Optimal relay node placement for throughput enhancement in wireless sensor networks

Gianni A. Di Caro and Eduardo Feo Flushing
Dalle Molle Institute for Artificial Intelligence (IDSIA)
Lugano, Switzerland
{gianni,eduardo}@idsia.ch

***Abstract*—In this work we consider the relay node placement problem in wireless sensor networks: where to position a limited number of available nodes that can act as additional relays to foward sensor data toward one or more base stations. The objective is to define relay node locations in order to improve network performance in terms of delivery ratio and end-to-end delay, and/or to provide connectivity in partially disconnected areas. Typical application scenarios include the repair of the network in face of failures, or the case of networks used in dynamic environments, such as network characteristics need to be dynamically adapted to the changing conditions. We formalize the problem by defining a linear, mixed integer mathematical programming model. We include a number of constraints and penalty components, aimed at closely modeling the specific characteristics of the wireless environment. Model solutions specify both where to place the relays and the optimal data paths to route the data. Through a comprehensive evaluation in simulation we show that our approach is effective in accomplishing the desired objectives. We compare the solutions provided by our scheme against a state-of-the-art dynamic routing protocol, to assess the quality of the routes, and against a relay node placement heuristic, to evaluate the positioning of the relay nodes.**

I. INTRODUCTION

A *wireless sensor network* (WSN) consists of a set nodes that are equipped with sensing and limited processing capabilities, and can locally communicate with each other through a wireless medium [1]. The *sensor nodes* (SNs) composing a WSN are usually inexpensive and low-powered, such as they can be deployed in large numbers to provide monitoring and sensing services for long time periods. In typical applications, the data generated by the sensor nodes need to be transmitted to and aggregated and processed at *base stations* (BSs). The general model for the forwarding of the data from SNs to BSs is based on the definition and use of *multi-hop routing paths.*

Since a WSN can operate for relatively long times and/or it can be embedded in dynamic or hostile environments, a core issue in WSNs is the definition of effective strategies for the time maintenance of network operativity and/or for its adaptation to external or internal changes. In this direction, a wealth of research has considered the use of special nodes, referred to as *relay nodes* (RNs), that can be deployed and added to the WSN after the network has been put in place. RNs can be positioned at precise locations by hand, or they can be part of a mobile robotic unit, such that they can be deployed autonomously or on-demand. Possible roles of RNs include the provisioning of connectivity and the enhancement of the network response (e.g., in terms of throughput, sensing, and lifetime), where and when needed. For instance, RNs can be used to provide connectivity when the network becomes partially disconnected from the BSs because of factors such as energy depletion, node failures, interaction with a hostile environment, etc. In other scenarios, such as the tracking of moving targets and the monitoring of dynamic environments, the role of the RNs is to help the network to locally adapt to the external changes. In this work we consider the problem of the deployment of RNs with the objective of general performance enhancement. That is, given a restricted number of available RNs, we aim to determine the locations where these additional nodes can be positioned in order to improve throughput and end-to-end packet delays for the data gathered at BSs. When the network is partially connected (i.e., some SNs are not be able to send data to BSs), the deployment of RNs is aimed both at bringing connectivity and optimizing global network performance. We formalize the problem of RN positioning as a *linear, mixed integer mathematical program* (MIP). The formulation includes a number of constraints and penalty components, aimed at closely modeling the specific characteristics of the wireless environment, as well as a number of heuristics, aimed at speeding up the computations. The model is solved to optimality using a standard solver, finding the best locations where to position the available RNs. Although our primary objective is determining the physical locations at which RNs should be placed, the solution also specifies the way these RNs should be used. This specification comes in the form of *optimal routing paths* from SNs to BSs to forward data flows.

By considering a large set of different network topologies and traffic loads, we performed extended simulation studies to assess the efficacy of the proposed model. The TOSSIM simulator for WSNs [2] was used. The throughput and delay performance of the static data routes obtained by the solution of the MIP model is compared to the performance obtained by the *Collection Tree Protocol* (CTP) [3], a state-of-the-art adaptive routing algorithm for WSNs. In addition, the effectiveness of the found RN locations is evaluated by comparing the network performance with and without the additional RNs and, for the case of initially disconnected topologies, we compare the performance of our RN placement versus the placement defined by a state-of-the-art *minimum spanning tree* (MST) heuristic. Simulation results show that our approach outperforms the selected competitors, effectively achieving the network performance enhancement objectives and the provisioning of connectivity, when needed.

The major contribution of this work is the MIP formulation, that realistically models the core aspects of a WSN and defines both optimal RN locations and routing paths to BSs, to provide maximal network performance and connectivity.

II. RELATED WORK

In the recent years, a number of studies in WSNs have considered the *relay node placement problem* (RNP) under different requirements and objectives. Most of the existing work has focused on the deployment of RNs to provide *connectivity* [4]–[12], *extend the network lifetime* [13], [14], *energy-efficient or balanced data gathering* [15]–[17], and to provide *survivability and fault tolerance* [10], [12], [18]–[21].

Studies based on connectivity requirements aim to place a minimum number of RNs in order to achieve different goals. In the *Connected* problem, the objective is to make the induced network topology globally connected, assuming the nodes were initially partially disconnected. In the *Survivable* or *Fault-tolerant* approach, the goal is to ensure that the network remains connected in the presence of $k \geq 1$ failures. For the Connect RNP, the problem can be described as deploying a minimum number of RNs in a WSN so that, for each pair of SNs, it exists path consisting of RNs and/or SNs and such that each hop of the path is no longer than the common transmission range of the SNs and the RNs. Lin and Xue [4] formulated the problem as the *Minimum Steiner Tree* with minimum number of Steiner points and bounded edge length problem. They proved that the problem is NP-hard and proposed a minimum spanning tree 5-approximation algorithm. Following this work, several authors tackled the Connected and the Survivable RNPs by proposing different approximation algorithms. Although our problem shares some similarities with these works with respect to the provisioning of connectivity when needed, we do not focus on optimality in terms of number of RNs. Also, none of the previously mentioned works considered the RNP from a practical point of view, that is, taking into consideration network properties and/or wireless channel characteristics.

Falck *et al.* [15] have considered the RNP in the context of balanced data gathering. They presented the problem of finding an optimal routing as a linear program, but with the objective of achieving load balancing. Patel *et al.* [16] examined the joint problem of deploying SNs, RNs and BSs on a set of feasible locations and finding bandwidth-constrained energy-efficient routes with guaranteed coverage, connectivity, bandwidth, and robustness. They also make the use of a linear program formulation. These authors have considered the objective of network performance in the form of maximize network utilization in which RNs can be deployed only in a set of feasible sites. Our work follows a similar approach, however, we perform extensive network simulations to determine the quality of the solutions and include the effects of interference and radio propagation in our model. Kashyap *et al.* [22] studied the placement of RNs with the goal of reducing the maximum link load for a given traffic imposed on the sensors. Ergen and Varaiya [17] considered the problem of determining optimal locations for RNs together with optimal energy provisioning, such that the network operates for the desired lifetime with minimum energy expenditure. They have also considered a non-linear programming model and established a set of possible locations for the RNs based on a grid partitioning. However, their work mostly focuses on energy-efficiency and considers a simplistic radio propagation model. Therefore, the quality of the solutions in terms of network performance is not really studied. Finally, Wang *et al.* [14] studied the deployment of RNs to maximize network lifetime in two-tiered WSNs with a single base station. We consider general flat topologies and the presence of multiple BSs.

III. SYSTEM MODEL

We model a WSN as a set of sensor nodes (SNs) and base stations (BSs) located in a set of known positions $\mathcal{S}$ and $\mathcal{B}$, respectively (e.g., obtained using global or relative positioning devices and techniques). SNs generate data packets and also forward packets received from other nodes towards one of the BSs. We assume that the data generation characteristics of each SN are known, or can be estimated. We are given a set of K relay nodes (RNs) which can be additionally placed in the sensing field. These nodes do not generate any information, their only task is to forward data received from other nodes, therefore serving as a bridge between disconnected parts of the network and/or improving the quality of the paths used to route SN data. The nodes in the network communicate with each other within the network *communication range* $\mathbf{r}$ using a shared wireless channel. The problem we consider consists in finding the optimal locations for a maximum number K of RNs and determining routing paths from each SN to a BS in order to maximize total network throughput. The way we tackle this objective is by minimizing end-to-end delays and the percentage of packet losses due to the characteristics of the shared medium. When no RNs are available to be deployed ($K = 0$), the problem becomes a pure *routing problem*, and its solution provides the optimal data routes. In the case the network is partially connected, the placement of RNs aims both to connect the network and improve its performance. We do not require the placement strategy to deploy all the RNs available. In fact, in some cases, the maximal performance can be achieved using less than K RNs. Moreover, we do not constrain the data to be transmitted over a single path, hence a data flow can be split over multiple paths.

Given that computing the RN optimal locations by allowing them to be placed anywhere in the embedding environment is known to be NP-hard [23], we restrict the placement to a *numerable set of candidate locations* $\mathcal{R}$ placed on a *2D grid* covering the area. The grid points are uniformly separated in x and y coordinates by a distance Δ. Reducing Δ might lead to better quality solutions, at the cost of increasing the complexity and the computation time of problem solving.

We formalize the problem of RN deployment by defining a *linear, mixed integer mathematical programming* (MIP) model based on a *network flow* formulation. The model is described in two steps. First, we introduce a basic network flow formulation in which the relay node locations are selected with the objective of minimizing the length, in terms of hops, of the SN-to-BS flow paths. The minimization of path lengths is a basic approach to optimize both end-to-end delays and throughput. Later on, we revise this approach by introducing link costs that take into account also other aspects (i.e., wireless channel access) other than just the hop count and we introduce in the model several other components that closely model the characteristics of a wireless environment.

A. MIP Formulation

To formulate the relay node placement problem, we define a variant of the *Minimum Cost Flow Problem* (MCFP). Let $G = (V, E)$ be a connected digraph representing a WSN, where $V = \mathcal{N}$ is the set of nodes and the set of edges E defines the topology in terms of communication links. Let $\gamma : E \mapsto \Re$ be a link cost function, and $\tau : \mathcal{S} \mapsto \Re$ be a data generation (traffic load) function, expressed in *data per second* a SN generates. For convenience (see below), we use the concept of *flow unit* (f_{unit}), measured in bytes/sec, as unit for τ values. In the following, since $e = (i,\ j) \in E$, we write γ_{ij} as $\gamma(e)$.

As *decision variables*, the model makes use of the following. The *flow variable* f_{ij} denotes the amount of flow through link (i, j), that is, the data traffic to be sent from node n_i to node n_j located at positions $i,\ j \in \mathcal{N}$ respectively. The values of f_{ij} are expressed in *flow units*, measured in bytes/sec. The *binary positional variable* y_i indicates whether location $i \in \mathcal{R}$ (a point in the 2D grid) is being used to circulate flow or not. When y_i is set to 1 in a solution, it indicates that a RN should be positioned at the corresponding location. A full solution to the relay placement problem is specified by the two sets of variables f and y: the locations of the relays to be deployed are determined by the set $\{i \in \mathcal{R} \mid y_i = 1\}$, the SN-to-BS routes are defined in the *routing-tree* induced by the set $\{(i,\ j) \in E \mid f_{ij} > 0\}$. The minimum cost flow problem with RN placement is formulated as:

$$\min \quad \text{RNP} = \sum_{(i,\ j) \in E} \gamma_{ij} f_{ij} \tag{1}$$

$$\text{subject to:} \quad \sum_{(i,\ j) \in E} f_{ij} - \sum_{(j,\ i) \in E} f_{ji} = \begin{cases} \tau_i & \text{if } i \in \mathcal{S}, \\ 0 & \text{if } i \in \mathcal{R} \end{cases} \tag{2}$$

$$\sum_{i \in \mathcal{B}} \sum_{(j,\ i) \in E} f_{ji} = \sum_{k \in S} \tau_k \tag{3}$$

$$y_i = 1 \iff \sum_{j \in \mathcal{N}} f_{ji} > 0 \quad \forall i \in \mathcal{R} \tag{4}$$

The following additional constraint limits the maximum number available K of RNs by using the variable y_i:

$$\sum_{i \in \mathcal{R}} y_i \leq K \tag{5}$$

Given that the optimal solution can be found for a number of RNs $k < K$ (e.g., using the remaining RNs can be redundant or even introduce unnecessary extra costs). Therefore, we include a penalty factor in the objective (1) to favor the use of a minimal amount of RNs: any optimal solution using n relays needs to provide a minimal gain with respect to the solution obtained using $n - 1$ relays. We term this minimum gain the *relay penalty factor* $\hat{R}$. The objective function (1) becomes:

$$\min \ \text{RNP} = \sum_{(i,\ j) \in E} \gamma_{ij} f_{ij} + \hat{R} \sum_{i \in \mathcal{R}} y_i \tag{6}$$

$\hat{R}$ is a parameter that can be adjusted according to the problem instance (e.g., relay node availability, economic cost). In the tests of Section IV $\hat{R}$ is set to one flow unit.

Shared wireless channels in WSNs are necessarily *bandwidth-limited.* Therefore, we must explicitly consider this limiting factor to avoid solutions that would saturate parts of the network producing, in practice, large packet losses and bad performance. To specify this as a constraint, we first introduce the notion of *link capacity* L_{cap}, which is the nominal amount of data that can be transmitted by a wireless link in the network (assumed that they all have the same nominal capacity). Link capacity is usually measured in bytes/sec, and, as in the case of τ, we express it in flow units. Using link capacity, the following constraints formulate bandwidth limitations:

$$\sum_{(i,\ j) \in E} f_{ij} + \sum_{(j,\ i) \in E} f_{ji} \ \leq \ L_{cap}, \quad \forall i \in \mathcal{N} \tag{7}$$

For a node n, the *routing in-degree* is the number of n's neighbors using n to relay data. Because of the shared wireless medium and contention access, this number strongly impacts on the effective capacity of node n and on network load balancing. For this reason, the following constraint adds to the model a restriction on the maximum routing in-degree allowed (b is a binary auxiliary variable, D = 6 in the experiments):

$$b_{ij} = 1 \iff f_{ij} > 0 \quad \forall i,\ j \in \mathcal{N} \tag{8}$$

$$\sum_{(j,\ i) \in E} b_{ji} \leq D \quad \forall i \in \mathcal{S} \tag{9}$$

To minimize wireless interference, a central issue in multi-hop WSNs, and produce at the same time balanced routing trees, which allow balanced energy depletion, we need to setup *minimally interfering flow paths*. We enforce the generation of such paths by introducing in the objective function a penalty component based on the *maximum local flow*, $\bar{F}_{max}$, defined as the maximum amount of flow that can circulate within a disk of radius $\mathbf{r}$ centered in an SN. The calculation of the flow circulating within the $\mathbf{r}$-disk of a sensor node i, requires to sum up the outgoing flows from all i's neighbors:

$$p_i = 1 \iff \left(\sum_{(i,\ j) \in E} \sum_{(j,\ k) \in E} f_{jk} \geq \bar{F}_{max} \right) \tag{10}$$

where p_i is a binary penalty variable that takes value 1 when the flow through the disk area around i is violating the maximum amount allowed. In order to use p to include the penalty for the maximum local flow in the objective function and give it a reasonable weight, we first derive an estimate of the order of magnitude of the objective function value RNP without any penalties, eq. (1) as follows. Given a link cost function γ, we can determine an upper bound on the *maximum link cost*, γ_{max}. A lower bound on the path length from an SN i to a BS can be defined by considering the minimum physical distance between i and the BS divided by the transmission range. Therefore, the following expression represents an estimate of the optimal solution value of RNP:

$$\hat{F} = \left(\sum_{i \in \mathcal{S}} \min_{j \in \mathcal{B}} \left(\frac{|i,\ j|}{\mathbf{r}} \right) \tau_i \gamma_{max} \right) \tag{11}$$

The penalty factor for the violation in maximum local flow can be therefore included in the objective function (1), that, together with all the other penalties become:

$$\min\ \text{RNP} = \sum_{(i,\ j) \in E} \gamma_{ij} f_{ij} + \hat{R} \sum_{i \in \mathcal{R}} y_i + \alpha \sum_{i \in \mathcal{S}} p_i \hat{F} \tag{12}$$

The parameter α weighs the penalty in the objective function. This parameter could be properly tuned by taking into consideration network characteristics such as density and topology, among others. In experiments we set $\alpha = 0.1$.

So far, *link costs* were generically defined as a mapping γ. In many related works (e,g. [24]), link costs are defined in terms of hop count in which all links have the same, unit cost. In our case, adopting such a way of proceedings, would be equivalent to minimize the number of links used by each flow, that is, to find the shortest routing trees (rooted at base stations). However, it is well understood that in WSNs indiscriminately selecting links following the min-hop rule ignores many critical aspects that strongly impact on the network performance. For instance, trying to minimize hop count often determines the selection of long distance links, that are less reliable in a wireless context, consequently increasing the risks of packet retransmissions and/or losses. Whether it is better to route over longer paths, with high quality links, or shorter paths with poor quality links, is still a fundamental question in multi-hop wireless networks [25], [26]. In order to reach a good trade-off between hop-count and link reliability, we adopted the following link cost function:

$$\gamma_{ij} = 1 + w_c c_{ij}\ , \tag{13}$$

where $c : E \mapsto [0,1]$ is a *link quality estimator* ($c_{ij} < c_{kl}$ means that link (i,j) has a better quality than (k,l)). For testing, we have used a link quality estimator which estimates the probability of a successful transmission at a link (i,j) by combining an interference model and a MAC model [27] for wireless sensor networks (see [28] for a detailed discussion). The weight w_c defines the relative importance between hop-count and link quality. Defining a good trade-off between these two aspects is in general a difficult task. On the other hand, formulating the value of γ as a linear combination of these two factors it allows to examine their relationship. In the evaluation, we perform experiments with different values of w_c and compare the resulting network routes in simulations, with the goal of empirically finding the most appropriate balance.

IV. MODEL EVALUATION

To evaluate the model efficacy in generating high quality solutions to the proposed RNP, we have considered a large number of randomly generated network instances with different characteristics and ran simulations to measure the performance of the obtained solutions. We analyze the performance of the model in terms of quality of routing paths and then assess the quality of the RN locations in cases where their deployment is guided by the need for connectivity. [1]

In order to cover a wide set of application scenarios, we considered networks embedded in area sizes from 50×50 m^2 to 75×75 m^2, the number of SNs was varied from 50 up to 225, and the number of BSs ranged from 12 to 56 to consider different node densities. The communication range of the network was set to 10 m. The networks were generated with different topological characteristics, i.e. uniform, clustered, and small world. The total number of considered network instances is 1500. Traffic load is modeled using periodic packet generations for each SN and the corresponding data generation rate ranges from 320 to 960 bytes/sec. For each instance 30 runs were executed to account for randomness. In the plots we report the *empirical cumulative distribution function* (indicated with CDF) considering the results over all instances.

To solve the MIP we used the CPLEX® solver. The solver operates under its default settings and the maximum solving time is set to 60 minutes. We perform simulations using TOSSIM [2] and consider packet delivery ratio and end-to-end delays as performance metrics. The main reason behind the choice of TOSSIM as simulation environment for our work is because it provides accurate results due to its realistic wireless channel model. TOSSIM also profits from the component based architecture of TinyOS [29] to transparently define a hardware abstraction layer that simulates the TinyOS network stack at the processor level. The log-distance path loss model is used to compute the link gain values, which are set at the start of each simulation run.

A. Routing paths and link cost metric

To evaluate the quality of the routing paths defined by our model solution, we measure in simulation the network performance (network delivery ratio and end-to-end delays) obtained using these paths and we compare it with the performance obtained by using the *Collection Tree Protocol* (CTP) [3]. CTP is a state-of-the-art routing and data collection protocol distributed with TinyOS. It performs best-effort, any-cast, multi-hop deliveries from SNs to BSs (which are referred to as *root nodes*). We have chosen CTP because: (i) it is suitable to our problem due to its data collection-based functionality, (ii) it has been used by other works either for comparison analysis [30] or as a framework for testing and evaluation [31],

---

[1] The interested reader can found in [28] additional results and studies on the computational performance of solving the model and the impact of using heuristics (use of different grid sizes and adaptive grid methods).

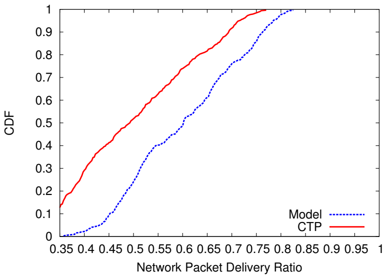

Fig. 1. Comparison between the performance obtained using the static routing paths defined by our model and CTP (Network delivery ratio).

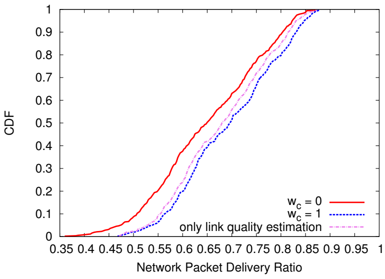

Fig. 2. Effect of different values of w_c (Network delivery ratio).

and (iii) it is included in the TinyOS suite, becoming a de-facto standard for applications based on this operating system.

By default, the forwarding engine of CTP requests packets acknowledgments and performs retransmissions when they are not received. As we consider best-effort transmissions in our work, and do not contemplate the use of packet retransmissions, we have disabled this behavior in CTP to make fair comparisons with our model. In this way, our comparison analysis will be based only in the path quality estimation of CTP and the use of probe-based routing metrics. Moreover, CTP, like many other probe-based routing protocols, requires some initialization time to stabilize the routes. Therefore, we have introduced a time lag before the start of the simulation, which has the same length as the simulation time, to allow the protocol to establish the routes, therefore in this way, we perform a fair comparison. Figure 1 shows that our model clearly outperforms CTP in terms of delivery ratio (an equivalent result was observed for end-to-end delay), in spite of the fact that the paths defined by our model are statically assigned, while CTP define them in an adaptive way.

The results shown in Figure 1 are obtained setting $w_c = 1$ (eq. 13). Figure 2 shows the effect of using different values of w_c to weight the importance of hop count vs. link quality.

B. Relay node placement to improve network performance

To evaluate the efficacy of the model on deploying RNs to enhance network performance, we have generated *connected* topologies and set the problem to deploy a maximum of K RNs. The values for K are specified in $\%$ of the total nodes (SNs and BSs) in the initial topology. Selected values for K are 5%, 10% and 15%. For each value of K, we solve the problem instances and evaluate the difference in network

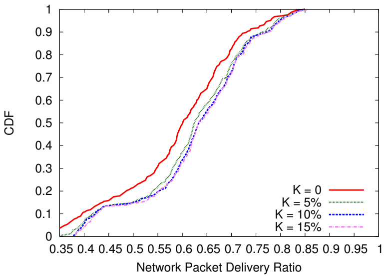

Fig. 3. Effect of increasing values of K (Network delivery ratio).

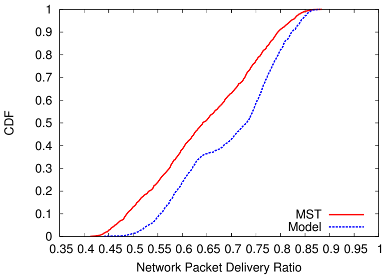

Fig. 4. Comparison between our model and the MST based heuristic to provide connectivity (area size 50×50 m^2).

performance of the solutions provided. Figure 3 shows the network delivery ratio CDF considering different values of K. The value $K = 0$ means that selected instances were solved without any relays, and serve as comparison to assess the performance improvement obtained by deploying the RNs. The performance gain obtained by the RNs deployment is quite evident. However, not all instances are prone to be significantly improved by RNs (e.g., the the original network already enjoys an excellent performance, or the connectivity is too sparse or too dense in certain areas). We found difficult to determine a way to distinguish these cases, and we consider this issue as an interesting and potential extension of our work. Despite of that, we have also analyzed the results also calculating the number of instances in which each setting, in terms of number of RNs, achieved the best performance. For the solutions corresponding to values of $K = 0$, $K = 5\%$, $K = 10\%$ and $K = 15\%$, the percentage of cases where each of them achieved the best performance in terms of packet delivery ratio was: 16%, 84%, 90%, 94% respectively. This confirms that in general adding RNs is effective and that their positioning is done properly using the proposed model.

C. Relay node placement to provide connectivity

Apart from the evaluation of our model in the context of network performance, as done in the previous section, we aim to assess the quality of our approach in determining good locations for RNs, even in situations where their deployment is mainly guided by the need of connectivity. First, we consider *partially connected* topology instances and use the *Minimum Spanning Tree Heuristic* (MST) introduced by [4] to determine the RN positions to achieve connectivity. Given that the MST heuristic only provides RN positions, we use our model to solve the routing problem. Secondly, we determine RN locations and solve these same instances with our model to compare the performance through network simulations. Figure 4 shows the performance evaluation results for both strategies. The superior performance obtained by the solutions provided by the model in terms of throughput is substantial. The main reason is likely the lack of notion of network paths in the MST heuristic, in which the placement is focused on minimizing the number of RNs without explicitly taking into account the quality of the routes from SNs to BSs.

V. CONCLUSION

In this paper we formalized the problem of relay node placement in WSNs by defining a linear, mixed integer mathematical programming model. The formulation is based on a flow model and includes a number of constraints and penalty components aimed at closely modeling the specific characteristics of the wireless environment. The model is efficiently solved using a standard solver, CPLEX, and outputs both relay node positions and routing paths to base stations. The quality of the calculated routing pats has been evaluated in extensive simulations in terms of network delivery ratio and end-to-end delays, and compared to the performance obtained using the Collection Tree Protocol (CTP), a state-of-the-art adaptive routing algorithm for WSNs. Results show that our model clearly outperforms CTP. In partially connected topologies, the effectiveness of node relay placement has been compared against the placement defined by a minimum spanning tree heuristic. Also in this case, our model shows superior performance. Finally, we have also studied how the network performance improves by adding relay nodes.

This work introduces several original contributions in the context of relay node placement. Compared to previous MIP models [16], [17], this is the first comprehensive study focusing both on network performance enhancement and connectivity issues, and explicitly including in the model most of the critical aspects of wireless environments.

In the future, we will test our approach using real WSNs and mobile robots, and we will work out its evolution into a fully distributed scheme. On-line estimators will be considered to measure node positions, link costs, and traffic loads.

REFERENCES

[1] I. Akyildiz, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, Mar. 2002.

[2] P. Levis, N. Lee, M. Welsh, and D. Culler, "Tossim: Accurate and scalable simulation of entire tinyos applications," in *Proc. of SenSys'03*. ACM, 2003, pp. 126–137.

[3] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection tree protocol," in *Proc. of SenSys'09*. ACM Press, 2009, pp. 1–14.

[4] G. Lin, "Steiner tree problem with minimum number of steiner points and bounded edge-length," *Information Processing Letters*, vol. 69, no. 2, pp. 53–57, Jan. 1999.

[5] D. Chen, "Approximations for steiner trees with minimum number of steiner points," *Theoretical Computer Science*, vol. 262, no. 1-2, pp. 83–99, Jul. 2001.

[6] X. Cheng, D.-Z. Du, L. Wang, and B. Xu, "Relay sensor placement in wireless sensor networks," *Wireless Networks*, vol. 14, no. 3, pp. 347–355, Jan. 2007.

[7] E. Lloyd and G. Xue, "Relay node placement in wireless sensor networks," *IEEE Trans. Comput.*, vol. 56, no. 1, pp. 134–138, Jan. 2007.

[8] A. Efrat, S. Fekete, P. Gaddehosur, J. Mitchell, and V. Polishchuk, "Improved approximation algorithms for relay placement," *Algorithms–ESA*, 2008.

[9] S. Li, G. Chen, and W. Ding, "Relay node placement in heterogeneous wireless sensor networks with basestations," in *Proc. of CMC'09*, no. 1. IEEE, Jan. 2009, pp. 573–577.

[10] S. Misra, S. D. Hong, G. Xue, and J. Tang, "Constrained relay node placement in wireless sensor networks: Formulation and approximations," *IEEE/ACM Transactions on Networking*, vol. 18, no. 2, pp. 434–447, 2010.

[11] B. Hao, H. Tang, and G. Xue, "Fault-tolerant relay node placement in wireless sensor networks: formulation and approximation," in *2004 Workshop on High Performance Switching and Routing*. IEEE, 2005, pp. 246–250.

[12] J. Tang, B. Hao, and a. Sen, "Relay node placement in large scale wireless sensor networks," *Computer Communications*, vol. 29, no. 4, pp. 490–501, Feb. 2006.

[13] Y. Hou, Y. Shi, H. Sherali, and S. Midkiff, "Prolonging sensor network lifetime with energy provisioning and relay node placement," in *Proc. of IEEE SECON'05*, 2005, pp. 295 – 304.

[14] G. Wang, L. Huang, H. Xu, and J. Li, "Relay node placement for maximizing network lifetime in wireless sensor networks," in *Proc. of WiCOM'08*. IEEE, Oct. 2008, pp. 1–5.

[15] E. Falck, P. Floréen, P. Kaski, J. Kohonen, and P. Orponen, "Balanced data gathering in energy-constrained sensor networks," *Algorithmic Aspects of Wireless Sensor Networks*, vol. 3121, pp. 59–70, 2004.

[16] M. Patel, R. Chandrasekaran, and S. Venkatesan, "Energy efficient sensor, relay and base station placements for coverage, connectivity and routing," in *Proc. of IEEE IPCCC'05*. IEEE, 2005, pp. 581–586.

[17] S. Ergen and P. Varaiya, "Optimal placement of relay nodes for energy efficiency in sensor networks," in *Proc. of ICC'06.*, vol. 8. IEEE, 2006, pp. 3473–3479.

[18] G. Wang, L. Huang, H. Xu, and Y. Wang, "Reliable relay node placement in wireless sensor network," in *Proc. of ICST*. IEEE, 2008, pp. 371–375.

[19] W. Zhang, G. Xue, and S. Misra, "Fault-tolerant relay node placement in wireless sensor networks: Problems and algorithms," in *Proc. of IEEE INFOCOM'07*. IEEE, May 2007, pp. 1649–1657.

[20] A. Kashyap, S. Khuller, and M. Shayman, "Relay placement for higher order connectivity in wireless sensor networks," in *Proc. of IEEE INFOCOM'06*, Apr. 2006, pp. 1–12.

[21] X. Han, X. Cao, E. Lloyd, and C. Shen, "Fault-tolerant relay node placement in heterogeneous wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 1, 2009.

[22] A. Kashyap and M. Shayman, "Relay placement for minimizing congestion in wireless backbone networks," in *Proc. of IEEE WCNC'06*. IEEE, 2006, pp. 159–164.

[23] D. Yang, S. Misra, X. Fang, G. Xue, and J. Zhang, "Two-tiered constrained relay node placement in wireless sensor networks: Efficient approximations," in *Proc. of IEEE SECON'10*. IEEE, 2010, pp. 1–9.

[24] Y. Zhu and Q. Han, "The more relay nodes, the more energy efficient?" in *Proc. of IEEE CSE'09*. IEEE, Aug. 2009, pp. 87–92.

[25] D. S. J. D. Couto, B. A. Chambers, and R. Morris, "Performance of multihop wireless networks : Shortest path is not enough," *Computer Communications*, vol. 33, no. 1, 2003.

[26] M. Haenggi and D. Puccinelli, "Routing in ad hoc networks: a case for long hops," *IEEE Commun. Mag.*, vol. 43, no. 10, pp. 93–101, 2005.

[27] E. Feo Flushing and G. A. Di Caro, "An analytical model of IEEE 802.15.4 non-beacon enabled CSMA/CA in multi-hop wireless sensor networks," IDSIA, Lugano, Switzerland, Tech. Rep. 05-11, 2011.

[28] E. Feo, "Optimal relay node placement for throughput enhancement in wireless sensor networks," Master's thesis, Trento, Italy, December 2010. [Online]. Available: http://www.idsia.ch/~gianni/Papers/Feo_thesis.pdf

[29] J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. Culler, and K. Pister, "System architecture directions for networked sensors," *SIGPLAN Not.*, vol. 35, pp. 93–104, November 2000.

[30] S. Moeller, A. Sridharan, B. Krishnamachari, and O. Gnawali, "Routing without routes: The backpressure collection protocol," in *Proc. of IPSN*. ACM, 2010, pp. 279–290.

[31] Y. Chen, O. Gnawali, M. Kazandjieva, P. Levis, and J. Regehr, "Surviving sensor network software faults," in *Proc. of ACM SIGOPS*. ACM, 2009, pp. 235–246.