A Multi-objective Approach for Data Collection in Wireless Sensor Networks

Christelle Caillouet, Xu Li, Tahiry Razafindralambo

A Multi-objective Approach for Data Collection in Wireless Sensor Networks

Christelle Caillouet[*1], Xu Li[2], and Tahiry Razafindralambo[2]

[1] Lehrstuhl II für Mathematik - RWTH Aachen University, Germany
caillouet@math2.rwth-aachen.de
[2] INRIA Lille - Nord Europe, France
xu.li, tahiry.razafindralambo@inria.fr

Abstract. Wireless sensors networks (WSNs) are deployed to collect huge amounts of data from the environment. This produced data has to be delivered through sensor's wireless interface using multi-hop communications toward a sink. The position of the sink impacts the performance of the wireless sensor network regarding delay and energy consumption especially for relaying sensors. Optimizing the data gathering process in multi-hop wireless sensor networks is, therefore, a key issue. This article addresses the problem of data collection using mobile sinks in a WSN. We provide a framework that studies the trade-off between energy consumption and delay of data collection. This framework provides solutions that allow decision makers to optimally design the data collection plan in wireless sensor networks with mobile sinks.

Keywords: mobile sink, data collection, energy

1 Introduction

Wireless sensor networks (WSNs) have received a lot of attention in recent years due to their potential applications in various areas such as environment monitoring or tracking [1, 5, 12]. In order to get useful and up-to-date information from the environment, the network is composed of a large number of low-capacity (processor, memories, battery) sensors. As the number of sensors increases, the amount of data in the network also increases. The data generated by the sensors has, then, to be sent to a central entity, called sink, for storage and processing. Thanks to the wireless communication capabilities and the protocols developed in the literature, multi-hop transmissions can be used to route data from a sensor to the sink if no direct connection is available. However, this classical N to 1 communication paradigm rapidly consumes the energy of intermediate sensors and provides unfair delay distribution depending on the distance to the sink [8]. Therefore, data collection becomes a key issue in wireless sensor networks.

The data collection in wireless sensor networks consumes energy and needs low delay depending on the application. Reducing the energy consumption while

---

increasing the amount of generated data to have a correct view of the environment is a great challenge. Due to this conflicting goals, the trade-off between energy consumption and environment observation accuracy is still an hot topic in wireless sensor networks. Moreover, WSNs are more and more used for delay sensitive applications such as battlefield monitoring. In these applications, delay reduction between data generation and data processing becomes mandatory.

The literature shows different ways to reduce the amount of transiting data in the network. On one hand, data aggregation techniques [9] limit the generated data by using forecasting. These techniques have some strong assumptions regarding the data. However, forecasting strongly reduces the amount of generated data and thus increases the network lifetime. On the other hand, the use of mobile sinks reduces the number of forwarding sensors [11] but needs motion capabilities for the sinks. In this paper, we focus on data collection using mobile sinks in wireless sensor networks with the objective of minimizing the energy consumption and the delay of data collection.

Our purpose is to determine where to place a set of gateways (or collection points) that are defined to collect the produced data of a region in the WSN field, and compute the route of a mobile sink moving along the gateways to gather the data from the sensors. To answer these questions, we propose a Multi-objective Linear Programming (MLP) framework that allows to optimally place the gateways and minimize jointly the energy spent in the WSN and the route of the mobile sink. Multi-objective optimization does not compute an unique solution, but a set of "best" solutions, called the *Pareto front*, capturing the trade-offs between the different metrics. Solving a multi-objective problem consists in finding the Pareto front, from which the decision maker chooses the solution that fits the best his needs. In this work, each point of the Pareto set is obtained by solving an optimization problem. The main contribution of this work is to give a multi-criteria vision of the data collection problem in WSNs. As far as we know there is no multi-objective analysis in this subject. The developed solutions reduce the overall energy consumption but also reduce the delay of data delivery from the sensors to the sink. Unlike the works proposed in the literature, the results of this paper are twofold. First, we tackle the problem of optimal placement of data collection points in an energy-efficient WSN. Second, we optimize the data collection tour by the sink to minimize the delay. In the first contribution, our aim is to reduce the energy spent by intermediate sensors and in the second contribution, we focus on delay reduction.

In Section 2, we review the previous proposed solutions for optimizing the energy consumption and data collection in WSNs, and describe our assumptions for this work. Section 3 presents the formulation of our proposed MLP. Section 4 presents the method to optimaly solve our MLP. Section 5 shows the experimental results and Section 6 concludes the paper.

2 Background and assumptions

2.1 State of the art

Various solutions have been proposed to extend the network lifetime and reduce delay for data collection. Some solutions have proposed to deploy static sinks in order to reduce the traffic bottlenecks which affect the energy efficiency and the WSN lifetime. In [3], authors propose an heterogeneous view of the network. They develop an ILP (Integer Linear Program) formulation for placing a minimum number of collection points (or gateways) and ensure connectivity among them and the sink to form a wireless mesh network to deliver the data. They minimize the number of collection points and the maximum distance between the sensors and a gateway so that energy consumption is minimized. The use of mobile sinks instead of static sinks to collect the data is more efficient and significantly increases the lifetime of the sensor network [4, 11]. In these works, the location of the mobile sinks is periodically computed so that the network lifetime is maximized. Some research efforts have focused on approaches either minimizing the energy consumed by the sensors [4], or maximizing the global network lifetime [2, 7].

Considering the route of the mobile sinks in WSNs instead of its periodical relocation has not been addressed in previous work to the best of our knowledge. We propose a different formulation of the problem, seeking to jointly compute an optimal placement of the collection points (or gateways) in an energy-efficient way, while minimizing the length of the route of a mobile sink. The extension of our model to deal with several mobile sinks is straightforward as described in the next section. We do not claim to provide a unique *best* solution of the problem, but we study the trade-off between the energy spent by the sensor nodes in data collection, and the delay induced by a mobile sink collecting data at collection points in the WSN. Thus, our work provides solutions that allow decision makers to optimally design the data collection plan in wireless sensor networks with mobile sinks.

2.2 Motivating application

We consider a multi-tiered network structure in order to improve energy efficiency. The hierarchical architecture considered in this paper has been designed for WSN and consists of multi-tiers:

- `Sensors` are static devices that are capable of collecting information and are resource constrained. The sensors are spread out over the sensing field following a distribution that may be probabilistic or deterministic.
- `Gateways` (or `collection points`) are static devices deployed to collect the traffic of a region of the network.
- `Sinks` are devices with motion capabilities which gather data from sensors.

In our application, the sinks gather data at different collection points of the sensing field. Since the collection points may not be inside the communication

range of a given sensor, sensors use multi-hop communications to send data to a given collection point. During the bootstrapping phase of our application, a set of possible collection points are defined. An optimal subset of collection points is chosen in order to reduce the application delay and to enhance the energy efficiency of the whole network. Moreover, since we consider mobile sinks, it is also important to optimize the route of the mobile sink in order to reduce its energy consumption and the delay for data delivery.

2.3   Model and assumptions

We assume that the routing in the WSN (from sensors to collection points) is given so that our model is independent on the specific routing strategy. We select and place the gateways in order to minimize the energy consumption of the WSN.

The energy model considered for the sensors is based on the *first order radio model* described in [6]. A sensor consumes $\epsilon_{elec} = 50\ nJ/bit$ to run the transmitter or receiver circuitry, and $\epsilon_{amp} = 100\ pJ/bit/m^2$ for the transmitter amplifier. Thus, to receive a k-bit message, sensor i consumes:

$$E_r = \epsilon_{elec}k, \qquad \text{and} \qquad E_t = \epsilon_{elec}k + \epsilon_{amp}dist^2(i,j)k, \qquad (1)$$

to transmit this message to a neighbor j, where $dist(i,j)$ is the euclidean distance between i and j.

In the following, we propose a Multi-objective Linear Programming (MLP) framework that allows to study the trade-offs between the length of the route of the mobile sinks associated with a computed gateway placement, and the overall energy consumption in a wireless sensor network.

3   Problem definition

In this work, we focus on data collection in WSNs minimizing the energy consumption and the delay. A set of gateways is chosen to collect the traffic of a region in the WSN field, and a mobile sink moves along the gateways to gather the data from each region.

Given a wireless sensor network represented by a set of sensor nodes $\mathcal{S}$, we define a set of candidate sites $\mathcal{C}_{\mathcal{S}}$ which can potentially be a collection point. We want to determine the gateways' location such that each sensor is associated with its closest one. In order to associate each sensor with its closest gateway, we order, for each sensor node $i \in \mathcal{S}$, the reachable gateways in the vector O_i: if $j < k$, then $dist(i,j) \leq dist(i,k)$ and $O_i(j)$ is before $O_i(k)$ in the vector. Sets $\mathcal{J}_i$ are index sets of vector O_i.

The routing table is an input of our optimization problem. We thus have the set P of paths between the sensors and the candidate sites where we can potentially deploy a gateway. $\mathcal{O}(p)$ (resp. $\mathcal{D}(p)$) denotes the source node (resp.

the destination node) of path $p \in P$. From the set of paths, we introduce the matrix C to indicate the connectivity between the sensors and the gateways:

$$C_{ic} = \begin{cases} 1 & \text{if there exists a route between sensor } i \text{ and candidate site } c \\ 0 & \text{otherwise} \end{cases}$$

The decision variables used in our MLP are the following:

$$E_{max} = \text{the maximum amount of energy consumed by a sensor}$$

$$x_{ij} = \begin{cases} 1 & \text{if sensor } i \text{ is assigned to gateway } j \\ 0 & \text{otherwise} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if a gateway is installed at candidate site } j \\ 0 & \text{otherwise} \end{cases}$$

$$\chi_{ij} = \begin{cases} 1 & \text{if two gateways are installed at candidate sites } i \text{ and } j \text{ and} \\ & \text{the link } (i,j) \text{ is selected for the route of the mobile sink} \\ 0 & \text{otherwise} \end{cases}$$

To evaluate the overall quality of our solutions, we use the following metrics:

- *MinMaxE (f^1)*: Balancing the energy spent by the sensor nodes, that can be viewed as WSN lifetime maximization. From the energy model described in Section 2.3, we seek to minimize the maximum energy spent by each sensor node: $\min \max_{s \in \mathcal{S}}(E_{tot}(s))$, where $E_{tot}(s)$ is the total amount of traffic sent and relayed by sensor s in the WSN. The formal definition of $E_{tot}(s)$ is given in the next subsection.
- *MinRoute (f^2)*: Minimizing the length of the route of the mobile sink between the different installed gateways: $\min \sum_{i \in \mathcal{C}_\mathcal{S}} \sum_{j \in \mathcal{C}_\mathcal{S}} dist(i,j)\chi_{ij}$.

3.1   Multi-objective linear program

The optimization problem of placing the gateways such that we jointly minimize the length of the mobile sink route, and the energy spent by the sensor nodes is the following:

$$\begin{cases} \text{(i)} & \min f^1 = E_{max} \\ \text{(ii)} & \min f^2 = \displaystyle\sum_{i \in \mathcal{C}_\mathcal{S}} \sum_{j \in \mathcal{C}_\mathcal{S}} dist(i,j)\chi_{ij} \end{cases} \tag{2}$$

$$\sum_{j \in \mathcal{C}_{\mathcal{S}}} x_{ij} = 1 \qquad \qquad \forall i \in \mathcal{S} \qquad (3)$$

$$x_{ij} \leq C_{ij} y_j \qquad \qquad \forall i \in \mathcal{S}, j \in \mathcal{C}_{\mathcal{S}} \ (4)$$

$$y_{O_i(k)} + \sum_{h \in \mathcal{J}_i, h > k} x_{iO_i(h)} \leq 1 \qquad \forall i \in \mathcal{S} \qquad (5)$$

$$\sum_{\substack{p \in P, \, i \in p, \\ i \neq \mathcal{O}(p)}} (E_r + E_t) \, x_{\mathcal{O}(p)\mathcal{D}(p)} + \sum_{\substack{p \in P, \\ i = \mathcal{O}(p)}} E_t \, x_{i\mathcal{D}(p)} \leq E_{max} \quad \forall i \in \mathcal{S} \qquad (6)$$

$$\sum_{i \in \mathcal{C}_{\mathcal{S}}} \chi_{ij} = y_j, \sum_{j \in \mathcal{C}_{\mathcal{S}}} \chi_{ij} = y_i \qquad \forall i, j \in \mathcal{C}_{\mathcal{S}} \qquad (7)$$

$$\sum_{i,j \in K} \chi_{ij} \leqslant \sum_{i \in K \setminus \{k\}} y_i + 1 - y_c \qquad \begin{array}{l} \forall k \in K \subset \mathcal{C}_{\mathcal{S}}, \\ c \in \mathcal{C}_{\mathcal{S}} \setminus K \end{array} (8)$$

The objective (2) seeks to minimize *(i)* the maximum energy consumed by the sensor nodes, and *(ii)* the length of the route of a mobile sink along the placed gateways, subject to the euclidean distance between the gateways. Constraints (3) and (4) ensure that each sensor must be associated with an installed gateway that can be reached using a given existing path. Constraints (5) force each sensor to be assigned to its closest gateway.

The objective (2)(i) and constraints (6) try to minimize the maximum amount of energy spent by the sensors. Each sensor has to send its own traffic to its associated gateway, and it has also to forward traffic received by other sensors destinated to their associated collection points. According to the definitions in Section 2.3, the total energy spent by sensor i to relay the data from another sensors associated to an installed collection point j equals:

$$E_{Ftot}(i) = \sum_{p \in P, \, i \in p, \, i \neq \mathcal{O}(p), \, y_{\mathcal{D}(p)} = 1} (E_r + E_t) \, x_{\mathcal{O}(p)\mathcal{D}(p)}.$$

We have to add i's traffic to this formula:

$$E_{Stot}(i) = \sum_{p \in P, \, i \in p, \, i = \mathcal{O}(p), \, y_{\mathcal{D}(p)} = 1} E_t \, x_{i\mathcal{D}(p)}.$$

Leading to the total amount of energy spent by i: $E_{tot}(i) = E_{Ftot}(i) + E_{Stot}(i)$.

In constraints (4), the sensor-gateway association cannot exist if the candidate site is not chosen, i.e. $y_j = 0 \Rightarrow x_{ij} = 0$, $\forall i \in \mathcal{S}$. $E_{tot}(i)$ can thus be replaced by the left-hand side of constraints (6). We then bound this amount by E_{max} which is to be minimized in the objective function. Equalities (7) force the mobile sink to visit all the chosen collection points to collect the data generated by the sensors. They refer to the NP-complete *Travelling Salesman Problem (TSP)*. Since our problem combines gateway placement and TSP, we do not know *a priori* the number of deployed gateways that have to be part of the TSP tour but only the number of candidate sites. The classical linear constraint for subtour elimination is not enough to ensure subtour elimination within the

deployed gateways. Constraints (8) are so-called *generalized subtour elimination contraints* and ensure that there must be a selected edge between a subset K and $\overline{K} = \mathcal{C}_\mathcal{S} \setminus K$ only if there is at least a deployed gateway inside and outside of K. Unfortunately, the number of subsets $K \subset \mathcal{C}_\mathcal{S}$ is exponential in the cardinality of $\mathcal{C}_\mathcal{S}$. To avoid the complete enumeration of the subsets, we proceed as follows:

1. Solve the MLP without subtour elimination.
2. Check the solution: if it has no subtour, we are done.
3. If there is a subtour, add the corresponding constraint (8) for the subtour and solve the program again.
4. Iterate until a solution without subtours is found.

3.2 Multiple mobile sinks

The extension with several mobile sinks is straightforward in our model and can be done in the following way:

– Introduce the set $\mathcal{M}$ of mobile sinks, and the decision variable χ_{ij}^m to be the selection of link (i, j), $i, j \in \mathcal{C}_\mathcal{S}$ in the route of mobile sink $m \in \mathcal{M}$.
– Add the set of constraints specifying that each link between two gateways must be selected by exactly one mobile sink:

$$\sum_{m \in \mathcal{M}} \chi_{ij}^m = \chi_{ij} \qquad \forall i, j \in \mathcal{C}_\mathcal{S}. \tag{9}$$

– Ensure that the route of each mobile sink is a tour:

$$\sum_{j \in \mathcal{C}_\mathcal{S}} \chi_{ji}^m = \sum_{j \in \mathcal{C}_\mathcal{S}} \chi_{ij}^m \qquad \forall i \in \mathcal{C}_\mathcal{S},\ m \in \mathcal{M}. \tag{10}$$

Investigation regarding multiple mobile sinks is out of the scope of this paper and is left as future works. We can also notice that this simple extension can be enhanced by a coordination of the different mobile sinks' tours.

4 Energy-delay trade-off

In order to determine the optimal solutions of our problem, we have solved the proposed MLP with IBM CPLEX solver[3] version 12 on an INTEL Core 2 2.4 GHz with 2 Gb of memory.

Combining the two metrics of the objective function (2) for our problem is not relevant. Indeed, there exists confliction between route length and energy consumption in our optimization framework, i.e., pursuing the optimization of the route length of the mobile sink inextricably damages the performance of energy spent by the sensors. Saving energy enforces to deploy more collection points in the network, thus increasing the length of the route for the mobile sinks. As a result, the length of the route would be damaged, and vice versa.
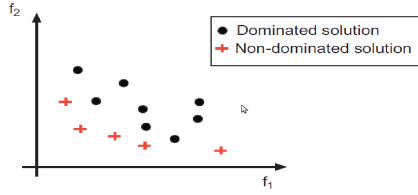
---

[3] http://www.ibm.com/software/integration/optimization/cplex-optimizer/

Fig. 1. Non-dominated and dominated solutions for a 2-function minimization problem.

Consequently, for such a multi-objective optimization problem in which the objectives cannot be optimized simultaneously, the concept of Pareto optimality was introduced into the evaluation system.

The main idea to study the trade-offs between the two metrics *MinMaxE* and *MinRoute* is to find out all the possible non-dominated solutions of the optimization problem. In a general multi-objective problem of the form:

$$\begin{cases} \min f(x) = (f^1(x), f^2(x), \ldots, f^k(x)) \\ s.t. \quad g_i(x) \leqslant 0, \ i = 1, \ldots, m \end{cases}$$

Where $x \in \mathbb{R}^n$ is the decision vector belonging to the feasible region $F = \{x \in \mathbb{R}^n \mid g_i(x) \leqslant 0, i = 1, 2, \ldots, m\}$, a solution $x_2 \in \mathbb{R}^n$ is *dominated* if there is another solution $x_1 \in \mathbb{R}^n$ such that:

- The decision vector x_1 is not worse than x_2 in all objectives:
  $f^i(x_1) \leqslant f^i(x_2), \quad \forall i = 1, 2, \ldots, k.$
- The decision vector x_1 is strictly better than x_2 in at least one objective:
  $f^i(x_1) < f^i(x_2)$ for at least one $i = 1, 2, \ldots, k.$

A solution is non-dominated if there is no other solution dominating it. Informally, this means that if a solution is non-dominated within the whole solution space, it is not possible to improve one of the metrics without worsening at least one of the other metrics. Each multi-objective problem has a set of Pareto-optimal solutions defined as the set of non-dominated solutions. The set of all non-dominated solutions is the *Pareto front* [10]. The Pareto front provides a set of solutions that can be chosen depending on the application requirements.

More precisely, each non-dominated solution represents a different optimal trade-off between the objectives. In this paper, the objective functions of our data collection problem for WSNs are $f^1 = MinMaxE$ and $f^2 = MinRoute$, which are used in the Pareto dominance comparison as in Figure 1. In order to generate Pareto-optimal solutions on the Pareto front, we use the ϵ-constraint method that transforms the multi-objective problem into a sequence of parameterized single-objective problems such that the optimum of each single-objective problem corresponds to a Pareto-optimal solution [10]. We thus generate and solve mono-objective optimization problems of the form:

$$\begin{cases} \min f^i(x) \\ s.t. \quad f^j(x) \leqslant \epsilon^j, \ \forall j \neq i \end{cases}$$
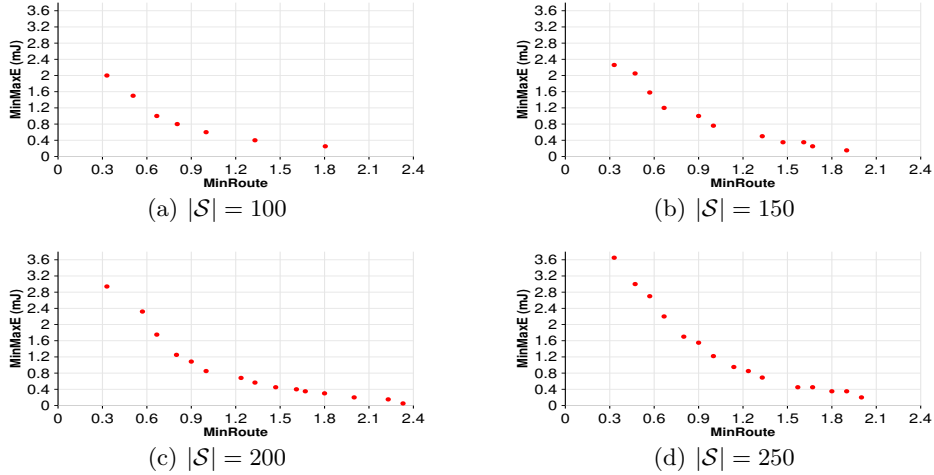
Fig. 2. Pareto fronts for random networks with 25 candidate sites and different number of sensors $|\mathcal{S}|$.

The ϵ^i are chosen such that $f^{i*} \leqslant \epsilon^i$, where f^{i*} corresponds to the optimum value of the mono-objective problem minimizing only objective f^i.

5   Performance evaluation

In this section, we present results obtained by solving our MLP with the ϵ-constrained method in order to get optimal solutions on the Pareto front. We study networks of size between 50 and 250 sensors whose position is randomly chosen in a square area. The length is normalized so that the WSN is deployed in a unitary squared area. For each random network computed, we use two policies of candidate sites locations: a regular and a random one. The first model divides the area considered into equal squares in which one candidate site is placed in the center of the square. In this way, the candidate sites for placing the collection points form a regular grid. The second policy chooses randomly the location of the candidate sites in the area. We have tested the two policies with $\{3^2, 4^2, 5^2\}$ candidate sites. The set of paths considered is comprised of shortest paths between each pair (sensor, candidate site).

5.1   Effect of candidate sites and network density

To demonstrate the utility of the approach, we generate sets of non-dominated solutions by iteratively solving ϵ-constrained mono-objective optimization problems. Results are depicted in Figures 2 and 3 for various network sizes and numbers of candidate sites.

We can easily see that limiting the energy spent by each sensor node increases the length of the mobile sink route. Moreover, the number of installed gateways
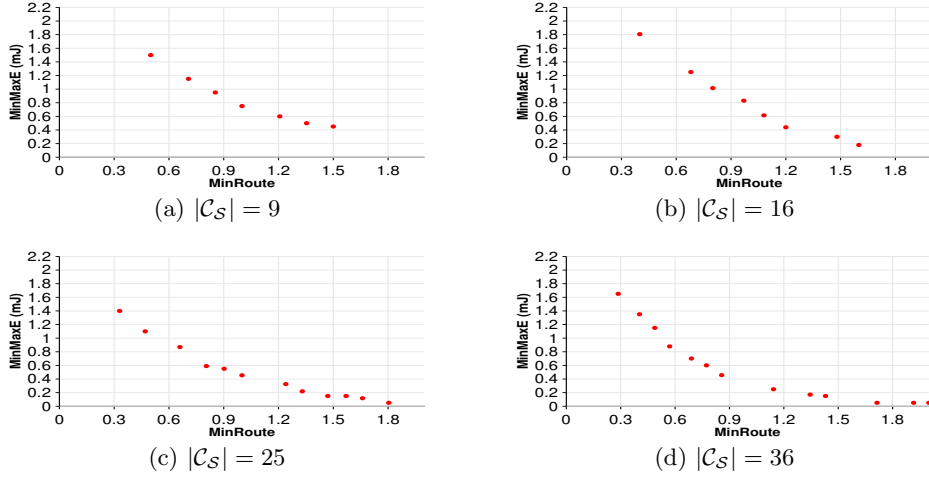
(a) $|\mathcal{C}_\mathcal{S}| = 9$

(b) $|\mathcal{C}_\mathcal{S}| = 16$

(c) $|\mathcal{C}_\mathcal{S}| = 25$

(d) $|\mathcal{C}_\mathcal{S}| = 36$

Fig. 3. Pareto fronts for random networks with 50 sensors and different number of candidate sites $|\mathcal{C}_\mathcal{S}|$.

(a) Number of deployed gateways depending on E_{max}.

(b) Number of paths through the most loaded sensor depending on gateways.
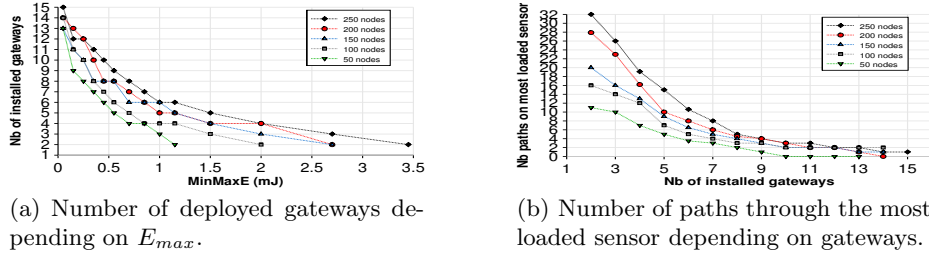
Fig. 4. Results for random networks with 25 candidate sites.

strongly depends on the limit of the energy spent ($MinMaxE$): The amount of forwarding traffic from other sensors is strongly limited. In particular, when we focus on energy (optimizing only $MinMaxE$, without any constraint on the number of deployed collection points) the optimal solution of our MLP minimizes the energy spent by each sensor essentially by limiting its forwarding traffic. We can thus see that the placement of the collection points verifies that each sensor is a neighbor of its associated gateway (when possible).

In order to get a better analysis about the forwarding traffic related to the energy consumed by the sensors, we have computed the number of paths going through the most loaded sensor node for a given placement of collection points. More formally, given the set of paths as an entry of our problem, we compute for each sensor i the number of paths containing i that have a gateway placed as destination node, and so that the source node of the path is associated with

this gateway:

$$Load(i) = \sum_{p \in \mathcal{P} | i \in p} x_{\mathcal{O}(p)\mathcal{D}(p)}.$$

The most loaded sensor node is therefore the one that has the maximum number of paths going through it: $Load(\mathcal{S}) = \max_{i \in \mathcal{S}} Load(i)$. Figure 4(b) represents another way of viewing the trade-off between the two objectives, by depicting the maximum number of paths going through the most loaded sensor node in function of the number of deployed gateways in the WSN. We can see that deploying more gateways allows to limit the amount of forwarding traffic at each sensor. When the number of gateway is large enough so that each sensor is a neighbor of one gateway (when possible), then the forwarding traffic becomes null and the sensor spends energy only for sending its own traffic.

When the size of the network increases, then the energy consumption of the sensor nodes also increases (see Figure 2). Indeed, the total traffic in the WSN is more important, so the sensors have more forwarding traffic to relay which increases their load. When the energy is limited to $E_{max} = 2$ mJ for each sensor, it is worth noting that the length of the mobile sink route also increases with the network size. The average route length of the mobile sink equals respectively 0.33, 0.47, 0.6, and 0.65 for a WSN of 100, 150, 200, and 250 nodes. However, the maximum amount of energy spent by the sensor nodes decreases when the number of candidate sites in the network increases as depicted in Figure 3. On one hand, placing a gateway in a candidate site reduces the relaying traffic and, therefore, reduces the energy spent. On the other hand, the length of the mobile sink's route increases, especially when the energy consumed by each sensor is low. This assertion is confirmed by Figure 4(a) that depicts the number of deployed gateways depending on the maximum energy spent by the sensors.

5.2 Graphical trade-off interpretation

The trade-offs among the different metrics are shown using value paths in Figure 5. Value paths have proven to be an effective way to present trade-offs in multi-objective problems [13]. In the figure, there is a vertical axis for each objective. The value assigned to each non-dominated solution on a particular axis is that solution's value divided by the best solution possible for that objective: i.e. f^i/f^{i*} for each $i = 1, 2$. The minimum value for each axis is 1.00, corresponding to the optimal mono-objective solution f^{1*} and f^{2*}: The minimum energy consumption for $MinMaxE$ and the minimum possible route length $MinRoute$ for the mobile sink.

For a network of 50 sensors and 9 candidate sites (Fig. 5(a)), the optimal solution for $MinMaxE$ leads to a route 3 times longer than the optimal one. On the contrary, having the optimal route length incurs 233 per cent more in energy consumption than does the best solution. For a fair trade within the two-objectives, the best solution would be of minimizing the maximum value on all the axis: $\min_v \max_{f^1, f^2} \{f^1(v)/f^{1*}, f^2(v)/f^{2*}\}$. The non-dominated solution optimizing the fair trade is thus the one incuring 66 percent more in energy
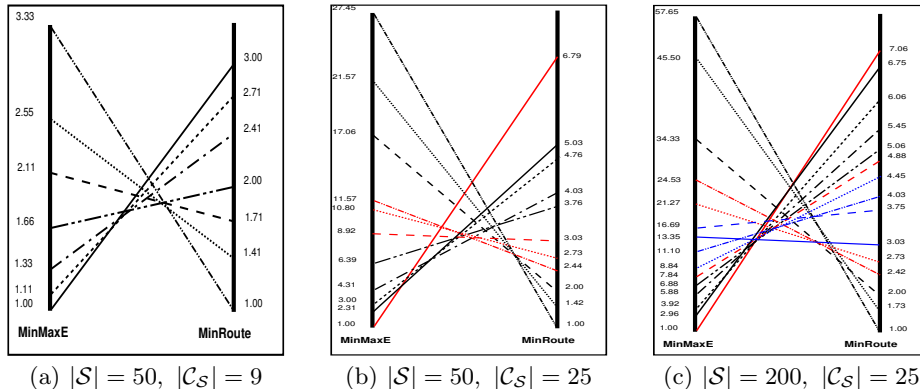
(a) $|\mathcal{S}| = 50$, $|\mathcal{C}_{\mathcal{S}}| = 9$ (b) $|\mathcal{S}| = 50$, $|\mathcal{C}_{\mathcal{S}}| = 25$ (c) $|\mathcal{S}| = 200$, $|\mathcal{C}_{\mathcal{S}}| = 25$

Fig. 5. Percentage path analysis for delay-energy trade-off.

consumption and 100 per cent more in route length. On the Pareto front this solution is located in the bottom of the curve of the Pareto front, with value $MinMaxE = 0.75mJ$ and $MinRoute = 1$ (see Fig. 3(a)). When the number of candidate sites increases in the network, the maximum gap between the optimal mono-objective and the multi-objective values for $MinMaxE$ becomes really large (Fig. 5(b) and 5(c)). The number of possibilities to deploy the gateways grows exponentialy, therefore this can lead to really bad results in terms of energy, when the route length of the mobile sink is optimized. Value 1 for $MinRoute$ corresponds to value 27.45 for $MinMaxE$ in Figure 5(b). The fair trade non-dominated solution for a 50-sensors network with 25 candidate sites has value 4.3 and 4 for $MinMaxE$ and $MinRoute$ percentages respectively, with exact values $MinMaxE = 0.22mJ$ and $MinRoute = 1.33$ (Fig. 3(c)). But when the number of sensors increases, while the number of candidate sites stays the same, then the scales in the percentage value paths have the same order for $MinRoute$, and does not increase a lot for $MinMaxE$, because the routing stays roughly the same with the same number of candidate sites. The fair trade non-dominated solution is $MinMaxE = 0.3mJ$ and $MinRoute = 1.8$ for a network with 200 sensors and 25 candidate sites (Fig. 5(c) and 2(c)), which is quite effective. One can remark that the fair trade optimal solution is always closer to the optimal solution of $MinMaxE$ because when the route length is optimized, then the energy consumption becomes really bad in the WSN.

After analysing such trade-offs, a network manager can then decide which option is preferred. The network manager can also choose the required number of collection points deployed and can then find the best way to route the data with minimum energy consumption.

5.3 Impact of the candidate site positioning

The location of the candidate sites among the WSN is also important regarding the network lifetime and the delay of data collection. In Table 1, we present

Topology		Grid placement				Random placement			
$|\mathcal{S}|$	$|\mathcal{C_S}|$	f^{1*} (mJ)	f^{2*}	$\overline{Dpl.\ Gtw}$	$\overline{Load(\mathcal{S})}$	f^{1*} (mJ)	f^{2*}	$\overline{Dpl.\ Gtw}$	$\overline{Load(\mathcal{S})}$
50	9	0.45	0.5	3.5	8.93	0.45	0.21	2.96	13.24
50	16	0.15	0.4	4.25	8	0.35	0.22	4.63	9.83
50	25	0.05	0.33	4.29	7.59	0.15	0.02	4.93	12.57
100	9	0.45	0.5	3.33	10.2	0.65	0.28	3.71	21.26
100	16	0.35	0.4	4.81	9.87	0.65	0.21	4	21
100	25	0.05	0.33	4.7	9.05	0.25	0.03	4.24	23.74
150	9	0.75	0.5	2.85	17	0.95	0.08	3.05	27.88
150	16	0.45	0.4	4.24	12.76	0.75	0.01	4.41	20.53
150	25	0.15	0.33	4.77	14.24	0.25	0.04	3.76	32.78
200	9	0.85	0.5	3.9	16.3	0.95	0.22	3.17	20.79
200	16	0.25	0.4	4.77	13.5	1.05	0.11	3.53	27.4
200	25	0.05	0.33	4.4	15.84	0.35	0.07	4.43	34.3

Table 1. Comparison between regular and random placement of the candidate sites.

results for various topologies in order to compare the two policies of candidate site placement described at the beginning of the section. For each policy (regular grid and random), we compute the optimal value of the 2 mono-objective optimizations f^{1*} and f^{2*}. On one hand, a regular placement of the collection points always saves energy in the WSN, thus increasing the network lifetime. These results are explained by the fact that the distance between a sensor and a gateway can be bounded due to the regular placement of candidate sites. On the other hand, a random placement of the collection points usually leads to a smaller route for the mobile sink, thus leading to a better data collection. This behavior is due to the possible useless placements of gateways since these placements are randomly chosen. Moreover, due to the same reason, the average number of deployed gateways among the Pareto optimum of the bi-objective optimization $\overline{Dpl.\ Gtw}$ is smaller for the random policy. However, the reduction of $\overline{Dpl.\ Gtw}$ increases the sensor's average load $\overline{Load(\mathcal{S})}$ (the mean value $Load(\mathcal{S})$ over the Pareto optimum) compared to a regular placement of candidate sites.

When the candidate sites are regularly placed in the area (i.e. the regular policy as described in Section 5.1), then the sensor's load is significantly reduced in comparison to a completely random placement (see Table 1). This value is always greater of at least 20% when the gateway placement is performed among candidate sites chosen randomly, leading to more loaded sensors in the WSN.

6 Conclusion

In this paper, we have presented a framework for efficient data collection in wireless sensor networks. We have developed a Multi-objective Linear Program with two metrics to evaluate the trade-off between the maximum amount of energy spent by the sensor nodes, and the length of the route of mobile sinks collecting data at collection points that we jointly deploy. We can see from these particular results that the load of the sensors rapidly decreases with the number of gateways and that above a given number of deployed gateways the load remains stable (or decreases slowly). This allows to save energy in the WSN and maximize its lifetime. When the number of deployed gateways is important, then the mobile sink collecting data at the gateways has a longer route to perform, therefore increasing the delay of data collection until processing.

The proposed model therefore provides a means by which several objectives must be evaluated by a network manager. A fair trade optimal solution is drawn from the obtained Pareto fronts to fairly optimize the metrics. If the energy is the major concern, then the network lifetime objective may be favorable. If the decision maker wants to reduce the delay of collecting the data to ensure fast processing, he or she may give a higher priority to the route length objective when compared to other objectives. Hence, there is a critical necessity to incorporate all the different objectives when we model the data collection plan in WSNs.

References

1. Barrenetxea, G., Ingelrest, F., Schaefer, G., Vetterli, M.: The hitchhiker's guide to successful wireless sensor network deployments. In: ACM Conference on Embedded Networked Sensor Systems (ACM SenSys 2008). pp. 43–56. Raleigh, USA (2008)
2. Ben Saad, L., Tourancheau, B.: Multiple mobile sinks positioning in wireless sensor networks for buildings. In: 3rd International Conference on Sensor Technologies and Applications (SENSORCOMM 2009). pp. 264–270. Athens, Greece (2009)
3. Capone, A., Cesana, M., Donno, D., Filippini, I.: Deploying multiple interconnected gateways in heterogeneous wireless sensor networks: An optimization approach. Elsevier Computer Communications 33(10), 1151–1161 (Jun 2010)
4. Gandham, S., Dawande, M., Prakash, R., Venkatesan, S.: Energy efficient schemes for wireless sensor networks with multiple mobile base stations. In: IEEE Globe-Com. pp. 377–381. San Francisco, USA (2003)
5. Hart, J.K., Martinez, K.: Environmental sensor networks: A revolution in the earth system science? Earth-Science Reviews (Elsevier) 78, 177–191 (2006)
6. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless micro sensor networks. In: Hawaii International Conference on System Science (HICSS 2000). pp. 3005–3010. Maui, Hawaii (2000)
7. Kalpakis, K., Dasgupta, K., Namjoshi, P.: Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks. Computer Networks (Elsevier) 42, 697–716 (2003)
8. Khadar, F., Razafindralambo, T.: Performance evaluation of gradient routing strategies for wireless sensor networks. In: International IFIP-TC 6 Networking Conference (Networking 2009). pp. 535–547. Aachen, Germany (2009)
9. Krishnamachari, B., Estrin, D., Wicker, S.B.: The impact of data aggregation in wireless sensor networks. In: International Conference on Distributed Computing Systems (ICDCS 2002). pp. 575–578. Washington, DC, USA (2002)
10. Laumanns, M., Thiele, L., Zitzler, E.: An adaptive scheme to generate the pareto front based on the epsilon-constraint method. In: Branke, J., Deb, K., Miettinen, K., Steuer, R.E. (eds.) Practical Approaches to Multi-Objective Optimization. No. 04461 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany (2005)
11. Luo, J., Hubaux, J.P.: Joint mobility and routing for lifetime elongation in wireless sensor networks. In: IEEE International Conference on Computer Communications (IEEE INFOCOM 2005). pp. 1735 – 1746. Miami, USA (2005)
12. Martinez, K., Ong, R., Hart, J.: Glacsweb: a sensor network for hostile environments. In: IEEE SECON. pp. 81–87. Santa Clara, USA (2004)
13. Miettinen, K.: Graphical illustration of Pareto optimal solutions. In: Tanino, T., Tanaka, T., Inuiguchi, M. (eds.) Multi-Objective Programming and Goal Programming: Theory and Applications. pp. 197–202. Springer-Verlag, Berlin (2003)