

Application-Driven Node Management in Multihop Wireless Sensor Networks

Flávia Delicato^{1,3}, Fabio Protti², José Ferreira de Rezende³, Luiz Rust¹, Luci Pirmez¹

Núcleo de Computação Eletrônica¹, Computer Science Department², Grupo de Teleinformática e Automação³ - Federal University of Rio de Janeiro
P.O Box 2324, Rio de Janeiro, RJ, 20001-970, Brazil
{ fdelicato, fabiop } @ nce.ufrj.br, rezende@gta.ufrj.br, { luci, rust } @ nce.ufrj.br

Abstract. A strategy for energy saving in wireless sensor networks is to manage the duty cycle of sensors, by dynamically selecting a different set of nodes to be active in every moment. We propose a strategy for node selection in multihop sensor networks that prioritizes nodes with larger residual energy and relevance for the application. The proposed scheme is based on an implementation of the knapsack algorithm and it seeks to maximize the network lifetime, while assuring the application QoS. An environmental monitoring application was simulated and huge energy savings were achieved with the proposed scheduling algorithm.

1. Introduction

WSN applications often request the deployment of sensors in hard access areas, turning battery recharge or sensor replacement so difficult that it is important to keep sensor nodes alive as long as possible. Therefore, the network operational lifetime is severely constrained by the battery capacity of its nodes. Energy saving becomes a paramount concern in WSNs, particularly for long running applications [5].

WSNs often have a large density of nodes, generating redundant data. Recent works [8,9,10] argue that, instead of providing such unnecessary redundancy to the application, the large density of nodes can be exploited to achieve significant energy savings by dynamically activating a reduced set of sensors (i.e. some nodes are assigned “to sleep”).

This work analyses the potentiality of adopting an enhanced sensor management in multihop WSNs, based on the strategy of turning off redundant sensors to extend the network lifetime while satisfying application requirements. The fundamental problem concerns the election of nodes that should remain active. Basically, the election process is formulated as an optimization problem, which is solved by the knapsack algorithm [2]. The major goal is to maximize relevance (for the application) and residual energy of active nodes, constrained by connectivity, coverage and energy issues.

Several researchers have been investigating the problem of WSN management in the last years, most of them seeking to achieve high levels of energy efficiency and considering the guarantee of coverage and connectivity as the unique QoS requirement for WSNs. In [1] techniques of linear programming are used to select the minimum set of active nodes able to maintain the sensing coverage of the network. Application specific requirements were not considered in these works. In [9] and [8] the problem of maximizing the lifetime of a WSN while guaranteeing a minimum level of quality at the application level is addressed. In those works, node selection and data routing are jointly addressed, and solved as a problem of generalized maximum flow. They present both an optimal and a heuristic solution with a totally centralized approach.

In contrast, our work addresses the active node selection as a problem independent from the network routing protocol. The proposed scheme for node selection considers as QoS requirements, besides coverage and connectivity requirements, network-related and application-related parameters, such as network lifetime and data accuracy. Furthermore, differently of approaches based on computational intensive techniques of linear programming, which are restricted to run off-line, the proposed approach is light enough to be executed inside the sensor network.

The rest of this paper is organized as follows. In Section 2 we present the problem description and formulation. Section 3 describes the performed simulations and results. Finally, Section 4 presents our conclusions.

2 Node Election in MultiHop Wireless Sensor Networks

Given an application submitting a sensing task to the WSN, the node election algorithm decides which sensors should be active for the task execution. In the proposed algorithm, time is divided in rounds. During each round r the subset of selected nodes and the role of each node (sensor/router) do not change. A task launching at the round initiation can last a time interval equal to an integer number multiple of p , where p is the round extent.

The algorithm of node election is firstly executed when interests from a new application are submitted to the network. Application interests consist of the task descriptor and QoS requirements. The first round of a task starts just after the election algorithm is concluded. The algorithm is executed again in the following cases: (i) on-demand by the application to change some QoS parameter; (ii) proactively by the network, for purposes of energy savings; or (iii) reactively by the network whenever some QoS violation is detected.

2.1 Network and Application Models

A WSN is usually composed of hundreds of sensor nodes and one or more sink nodes. Sink nodes are entry points of application requests and gathering points of sensor-collected data. The data communication in WSNs is accomplished through multiple hops from data sources to sink nodes.

The energy model assumes that sensors are capable to operate in a sleep/inactive mode or according to K predefined active modes. Two main roles are assumed by active nodes: (i) source, for nodes placed inside the target area; (ii) router, for nodes outside the target area, responsible for forwarding their neighbors data. Furthermore, a sensor can play both roles, simultaneously. In each mode, a sensor spends a different amount of energy [3].

An application of environmental monitoring (continuous measurements about a given physical phenomenon) was chosen as the target of our work. The application defines a data-sending rate, a geographical area of interest, monitoring time and, optionally, one or more data aggregation functions. Furthermore, the application defines a minimum value for the accuracy and for the spatial precision of the sensor-collected data.

2.2 Problem Formulation

The proposed scheme for node selection aims to maximize the lifetime of a network containing N multi-mode sensors while guaranteeing a required level of application quality. The adopted algorithm seeks out the best set of sensors to be activated for accomplishing a specific sensing task. Two strategies can be used to extend WSN lifetime: (i) to minimize the network energy consumption by choosing the smallest possible number of nodes capable of providing the requested level of QoS; and (ii) to maximize the residual energy of the selected nodes, that is, to consume energy in a uniform way among sensors along time, thus avoiding the premature collapse of excessively used nodes. Both strategies are used in the proposed algorithm. Further, the algorithm takes into account the potential relevance of data reported by each sensor, from the application point of view.

The proposed scheme for node selection was modeled as a knapsack problem [2], with some additional constraints. With the knapsack algorithm applied to the problem of active node selection, the sum of the utilities of nodes placed in the knapsack is optimized under the constraint of the energy budget considered. The algorithm seeks to maximize the relevance R_i and the final residual energy U_i of the selected nodes. The objective function of the problem is given below:

$$\begin{aligned} \text{Max } \sum x_i (\alpha R_i + \beta (U_i - w_i)) \\ \text{st. } \sum x_i w_i \leq M, \text{ where } x_i \in \{0,1\} \end{aligned} \quad (1)$$

A value 1 for x_i indicates that sensor i is selected to participate of task T. The term $U_i - w_i$ denotes the final energy of sensor i , if it was chosen to participate of the task (initial residual energy U_i minus the energy spent for the sensor in the task, w_i). The coefficients α and β are used to balance the priorities given for each term of the equation, and they depend on the application QoS requirements. In the general case, $\alpha = \beta = 1$.

The relevance of a node i depends on its physical and topology characteristics, given by its nominal precision (NP_i); the environmental noise of its measurements (F_i); its set of sensing neighboring nodes (N_i) and its proximity of the target area defined by the application (A_i). Each parameter contributes with a different balancing factor for the computation of R_i . The value of NP_i is a physical feature of each sensor. We assumed that NP_i have the smallest balancing factor among all terms for computing R_i . The parameter F_i

is mainly influenced by the physical characteristics of the place where i was deployed. The parameter F_i is in fact a normalized value that depends upon the actual level S_i of environmental noise, where S_i ranges from 0 to 100. We applied the formula $F_i = 1 - S_i / 100$ (2).

The largest balancing factors were assigned to the parameters A_i and N_i . The values of those two parameters are highly correlated. The value of N_i is inversely proportional to the amount of neighbors of the sensor. The importance of the value measured by a node in a location X, Y is proportional to the contribution of that sensor for sensing such location.

For calculating the value of A_i , sensors with distances d_i from the target area larger than the radio range Rr are automatically excluded from selection. Since it is desired to assign a smaller value of relevance for sensors located at larger distances from the target area, we applied the formula $A_i = 1 - d_i / Rr$ (3).

From the observed correlation between A_i and N_i , and considering the different balancing factors of each parameter in the calculation of R_i , the following equation is used:

$$R_i = \delta NP_i + \phi F_i + \gamma \left(\frac{1}{A_i N_i} \right) \quad (4)$$

where ϕ , δ and γ are coefficients that represent the balancing factors of each parameter, and $\delta < \phi < \gamma$.

2.2.1 Including QoS Profiles

Applications can choose to prioritize the lifetime in favor of the accuracy, or to prioritize the accuracy in favor of the monitoring period, or they can choose to balance both parameters. In the present work, the application QoS requirements, along with the parameter that it chooses to prioritize, compose a QoS profile. There are 3 possible QoS profiles: (i) precision-based, which prioritizes the data accuracy or precision; (ii) lifetime-based, which prioritizes the network lifetime; and (iii) ratio-based, that seeks the best tradeoff between energy consumption and data accuracy.

Considering the QoS profiles above, the original objective function (4) is modified to include different weights according to the priority given by the application to the different QoS parameters. For precision-based profiles, larger values are assigned to the coefficient α ; for lifetime-based profiles, larger values are assigned to the coefficient β ; and finally, for ratio-based profiles, equal values are assigned to both the coefficients.

2.3 Constraints

The choice of active nodes in a WSN is subject to a set of constraints, which should be taken into account by any scheme for node selection.

2.3.1 Energy Constraints

The first constraint to be considered (R1) is the finite amount of energy of the network. At each round j , the energy spent by the selected set of sensors cannot be larger than the

budget of energy of the network for that round. The constraint R1 is already taken into account by the knapsack algorithm, since the value (capacity) of the knapsack is the total budget of the network in each given round.

A second energy-related constraint (R2) considers that a sensor node is only eligible to remain active in a round j if it has energy enough to remain alive up to the end of the round. To satisfy that constraint, we defined a minimum energy threshold, L , which a node should have to be eligible for selection. For establishing such threshold we assumed that, if the node is inside the target area, it should have at least energy enough for sensing at the defined rate and to transmit its data. Otherwise, it should have at least energy to forward its neighbors' data. The constraint R2 can be defined as follows: $x_i \leq U_i/L$ (R2)

Since x_i is a binary variable, if the residual energy U_i of sensor i is smaller than the threshold L , x_i is set to 0 (the sensor cannot be selected). Otherwise, if $U_i \geq L$, then the variable x_i may or may not be set to 1 (that is, the sensor i is eligible).

The constraint R2 can be included in the knapsack algorithm, by including an additional if, or it can be solved through a previously executed procedure.

2.3.2 Coverage and Connectivity Constraints

Since the primary goal of a WSN is to monitor the environment, it has to maintain a full sensing coverage, even when it operates in a power save mode. Besides, a successful WSN operation must also provide satisfactory connectivity so that all active nodes can communicate for data fusion and report to sink nodes.

In this work, a point p is assumed to be covered by a node i if the Euclidian distance between them is smaller than the sensing range of i , denoted by Sr . Another assumption is that the covering area CA of sensor i is the circular area with center in X,Y , where X,Y are the geographical coordinates of i , and whose ray is Sr . A convex area A is defined as having a degree of coverage K (that is, A is K -covered) if every point p inside A is covered by at least K nodes [10]. In addition, we assumed that any two nodes i and j can directly communicate if the Euclidian distance between them is smaller than the radio range of the nodes, Rr , i.e., $d(i,j) < Rr$.

The coverage and connectivity constraint R3 can be formulated as follows. Given a convex area A and a coverage degree K specified by the application, the number of inactive nodes should be maximized under the constraint that (i) active nodes guarantee that A is at least K -covered and (ii) all active nodes are connected. That is, for every point p of A :

$$\sum_{i \in A} x_i \geq K \text{ (coverage degree requested by the application) (R3)}$$

To satisfy such constraint, a procedure based on the disk-covering algorithm [6] was employed before executing the knapsack algorithm. That procedure consists of two stages, the first one aiming to guarantee the coverage of the target area and the second one to guarantee the network connectivity. In the first stage, the target area (a rectangular area defined by the application) is totally covered by disks whose diameter is defined as the spatial precision requested by the application. Afterwards, the procedure heuristically selects K nodes that must remain active inside each disk. That selection can be totally

random or it can take into account the residual energy of the nodes. In the second stage, the sensor field is totally covered by disks whose ray is equal to the radio range R_c . To assure the network connectivity, the procedure should guarantee that in each disk there is at least one active node.

3. Simulations

We ran simulations in the JIST simulator [7] to demonstrate the benefits of using our scheme for node selection in WSNs. A greedy heuristic for solving the knapsack algorithm was implemented [2]. The algorithm runs in an unconstrained sink node.

An application of environmental monitoring was simulated. The requested sensing task was to monitor the temperature of a target area during a period of time. The application was interested on raw data values (with no aggregation), with the following requirements: (i) a spatial resolution of 40m^2 with a 1-coverage degree (at least 1 sensor at each 40m^2); (ii) an acquisition rate of 10 seconds, and (iii) a data accuracy above a predefined threshold. Data accuracy is given by the Mean Square Error (MSE) value. The MSE is calculated as the difference between a set of values assumed as “real” values and the set of values generated by sensors, considering their nominal precisions and the environmental noise. The WSN lifetime has to be long enough to guarantee that data will be collected during all period of time requested by the application and respecting QoS requirements.

A sensor field was created with 300 nodes randomly distributed in a square area with $200\text{m} \times 200\text{m}$. Each node had a radio range of 40m and a sensing range of 20m . The energy dissipation model is as described in [3]. Sensors that generate data (sources) were randomly selected from nodes in a $100\text{m} \times 100\text{m}$ square (target area) within the sensor field. The sink node was located in the right upper bound of the field. Since we were not interested in simulating any specific routing or MAC protocol, we assumed hypothetical protocols, delivering data generated from sources to the sink node through the shortest path (in terms of geographical distance), without data loss. Each simulation runs for 1000 seconds, divided in 10 or more rounds, at the end of each the network residual energy and the MSE are computed. “Real” values of temperature data were randomly generated at every round ranging from 20 to 40 degrees Celsius. The size of data packets in all transmissions is fixed and equal to 100 bytes. All results correspond to the average of 10 simulation runs.

In the first simulations, we compare the results of scheduling different percentages of active nodes, in terms of final residual energy of the network and data accuracy. Our goal is to show that activating only a subset of nodes can satisfy the QoS requested by the application, leaving WSN resources for new tasks and applications. The network energy budget (knapsack capacity) is specified as a percentage of active nodes, which varies from 30% to 100%. A budget given in percentage means that the knapsack capacity is set to the sum of the weights of the respective percentage of nodes. In the greedy approach, we assume that the weights of all nodes were the same and equal to their initial energy. All sensors have an initial energy randomly chosen between 15 and 20J. Before running the selection procedure, nodes were sorted according to their relevance and residual energy, so

that the procedure prioritizes the selection of nodes with larger values for these parameters. After running the procedure, routes from sources to the sink were established and kept unchanged until the end of the rounds. The monitoring time requested by the application corresponds to 9 rounds and the maximum tolerated MSE was 0.3.

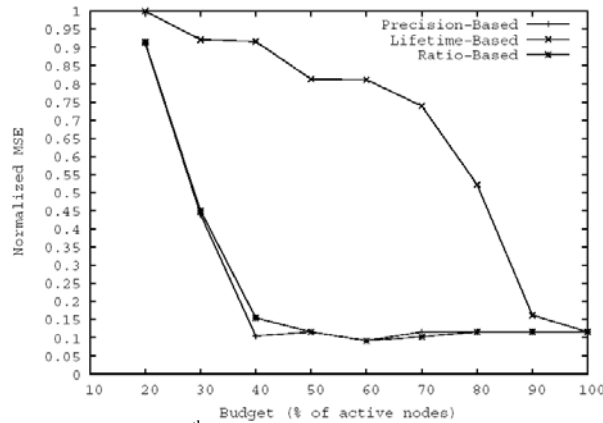


Fig. 1. Normalized MSE at the 10th round, for the different budgets, considering the three QoS profiles.

Results shown that a gain of 1000% in the final energy at round 9 is obtained when only 30% of nodes are activated, in contrast with activating 100% of nodes. We observed that from the 8th round the MSE starts increasing for all budgets. This is due to a large number of sensors being short of energy. Lifetime expiration of source nodes or nodes located in the path from sources to the sink prevents data delivery. Although the MSE increases, up to the 9th round it is still below the point tolerated by the application, for all budgets. From the next rounds, MSE increases to a value above the desired threshold, meaning that the application QoS is not being satisfied anymore. Since the monitoring time was requested as 9 rounds, results prove that with only 30% of nodes the application QoS was met, with a huge energy saving. Next, we varied the number of sensors while keeping the size of the sensor field, to analyze the effect of node density. Similar results were achieved for 400 nodes. For 200 nodes a smaller although significant energy saving of 300% was obtained. These results indicate that schemes for node scheduling are more suitable for high density WSNs.

All the previous simulations assumed a ratio-based QoS profile. Next, we evaluate the effect of using the different profiles considered in this work. For the precision-based profile the value of the coefficient α was set to 50, while β was set to 1. For the lifetime-based profile the value of the coefficient α was set to 1, while β was set to 50. Results show that the final energy does not significantly change among the different profiles. This result is due to the fact that the selection algorithm runs before the first round, when the residual energy of all nodes is very similar. A different result would probably be achieved if nodes were assigned energy values with larger ranges. On the other hand, the values of relevance

vary a lot among different nodes. Results shown in Fig.1 corroborate this fact. When the application decides to prioritize the relevance (precision-based profile) the final value of error was up to 90% smaller than when the network lifetime is prioritized.

4. Conclusions

We presented a scheme for node selection in multihop WSNs whose primary goal is maximizing residual energy and application relevance of active nodes. We formalized the problem of node selection as an optimization problem, and we adopted the knapsack algorithm for solving it. An application of monitoring environment was chosen to derive some specific requirements. We adopted a non-optimal, greedy approach for solving the knapsack problem, whose complexity is low enough to allow an online, in-network execution of the algorithm. Simulation results are very encouraging, and huge energy savings can be achieved while preserving application QoS requirements.

References

1. Chakrabarty, K. et al.: Grid coverage for surveillance and target location in distributed sensor networks. *IEEE Transactions on Computers*, 51(12), pp. 1448-1453 (2002)
2. Cormen, T. H. et al.: *Introduction to Algorithms*. MIT Press (2001)
3. Estrin, D., Sayeed, A., Srivastava, M.: *Wireless Sensor Networks*. Mobicom2002 Tutorial. Available in: <http://www.sigmobile.org/mobicom/2002/program/tutorial.html>
4. Frolik, J.: QoS Control for Random Access Wireless Sensor Networks. In Proc. of IEEE WCNC2004, Atlanta (2004)
5. Mainwaring, A. et al.: Wireless sensor network for habitat monitoring. In Proc. of WSNA2002, Atlanta (2002)
6. Pach, J., Agarwal, P.K.: *Combinatorial Geometry*. Wiley Pubs. New York (1995)
7. JIST: Java in Simulation Time. Available in <http://jist.ece.cornell.edu/>
8. Perillo, M., Heinzelman, W.: Optimal sensor management under energy and reliability constraints. In Proc. of the IEEE WCNC2003, New Orleans (2003)
9. Perillo, M., Heinzelman, W.: Sensor Management Policies to Provide Application QoS. *Elsevier AdHoc Networks Journal, Special Issue on Sensor Network Applications and Protocols*, 1 (2-3), pp 235-246 (2003)
10. Wang, X. et al.: Integrated Coverage and Connectivity Configuration in Wireless Sensor Networks. In Proc. of ACM SenSys03, Los Angeles (2003)