACTIVE & OFF states)
else if 5k time without packets then
  | enter OFF mode
else
  | enter OFF mode // no packet received
end

Algorithm 1: Distributed algorithm for every node in the network

BS start (initialization)
calculate position
communicate position
enter active mode
if neighbor’s IDs received then
  | send BS ID to neighbor nodes
if packet from target received then
  | if hop count received is less than hop count to target then
  |   | update path to target to path received
  | else if packet is information packet then
  |   | send information to server
end
if target has path to several BS then
  | Multipaths completed
  | stop path searching for this target
if timer-alive expired for path p then
  | send keep-alive through path p
  | replace node with backup node
  | if path p still does not work then
  |   | switch off path and switch on path on other BS
end
end

Algorithm 2: Centralized Algorithm for every Base Station in the network
4. Simulation Setup

The algorithms described in the previous section were implemented and tested in Java using an Intel Core i7-3610QM CPU @2.3 GHz with 8GB of RAM in a 64-bits Windows 8 Pro. Parameters for simulation were set as described in Table 3. 10,000 simulations were performed. Initial node positions were randomly chosen. From this 10,000 simulations, the average number of multipaths found, the average number of hops per multipath, the average number of nodes in multipath with backup nodes and the average number of backup nodes per node in the multipath were computed.

Table 3. Parameter values for simulation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td>100 m × 100 m</td>
<td>RC</td>
<td>15 m</td>
</tr>
<tr>
<td>Sensor number</td>
<td>100</td>
<td>RS</td>
<td>3 m</td>
</tr>
<tr>
<td>BS number</td>
<td>3</td>
<td>energy_{sensing}</td>
<td>1 mAh</td>
</tr>
<tr>
<td>Target number</td>
<td>1</td>
<td>initial − energy</td>
<td>10000 mAh</td>
</tr>
<tr>
<td>PL</td>
<td>3</td>
<td>availability</td>
<td>0.999</td>
</tr>
<tr>
<td>energy_{tx}</td>
<td>15 mAh</td>
<td>energy_{rx}</td>
<td>14 mAh</td>
</tr>
</tbody>
</table>

5. Analysis and Simulation Results

An initial simulation scenario is shown in figure 1a. Connections found for this scenario are shown in figure 1b. Paths from nodes to nearest Base Station are in figure 1c. Final paths to provide resilience are in figure 1d. 10,000 of these scenarios were generated and solved. The information computed from these simulations is shown in Table 4. Figure 1c shows the path to every node from the Base Stations, although this algorithm was tested, we take the decision of leaving out this part since it consumed a lot of energy and the clusters would make harder the creation of multiple paths from the targets.

Table 4. Simulation results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulations</td>
<td>10,000</td>
</tr>
<tr>
<td>Initial Energy</td>
<td>10,000 mAh x 100 nodes = 1,000,000 mAh (100%)</td>
</tr>
<tr>
<td>Energy after neighbor discovery</td>
<td>990.862 mAh (99.09 %)</td>
</tr>
<tr>
<td>Energy after target sensing</td>
<td>990.606 mAh (99.06 %)</td>
</tr>
<tr>
<td>Energy after multipath creation</td>
<td>733.322 mAh (73.33 %)</td>
</tr>
<tr>
<td>Energy after backup nodes lookup</td>
<td>730.072 mAh (73.01 %)</td>
</tr>
<tr>
<td>Multipaths found</td>
<td>3</td>
</tr>
<tr>
<td>Multipath average length</td>
<td>6.33</td>
</tr>
<tr>
<td>Average nodes with backup per multipath</td>
<td>1.33</td>
</tr>
<tr>
<td>Average backup nodes per node with backup per multipath</td>
<td>2.25</td>
</tr>
</tbody>
</table>

Figure 1d shows the final multipaths found and some backup nodes for the paths (the green links indicate connections to backup nodes). For this scenario, the algorithm found three multipaths, two of them are braided paths and the other one is a disjoint one. This hybrid scheme provides the highest resilience for the network, since the most vulnerable point is the sensing node (the node with the target), if we have two or more nodes covering a target, the dependence on one node can be eliminated, making the network more resilient and improving the lifetime of the network. Sadly in most cases there are not two or more nodes covering a target. When the node covering the target depletes its energy, the network is no longer useful.
Fig. 1. (a) Initial network deployed; (b) Network connection to neighbors; (c) Clusters from Base Stations; (d) Multipaths found.

6. Conclusions

A mathematical optimization model for resilient wireless sensor networks through multipath routing has been presented. The algorithm to solve it in a distributed and energy efficient way was described, implemented and tested. Results indicate that it is possible to create multiple paths to different Base Stations to provide maximum resilience to the network in the presence of failures. They also indicate that the network uses up to 27% of its energy in the initial setup, configuration, routes and backups establishment. Although this value can be high it depends on the energetic cost for transmitting and receiving packets, since this is the most common and energy consuming operation in the network and when packets are lost, they need to be retransmitted, causing the system to use more energy.

In all scenarios where a target was covered by at least one node, the algorithm found several paths, most of them where braided multipaths, but in some cases when several nodes covered the target the algorithm could find disjoint multipaths. Since the network is static, the algorithm does not implement keep-alive packets for the backup routes and nodes, but leaves them to save energy through ON-OFF cycles. These cycles can save a lot of energy at the cost of taking more time to reactivate the path to the target when compared with algorithms that keep alternative paths alive. But they can take less time to reactivate and surely less energy than reactive routing where the network will have to rediscover the path.

Finally, the backup nodes are a useful measure to improve resilience in the network, but when the network is sparse like the ones tested in this work, it is possible that only one or two backup nodes can be found by path, leaving the other nodes without backup thus making the path vulnerable. In networks of higher density the backup nodes could be a more interesting implementation, since they can improve resilience at low costs (3% in average for the simulations).
Acknowledgements

The authors would like to acknowledge the support of the Colombian “Departamento Administrativo de Ciencia, Tecnología e Innovación - Colciencias Conv. 528/2011”.

References


