An Analytical Study of Reliable and Energy-efficient Data Collection in Sparse Sensor Networks with Mobile Relays

Giuseppe Anastasi[1], Marco Conti[2], and Mario Di Francesco[1]

[1] Dept. of Information Engineering, University of Pisa, Italy
{giuseppe.anastasi,mario.difrancesco}@iet.unipi.it
[2] CNR-IIT, National Research Council, Italy
marco.conti@iit.cnr.it

Abstract. Sparse wireless sensor networks (WSNs) are emerging as an effective solution for a wide range of applications, especially for environmental monitoring. In this context, special mobile elements – i.e. mobile relays (MRs) – can be used to get data sampled by sensor nodes. In this paper we present an analytical evaluation of the data collection performance in sparse WSNs with MRs. Our main contribution is the definition of a flexible model which can derive the total energy consumption for each message correctly transferred by sensors to the MR. The results show that a low duty cycle is convenient and allows a significant amount of correctly received messages, especially when the MR moves with a low speed. When the MR moves fast, depending on its mobility pattern, a low duty cycle may not always be the most energy efficient option.

1 Introduction

Wireless sensor networks (WSNs) have become an enabling technology for a wide range of applications. The traditional WSN architecture consists of a large number of sensor nodes which are densely deployed over an area of interest. Sensor nodes sample data from their surrounding environment, process them locally and send the results to a data collection point, usually a sink node or an Access Point (AP). The communication between the sensors and the data collection point is multi-hop, which is possible due to the network density. More recently, a different WSN architecture has been introduced for application scenarios where fine-grained sensing is not required. In this case, nodes are sparsely deployed over the sensing field. As the number of nodes is moderate or low, in contrast with traditional solutions, the costs are reduced. However, since the network is sparse, neighboring nodes are far away from each other, so that they cannot communicate together directly nor through multi-hop paths and a different data gathering scheme is required.

In sparse sensor networks, data collection can be accomplished by means of *mobile relays* (MRs). MRs are special mobile nodes which are responsible for data gathering. They are assumed to be powerful in terms of data storage and processing capabilities, and not energy constrained, in the sense that their energy source can be replaced or recharged easily. MRs carry data from sensors to the sink node or an infra-structured AP [1]. Depending on the application scenario, MRs may be either part of the external environment [2, 3] (e.g., buses, cabs, or walking people), or part of the network infrastructure [4, 5] (e.g., mobile robots).

The communication between an MR and sensor nodes takes place in two different phases. First, sensor nodes have to discover the presence of the MR in their communication range. Then, they can transfer collected data to the MR while satisfying certain reliability constraints, if required. Different from the MRs, sensor nodes have a limited energy budget, so that both discovery and data transfer should be energy efficient in order to prolong the network lifetime [6]. As the radio component is usually the major source of energy consumption, the overall radio activity should be minimized. To this end, a duty cycle approach can be used, so that sensors alternate sleep and active periods. However, the effects of the duty cycle have to be properly investigated: if sensor nodes are sleeping when the MR comes, they cannot detect it neither transmit data, so that they are only wasting their energy.

In this paper we consider the joint impact of discovery and data transfer for reliable and energy efficient data collection in sparse WSNs with MRs. The major contribution of this paper is a detailed and realistic model for deriving the performance of the overall data collection process. The proposed methodology is general, so that it can be adapted to different discovery and data transfer protocols, and does not depend on the mobility pattern of the MR. To the purposes of our analysis, we consider a discovery scheme based on periodic wakeups and an ARQ-based data transfer protocol. Finally, we derive the performance of data collection in terms of both throughput (i.e. average number of messages correctly transferred to the MR) and energy efficiency (i.e. total energy spent per successfully transferred message) at each contact.

The results obtained show that, in general, a low duty cycle provides a better energy efficiency, especially if the contact time is large enough to allow the reliable transfer of a significant amount of data. However, when the contact time is limited, a very low duty cycle is not convenient as the energy saved during discovery is overcome by the decrease in the number of messages successfully transferred.

The rest of the paper is organized as follows. Section 2 presents an overview of the relevant literature in the field. Section 3 introduces the system model and the related assumptions. Section 4 and 5 present the analytical model for the discovery and the data transfer phases, respectively. Section 6 presents and discusses the obtained results. Finally, section 7 concludes the paper, giving insights for future work.

2 Related work

Many different papers have addressed the issues of data collection using MRs. In the context of opportunistic networks, the well known message ferrying approach has been proposed in [7]. Specifically, power management has been addressed by [8], where a general framework for energy conservation is introduced. The proposed solution, which can also exploit knowledge about the mobility pattern of the MR, is evaluated in terms of energy efficiency and delivery performance.

However, as the proposed solution is devised for opportunistic networks, it is not applicable in the scenario considered in this paper.

Indeed, many solutions have also been conceived specifically for WSNs. While many papers focus on the mobility of the MR [9, 10], some works actually address the problem of energy efficient data collection from the sensor node perspective. For example, [5] considers a periodic wakeup scheme for discovery and a stop-and-wait protocol for data transfer. A stop-and-wait protocol for data transfer is also used in [11], where the MR is assumed to be controllable. A different solution is investigated in [2], under the assumption that the MR has a completely predictable mobility. The above mentioned solutions, however, have only been analyzed with simulations, while in this paper we address the problem analytically. In addition, the solution proposed here is flexible enough to support different protocols and mobility patterns of the MR.

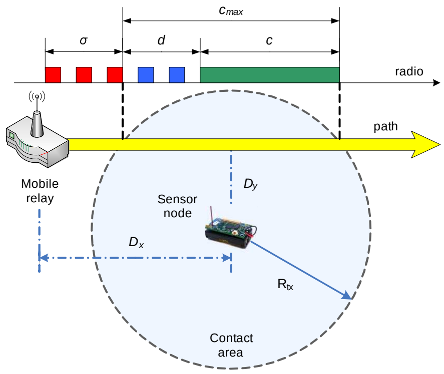
In the specific context of MRs, [3] considers MRs which are not controllable but move randomly, and models the success rate of messages arriving at the access point. However, [3] focuses on buffer requirements at sensors rather than on their energy consumption. Under the same scenario, [1] introduces a more detailed formulation, which considers both the discovery and the data transfer phases of data collection. Furthermore, it evaluates different mobility patterns of the MR and supports sensor nodes operating with a duty cycle during discovery. Although discussing the probability of data reception at the access point, both [3] and [1] assume an ideal channel and no specific data transfer protocol, so that their findings are mostly affected by buffering constraints. Instead, we explicitly consider data transfer – in addition to discovery – for reliable data collection. In addition, we take the message loss into account by using a model derived from real measurements.

The problem of reliable and energy efficient data collection has also been addressed in [12], where an adaptive and window-based ARQ transmission scheme is evaluated under a realistic message loss model derived from real measurements [13]. In detail, [12] analytically shows that the proposed scheme achieves not only a better throughput, but also a higher energy efficiency than a simple stop-and-wait protocol. However, the proposed approach is evaluated only in the application scenario where the sensor has only a limited amount of data to send. In addition, [12] does not consider the effect of discovery on the subsequent data transfer phase, as it focuses only on data transfer. On the contrary, the model presented in this paper jointly considers discovery and data transfer for deriving the overall energy efficiency.

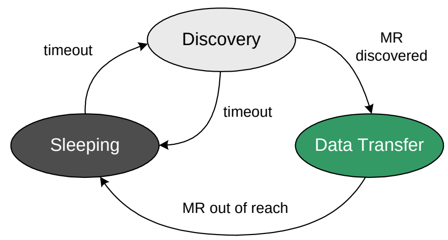
3 System model

In this section, we will first introduce the reference scenario considered in the analysis. Then, we will describe the discovery and the data transfer protocols used by the MR and the static sensors for data collection.

The reference network scenario is illustrated in Fig. 1(a). In the following, we will consider a single MR and assume that the network is sparse enough so that at any time at most a single (static) node can reach the MR. In addition, we will assume that the MR is part of the environment (e.g. a bus or a car), so



(a)



(b)

Fig. 1. Reference scenario (a) and state diagram for the static node (b)

that its mobility pattern cannot be controlled. Finally, we will assume that the MR moves along a linear path at a fixed vertical distance (D_y) from the static node, at a constant speed v. This assumption is reasonable for a sample scenario where the MR moves along a street or a road. Data collection takes place only during a *contact*, i.e. when the static node and the MR can reach each other. Furthermore, the area within the radio transmission range R_{tx} of the static node is called *contact area*[3], and the overall time spent by the MR inside the contact area is called *contact time*, and is referred to as c_{max}. During a contact, messages exchanged between the MR and the static node experience a certain message loss. We denote by $p(t)$ the probability that a message transmitted at time t is lost, and assume as $t = 0$ the instant at which the MR enters the contact area. Any message transmitted when the static node and the MR are not in contact is assumed to get lost, so that $p(t)$ is defined only within the contact area.

The overall data collection process can be split into three main phases. Fig. 1(b) shows the state diagram of the static sensor node [8]. As MR arrivals are generally unpredictable, the static node initially performs a discovery phase for the timely detection of the MR. Indeed, the successful MR detection by the static sensor is not immediate, but requires a certain amount of time, called *discovery time*, and denoted by d in Fig. 1(a). Upon detecting the MR, the static node switches from the discovery state to the data transfer state, and starts transmitting data to the MR. As a result of the discovery process, the static node cannot exploit the whole available contact time for data transfer. The portion of the contact time which can be actually used for subsequent data transfer is called *residual contact time* and is referred to as c. After the end of the data transfer phase, the static node may switch to the discovery state again in order to detect the next MR passage. However, if the MR has a (even partially) predictable mobility, the static node can exploit this knowledge to further reduce its energy consumption [8]. In this case, the static node can go to a sleep state

[3] Depicted with a circular shape in Fig. 1(a) only for convenience.

until the next expected arrival of the MR. In any case, the static sensor may be awake also when the MR is out of reach. The amount of time spent by the static node in the discovery state while the MR has not yet entered the contact area is called *waiting time*, and is indicated with σ in Fig. 1(a).

We now briefly describe the discovery and data transfer protocols used by the static sensor and the MR in the corresponding phases. In principle, different discovery and data transfer protocols could be used for data collection in the above scenario. In our analysis, however, we will consider a discovery protocol based on a periodic beacon transmission by the MR, and an ARQ-based protocol for reliable data transfer, as they are among the most frequent schemes in the field [5, 3, 11]. To advertise its presence in the surrounding area, the MR periodically sends special messages called *beacons*. The duration of a beacon message is equal to T_{BD}, and subsequent beacons are spaced by a *beacon period*, indicated with T_B. In order to save energy during the discovery phase, the static node operates with a duty cycle δ, defined by the active time T_{ON} and the sleep time T_{OFF}, i.e. $\delta = T_{ON}/(T_{ON} + T_{OFF})$. The static node follows a periodic wakeup scheme, with its activity time set to $T_{ON} = T_B + T_{BD}$, i.e. a value which allows the node to receive a complete beacon during its active time, provided that it wakes up when the MR is in the contact area. On the other hand, the sleep time T_{OFF} is set to a value which enforces the desired duty cycle δ.

As soon as it receives a beacon from the MR, the static node enters the data transfer state. While in this state, the static node remains always active to exploit the residual contact time as much as possible. On the other hand, the MR enters the data transfer phase as soon as it receives the first message sent by the static node, and stops beacon transmissions. The communication scheme adopted during the data transfer phase is based on Automatic Repeat reQuest (ARQ). The static node splits buffered data into messages, which are transmitted in groups (windows). The number of messages contained in a window, i.e. the *window size*, is assumed to be fixed and known both at the sender and at the receiver. The static node sends the messages in a window back to back, then waits for an acknowledgement sent back by the MR. Note that, in this context, the acknowledgement message is used not only for implementing a retransmission strategy, but also as an indication of the MR presence in the contact area. In the following, we assume that the static sensor has always data to send, so that the data transfer phase ends when the MR is not reachable any more (i.e. at the end of the residual contact time). However, the static node generally cannot know when the MR will leave the contact area, for instance because it cannot derive the residual contact time a priori. In practice, the static node assumes that the MR has exited the contact area when it misses N_{ack} consecutive acknowledgments. Similarly, the MR assumes that the communication is over when it does not receive any more message in a given period of time.

4 Discovery phase analysis

In this section we develop an analytical model for the discovery phase. The purpose of the analysis is to derive the distribution of the discovery time and, thus, the residual contact time as well. The analysis is split in two main parts.

First, the state of the static node (i.e. ON or OFF) over time is derived, by keeping in consideration the duty cycle. Second, the beacon reception process is modeled, i.e. the state transitions of the static node are characterized, on the basis of the probability that a beacon sent by the MR at a given instant will be correctly received by the static sensor.

With the help of Fig. 2(a) we introduce the framework for the subsequent analysis. As beacon transmissions do not depend on when the MR enters the contact area, the initial beacon transmission within the contact area is generally affected by a random offset (with respect to the beginning of the contact time). In detail, the time instant at which the MR transmits the first beacon while in the contact area is denoted as t_0. As beacon transmissions are periodic and start at t_0, the actual instants of subsequent beacon transmissions can be expressed as $t_n = t_0 + n \cdot T_B$, with $n \in [1, N-1]$ where $N = \lceil c_{max}/T_B \rceil$ is the maximum number of beacons the MR can send while in the contact area. Therefore, if the MR is discovered by means of the m-th beacon, the discovery time is $d = d_m(t_0) = t_0 + m \cdot T_B$, and the corresponding residual contact time is $c = c_m(t_0) = c_{max} - d = c_{max} - d_m(t_0)$.

The state of the static node at a given instant is completely specified by its composite state (s, r) where s denotes the radio state, i.e. ON or OFF, and r represents the residual time, i.e. the amount of time the node will remain in the same state s. The initial state of the static node at the time $t = 0$ is referred to as (s_0, r_0). Let us denote by $s(t)$ and $r(t)$ the radio state and the residual time, respectively, at a generic time t. Because of the duty cycle, both $s(t)$ and $r(t)$ evolve in a deterministic way. In detail, the radio state of the static node is periodic, with period equal $T_{ON} + T_{OFF}$. We focus now on the radio state $s(t_n)$ of the static node at beacon transmission times (t_n). As $s(t)$ is periodic, it is sufficient to investigate the remainder of the ratio between the beacon transmission times and the period of the duty cycle. By comparing this remainder against the initial residual state s_0 and the initial residual time r_0, it is possible to derive $s(t_n)$. Specifically, it is

$$\underset{s_0=\mathrm{ON}}{s(t_n)} = \begin{cases} \mathrm{ON} & \text{if } 0 \le t'_n < r_0 \\ \mathrm{OFF} & \text{if } r_0 \le t'_n < r_0 + T_{OFF} \\ \mathrm{ON} & \text{if } r_0 + T_{OFF} \le t'_n < T_{ON} + T_{OFF} \end{cases} \tag{1}$$

$$\underset{s_0=\mathrm{OFF}}{s(t_n)} = \begin{cases} \mathrm{OFF} & \text{if } 0 \le t'_n < r_0 \\ \mathrm{ON} & \text{if } r_0 \le t'_n < r_0 + T_{ON} \\ \mathrm{OFF} & \text{if } r_0 + T_{ON} \le t'_n < T_{ON} + T_{OFF} \end{cases} \tag{2}$$

where $t'_n = t_n \mod (T_{ON} + T_{OFF})$. Similarly, we can also derive the residual time $r(t_n)$

$$\underset{s_0=\mathrm{ON}}{r(t_n)} = \begin{cases} r_0 - t'_n & \text{if } 0 \le t'_n < r_0 \\ T_{OFF} + r_0 - t'_n & \text{if } r_0 \le t'_n < r_0 + T_{OFF} \\ T_{ON} + T_{OFF} + r_0 - t'_n & \text{if } r_0 + T_{OFF} \le t'_n < T_{ON} + T_{OFF} \end{cases} \tag{3}$$

$$\underset{s_0=\mathrm{OFF}}{r(t_n)} = \begin{cases} r_0 - t'_n & \text{if } 0 \le t'_n < r_0 \\ T_{ON} + r_0 - t'_n & \text{if } r_0 \le t'_n < r_0 + T_{ON} \\ T_{ON} + T_{OFF} + r_0 - t'_n & \text{if } r_0 + T_{ON} \le t'_n < T_{ON} + T_{OFF} \end{cases} \tag{4}$$

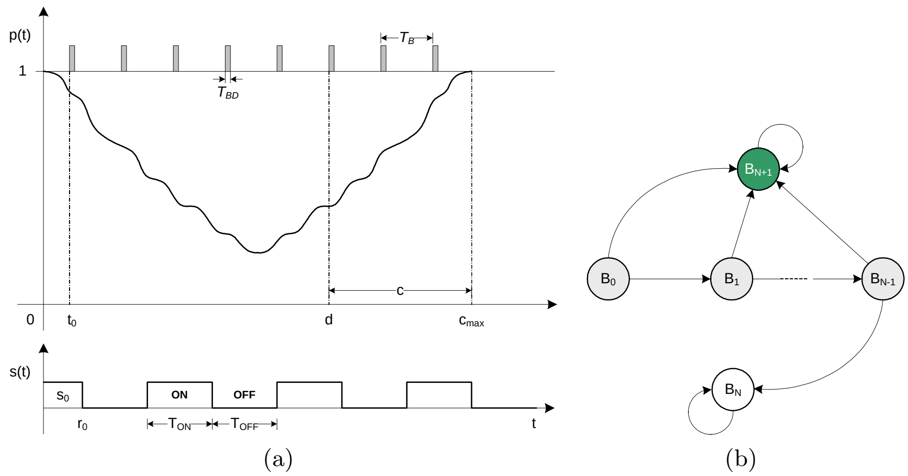


Fig. 2. Beacon discovery process (a) and beacon state (b)

Once the duty-cycled state of the static node has been fully characterized, we have to model the actual beacon reception process. To this end we introduce the state representation illustrated in Fig. 2(b), where the states B_i, $i \in [0, N+1]$, refer to the static node at beacon transmission times. In detail, B_0 is the initial state in which the static node is waiting for the MR to transmit the first beacon in the contact area. B_j is entered after missing the first j beacons sent by the MR, where $j \in [1, N-1]$. B_N is entered when the static node has not detected the MR presence at all, because it has not received any of the beacons. Finally, B_{N+1} is entered when the static node has successfully received a beacon. When it is in a state B_k, with $k \in [0, N-1]$, the static node can only move to the state B_{k+1} or to the state B_{N+1} if it has lost or got a beacon, respectively. Note that B_N and B_{N+1} are absorbing states. In addition, the state of the static node is completely specified by its current state.

Now, we can derive a joint characterization of the radio state of the static node and the beacon reception process. For simplicity, time has been discretized in slots with duration Δ, so that the whole process can be modeled as a discrete time Markov chain. For the sake of clarity, in the following we will not explicitly refer to time-dependent parameters by their actual discretized values, unless otherwise specified. The transition matrix H corresponding to the beacon reception process can be thus written as follows

$$\mathbf{H} = \begin{pmatrix} 0 & H_{01} & 0 & \cdots & 0 & H_{0,N+1} \\ 0 & 0 & H_{12} & & 0 & H_{1,N+1} \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & H_{N-1,N} & H_{N-1,N+1} \\ 0 & 0 & 0 & \cdots & H_{NN} & 0 \\ 0 & 0 & 0 & \cdots & 0 & H_{N+1,N+1} \end{pmatrix}$$

where the H_{kl} are sub-blocks denoting the transition probability from the state B_k to the state B_l. Note that the state B_0 is evaluated at time $t = 0$, while state B_i with $i \in [1, N]$ is evaluated at the i-th beacon transmission time, i.e. t_i. In addition to the state B related to the beacon reception, the H_{kl} blocks also keep track of the radio state of the static node. In detail, the elements of the H_{kl} block can be expressed as $h^{kl}_{(s_i,r_i),(s_j,r_j)} = \mathbb{P}\{B_l, (s_j, r_j) \mid B_k, (s_i, r_i)\}$. Since the state of the static node is deterministic, the only admissible transitions are those specified by the state equations (1-4), i.e. between the generic state (s_i, r_i) and the corresponding state (s^*_j, r^*_j) such that $s^*_j = s(t_k)$ and $r^*_j = r(t_k)$. Specifically, the transition probabilities are as follows

$$h^{kl}_{(s_i,r_i),(s^*_j,r^*_j)} = \begin{cases} 1 & \text{if } s^*_j = \text{OFF and } B_l \neq B_{N+1} \\ 0 & \text{if } s^*_j = \text{OFF and } B_l = B_{N+1} \\ p(t_k) & \text{if } s^*_j = \text{ON and } B_l \neq B_{N+1} \\ 1 - p(t_k) & \text{if } s^*_j = \text{ON and } B_l = B_{N+1} \end{cases}$$

The above probabilities can be justified as follows, assuming to be in the state B_k.

- If the radio will be OFF during the next beacon transmission (at time t_k), then the static node will miss the beacon for sure, so that it can only enter a state different from B_{N+1} (i.e. $B_l = B_{k+1}$).
- Otherwise, the static node will be active during the next beacon transmission time. The static node will miss the beacon with a probability $p(t_k)$, thus moving to the state $B_l = B_{k+1}$. Conversely, it will correctly receive the beacon with a probability $1 - p(t_k)$ thus entering the state $B_l = B_{N+1}$.

Let $\mathbf{X}^{(0)}$ be the initial state probability vector of the static node and $\mathbf{X}^{(k)}$ the state probability vector associated to the time of the k-th beacon transmission, with $k \in [1, N - 1]$,

$$\mathbf{X}^{(k)} = \left(X^{(k)}_0 \; X^{(k)}_1 \; \cdots \; X^{(k)}_{N-1} \; X^{(k)}_N \; X^{(k)}_{N+1} \right)$$
$$\mathbf{X}^{(0)} = \left(X^{(0)}_0 \; 0 \; 0 \cdots 0 \; 0 \; 0 \right)$$

where only the $X^{(0)}_0$ component of the initial state vector is not zero, as when the MR enters the contact area the static node is waiting for the first beacon to be sent. By definition of discrete time Markov chain, it follows that

$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} \cdot \mathbf{H} \quad \text{for } k = 0, 1, 2, \ldots, N - 1 \tag{5}$$

Note that the X^k_{N+1} component of the state vector represents the cumulative probability of the MR discovery after k beacon transmissions. Hence, the p.m.f. of the discovery time r.v. D, i.e. $d(m, t_0) = \mathbb{P}\{D(t) = m\}$, can be derived as

$$d(m, t_0) = \begin{cases} X^{(0)}_{N+1} & \text{if } m = t_0 \\ X^{(k)}_{N+1} - X^{(k-1)}_{N+1} & \text{if } m = t_k,\ k \in [1, N - 1] \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

To derive the p.m.f. of the discovery time by (5) we need to know the initial state probabilities for $\mathbf{X}^{(0)}$. As both beacon transmissions and activations of the static node are periodic and independent, it is reasonable to assume that the initial radio state and the initial residual time are uniformly distributed along all possible values. Hence, for both radio states and independent from the residual time, the initial probability is $\Delta/(T_{ON}+T_{OFF})$, where Δ is duration of a discretized time slot.

All the above discussion assumes a certain initial beacon transmission time t_0. To properly characterize the discovery time, the Equation (6) must be evaluated for all possible values of t_0. Again, as both beacon transmissions and activations of the static node are periodic and independent, we will assume that all possible values of t_0 are uniformly distributed in the range $0 \leq t_0 < T_B$. Note that t_0 has been discretized into $\hat{t}_0 \in \mathcal{T} \equiv \{0, \Delta, \ldots, n_{t_0} \cdot \Delta\}$, where $n_{t_0} = \lfloor T_B/\Delta \rfloor$ is the maximum number of discretized time slots Δ which fit into $[0, T_B)$. Hence, the p.m.f. $d(m)$ of the discovery time per contact is

$$d(m) = \sum_{\hat{t}_0 \in \mathcal{T}} d(m, \hat{t}_0) \cdot \mathbb{P}\left\{\hat{t}_0\right\} = \frac{\Delta}{T_B} \sum_{\hat{t}_0 \in \mathcal{T}} d(m, \hat{t}_0)$$

5 Data transfer phase analysis

In this section we derive the amount of messages correctly transferred by the static node to the MR. Recall that the static node enters the data transfer phase after a successful beacon reception. Since this depends on the discovery time, we will make use of the p.m.f. $d(m)$ obtained in the previous section to derive the number of correctly transferred data messages.

As anticipated in Section 4, while in the data transfer state, the static sensor is always on, and uses an ARQ-based communication protocol for data transfer. In the following, we will assume that both data and acknowledgments messages have a fixed duration T_s, referred to as *message slot*. In addition, we will assume a window size of w messages.

We focus now on a single window starting at the generic time t. As the message loss changes with the distance between the MR and the static sensor, every message will experience its own loss probability. However, we will assume that the message loss is constant during a message slot, i.e. that the i-th message in the window starting at time t will experience a message loss probability $p(t + i \cdot T_s)$. This is reasonable, given the short duration of the message slot. Let's denote by $N(i,t)$ the r.v. denoting the number of messages successfully received by the MR in a given slot i of the window starting at time t. Clearly, the p.m.f. of $N(i,t)$ is $n(i,t,m) = \mathbb{P}\{N(i,t) = m\}$, i.e.

$$n(i,t,m) = \begin{cases} 1 - p(t + i \cdot T_s) & \text{if } m = 1 \\ p(t + i \cdot T_s) & \text{if } m = 0 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Hence, the total number of messages received by the MR within a window, i.e. $N(t)$, is the sum of the $N(i,t)$ r.v.s $N(t) = \sum_{i=0}^{w-1} N(i,t)$, so that its p.m.f. is the convolution of the single p.m.f.s, i.e. $n(t,m) = \otimes_{i=0}^{w-1} n(i,t,m)$.

Furthermore, we denote by $R(t)$ the r.v. representing the number of messages correctly transferred to the MR when an ARQ-based mechanism is used. So, in this case $R(t)$ represents the number of messages acknowledged by the MR. In the following, we will consider a selective retransmission scheme, where acknowledgements notify the sensor node which messages sent in the last window have been correctly received by the MR. Hence, the reception of the acknowledgement has to be accounted as well, so that the messages within a window are correctly transferred if they are successfully received by the MR and the corresponding acknowledgement is not lost. We denote by $A(t)$ the r.v. indicating the number of acknowledgements correctly received by the MR for the corresponding window starting at time t. Hence, the p.m.f. of $A(t)$ is $a(t,m) = \mathbb{P}\{A(t) = m\}$, i.e.

$$a(t,m) = \begin{cases} 1 - p(t + w \cdot T_s) & \text{if } m = 1 \\ p(t + w \cdot T_s) & \text{if } m = 0 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

hence $R(t) = N(t) \cdot A(t)$ and, being them indepedent, we have that

$$\mathbb{E}[R(t)] = \mathbb{E}[N(t)] \cdot \mathbb{E}[A(t)] = \sum_{i=0}^{w-1} [1 - p(t + i \cdot T_s)] \cdot [1 - p(t + w \cdot T_s)]$$

5.1 Joint discovery and data transfer

The above discussion focuses on a single communication window. To get the number of messages transferred during the whole contact time, we have to characterize both the starting time t of the first window and the total number W of windows actually available in the residual contact time. Hence, the total number of messages correctly transferred during a contact is

$$R = \sum_{i=0}^{W} R(t + i \cdot (w+1) \cdot T_s)|_{t=D} \tag{9}$$

If D is the r.v. denoting the discovery time, whose p.m.f. has been derived in the previous section, clearly the start time of the first communication window is $t = D$. In addition, the number of windows in the residual contact time is $W = \lfloor (c_{max} - D)/((w+1) \cdot T_s) \rfloor$, under the assumption that the static node can exploit all the residual contact time for data transfer. We verified by simulation that with an appropriate setting of N_{ack} it is possible to exploit the residual contact time almost completely (see Section 6.2). Thus, the number of messages transferred successfully can be expressed as the r.v. R which depends only on the discovery time D.

Energy efficiency In this section we derive the total energy consumed by the static node per message successfully delivered to the MR, on the average. This metric provides an indication of the energy efficiency for the overall data collection process. The energy consumed in a given radio state is calculated as $P_{rs} \cdot T_{rs}$, where P_{rs} is the power drained in the state rs while T_{rs} is the time spent in the same state. As possible radio states we consider rx for receive, tx for transmit and sl for sleep (i.e. shutdown). In addition, we assume that the power consumption when the radio is idle (i.e. it is monitoring the channel) is the same as in the receive state. As the energy efficiency depends on both discovery and data transfer, we derive first the energy consumption for the discovery phase, and then its joint effect on the subsequent data transfer.

Since the MR arrival may be unknown a priori, the static node may spend a waiting time σ in addition to the discovery time (see Fig. 1(a)). Hence, the energy spent during the discovery phase is

$$\overline{E}_{disc} = P_{sl} \cdot (\sigma + \mathbb{E}[D]) \cdot (1 - \delta) + P_{rx} \cdot (\sigma + \mathbb{E}[D]) \cdot \delta$$

where the first term accounts for the energy spent in the sleep state, while the second one accounts for the time spent in the active state before the correct reception of the beacon message.

On the other side, the energy spent for data transfer – under the assumptions that the static sensor has always data to send and data transfer takes the entire residual contact time – is

$$\overline{E}_{dt} = \left(\frac{\mathbb{E}[c_{max} - D]}{w + 1} + \mathbb{P}\{D\} \cdot \frac{N_{ack}}{2} \cdot T_s \right) \cdot (w \cdot P_{tx} + P_{rx})$$

The first part of the equation denotes the number of windows in the contact time plus the windows wasted after the end of the contact time. Note that the wasted windows have to be considered only when the contact actually occurs, so that the related term is multiplied by the probability of correct detection of the MR (i.e. $\mathbb{P}\{D\} = X_N^{(N-1)}$). The second term, instead, denotes the amount of power spent for a single window in the receive and transmit states, respectively.

Finally, the average total energy consumed by the static sensor per each message correctly transferred to the MR is $\overline{E}_{msg} = (\overline{E}_{disc} + \overline{E}_{dt})/\mathbb{E}[R]$.

6 Results

In this section we will use the analytical formulas derived in the previous sections to perform an integrated performance analysis of the overall data collection process. To this end, we will consider the following performance metrics.

- *Residual contact ratio*, defined as the ratio between the average residual contact time and the contact time $\eta = \mathbb{E}[(c_{max} - D)/c_{max}]$.
- *Contact miss ratio*, defined as the fraction of MR passages not detected by the static sensor (i.e. $\mathbb{P}\{N\} = X_{N+1}^{(N-1)}$).
- *Throughput*, defined as the average number of messages (or bytes) correctly transferred to the MR at each contact (i.e. $\mathbb{E}[R]$).
- *Energy consumption per byte*, defined as the mean energy spent by the static sensor per each message (or byte) correctly transferred to the MR (i.e. $\overline{E}_{msg}$ as defined in Subsection 5.1).

In our analysis we used a message loss function derived from the experimental data presented in [13] and measured in the same scenario introduced in Section

Table 1. Interpolated message loss coefficients as functions of the MR speed (D_y = 15m (a) and other parameters used for analysis (b)

(a)

Coefficient	$v = 3.6$ km/h	$v = 40$ km/h
a_0	0.133	0.4492
$a_1(\mathrm{m}^{-1})$	0	0
$a_2(\mathrm{m}^{-2})$	0.000138	0.0077

(b)

Parameter	**Value**
Transmit power (0 dBm)	49.5 mW
Receive (idle) power	28.8 mW
Sleep power	0.6 μW
Message payload size	24 bytes
Message slot size	15 ms
T_B	100 ms
T_{BD}	9.3 ms

3. To get a more flexible model, we considered an interpolated polynomial packet loss function in the form

$$p(t) = a_2 \cdot \left(t - \frac{c_{max}}{2}\right)^2 + a_1 \cdot \left(t - \frac{c_{max}}{2}\right) + a_0 \tag{10}$$

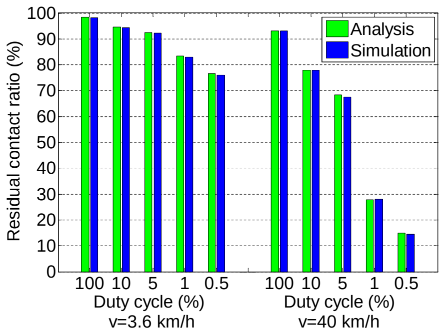
Equation (10) holds only within the contact area, i.e. for $0 < t < c_{max}$. For other values of t, $p(t)$ is assumed to be equal to one, as outside of the contact area any transmitted message is lost. To derive the coefficients in (10) – reported in Table 1(a) for different MR speeds v and for a vertical distance $D_y = 15$ m – we used the same methodology described in [12].

We evaluated the model derived in Sections 4 and 5 and validated the analytical results with a discrete event simulator written in C. In the following, we will show both analytical and simulation results. However, unless stated otherwise, we will refer to the analytical results. Table 1(b) shows the parameter settings for both analysis and simulation.

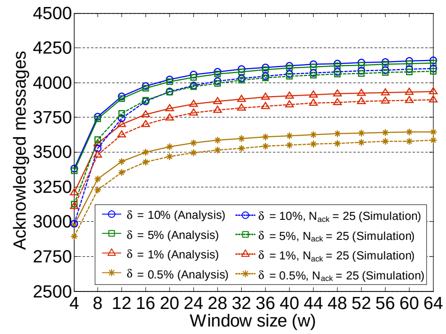
6.1 Discovery phase

In this section we evaluate the performance of the discovery protocol, in terms of residual contact ratio and missed contacts. First of all, the effectiveness of the discovery protocol strictly depends on the T_B and T_{BD} parameters. As a design criterion, both parameters should be reduced as much as possible. In fact, a low T_B increases the frequency of beacon transmission, hence the probability of a timely beacon discovery. On the other hand, a low T_{BD} reduces the overhead due to the beacon reception, so that the residual contact time is not significantly decreased. The proper setting of both parameters strictly depends on the actual system on which the data collection protocols are implemented. In the following discussion we set $T_B = 100$ ms and $T_{BD} = 9.3$ ms, which have been found as suitable values for a Mote class sensor platform [14].

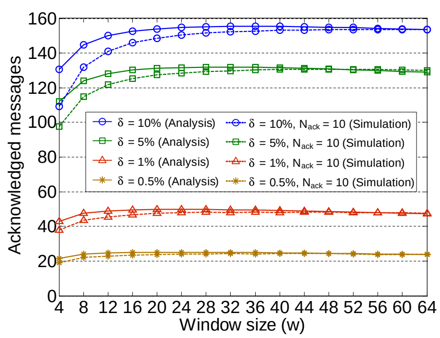
Figure 3(a) shows the residual contact ratio as a function of the MR speed, for different duty cycles. We start considering the MR moving at 3.6 km/h. The results clearly show that the duty cycle used for discovery does not significantly impact the residual contact ratio. For comparison purposes, we have also included the scenario in which the static sensor is always active during discovery.



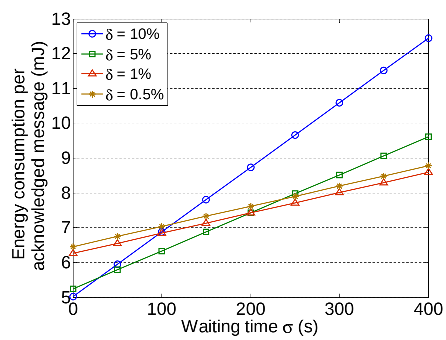
(a) Residual contact ratio



(b) Throughput as a function of the duty cycle for $v = 3.6$ km/h



(c) Throughput as a function of the duty cycle for $v = 40$ km/h



(d) Energy efficiency as a function of the waiting time for $v = 40$ km/h

Fig. 3. Analysis vs simulation results

In this case, the residual contact ratio is very close to the maximum achievable value. Actually, the static sensor can exploit most of the contact time also when it uses a moderate duty cycle. Also lower duty cycles can get satisfactory residual contact ratios, over the 75% for duty cycles of 1% and 0.5%, respectively. These results clearly depend on the average contact time – in this case equal to 158 s – which is much greater than the sleep time of the static node, so that the contact miss ratio is almost zero (i.e. 0.04%) even at a 0.5% duty cycle.

When the MR moves at 40 km/h we can see that the static node can still use a large part of the contact time – which is approximately 17 s – when it is always on during discovery, as the obtained residual contact ratio is above the 90%. However, the residual contact ratio drops when an even moderate duty cycle is used. For lower duty cycles (i.e. 1% and 0.5%), the residual contact ratio decreases significantly. In these cases, in fact, not only the residual contact ratio is much shorter, but there is also a high contact miss probability. This is because the 1% and the 0.5% duty cycles have a sleep time comparable to the contact time. As a consequence, the chance that the static node does not detect the passage of the MR at all is much higher than in the other cases.

6.2 Data transfer

In this section we evaluate the performance of of an ARQ-based transfer protocol in terms of throughput. It is worth recalling that the following results have been obtained for the actual residual contact time, so that they account for the effects of the discovery phase as well.

Figure 3(b) and Figure 3(c) show the throughput obtained by analysis and simulation when the MR moves at different speeds. Note that analysis assumes that the data transfer phase takes all the residual contact time, while simulation uses the loss of N_{ack} consecutive acknowedgements as end-of-contact indication. This is the reason for the slight difference between analytical and simulation results. Actually, this difference is higher for low values of the window size. This happens because the time after which the sensor assumes the MR as out of reach is $(w + 1) \cdot N_{ack} \cdot T_s$, so that it increases with the window size when N_{ack} is constant. When the window size is small, the actual value of N_{ack} should be increased in order to keep $N_{ack} \cdot (w + 1)$ of the same magnitude. Anyway, the value of N_{ack} should be tailored to the target scenario, which depends on the speed of the MR. For instance, in our simulation we used $N_{ack} = 25$ and $N_{ack} = 10$ for the 3.6 km/h and the 40 km/h scenarios, respectively.

We start considering the throughput in terms of messages acknowledged by the MR. Figure 3(b) shows the throughput as a function of the window size when $v = 3.6$ km/h. We can see that the throughput increases with the window size for all considered duty cycles, so that the maximum is obtained with the largest window size of 64 messages. This is due to the acknowledgement overhead, which decreases as the window size grows up. Specifically, the throughput reaches over 4000 messages per contact, corresponding to about 100 kB of data, for the 10% duty cycle. Similar results are also obtained with the lower 5% and 1% duty cycles. Such results can be explained on the basis of the residual contact ratios, which are similar for the different duty cycles, i.e. the residual contact time is not reduced significantly when a low duty cycle is used. The same is not true for the 0.5% duty cycle, which actually experiences a lower, but still reasonable, throughput of about 3500 messages per contact (nearly 88 kB).

Figure 3(c) shows the throughput as a function of the window size when $v = 40$ km/h. The throughput has a different trend in this case. It first increases when the window size is low, then decreases after a point which depends on the duty cycle. In addition, the obtained throughput changes significantly with the duty cycle. In fact, while the 10% and the 5% duty cycles both achieve a similar throughput over 100 messages per contact (around 3 kB), the 1% duty cycle gets only 50 messages (1.2 kB) per contact. The lowest 0.5% duty cycle even obtains a throughput of 25 messages per contact (0.6 kB), which is rather low, but may be enough for certain applications.

6.3 Energy efficiency

In this section we evaluate the energy efficiency of data collection. It should be noted that the considered energy consumption accounts for both discovery and data transfer, so that it fully characterizes the overall data collection process. While in the previous sections we have considered only what happens within the contact area, in the following we take into account also the time spent by the static sensor on waiting for the MR to enter the communication range. To this end, we measured the average energy spent per acknowledged message as a function of the waiting time.

Figure 3(d) shows the average energy consumption per acknowledged message when the MR moves at $v = 40$ km/h and the window size is 32. Clearly the energy consumption increases with the waiting time, but a very low duty cycle is not necessarily the most convenient option. In fact, when the average waiting time is below 25 s (i.e. the MR arrival can be predicted with rather good accuracy), the best option is the 10% duty cycle. Instead, when the average waiting time is between 25 s and 200 s, the most convenient duty cycle is 5%. From later on, i.e. when the MR presence is very difficult to estimate, the best duty cycle is 1%. In addition, the 0.5% duty cycle always gets a higher average energy consumption than the 1% duty cycle. These results are in contrast to the expected behavior, i.e. that the energy consumption decreases with the duty cycle, as it happens in the scenario where the MR moves at 3.6 km/h (we have omitted the correspondent figure for the sake of space).

Actually, these results can be explained as follows. Low duty cycles may delay the MR detection, leading to lower residual contact times. In addition, they may also produce high contact miss ratios, so that the energy spent during discovery is simply wasted, as the sensor node cannot transmit any data. In the considered scenario, where the MR moves with $v = 40$ km/h, the contact time and the residual contact time are short. When the waiting time is low, i.e. when the sensor knows the MR arrival times with a good accuracy, the energy overhead due to discovery is negligible, because the sensor spends most of its active time during data transfer. Instead, the advantages of low duty cycles become relevant when the waiting time is high. In this case, the sensor may spend a significant amount of time by looking for beacons when the MR is out of the contact area. Thus, a low duty cycle reduces the activity of the sensor during the waiting time, which is the highest share of the overall energy consumption. Regarding the 0.5% duty cycle, it is always unsuitable in this scenario, because the energy gain due to the lower activity time is thwarted by the decrease in the throughput (see Fig. 3(c)).

7 Conclusions

In this paper we have developed an analytical model of the overall data collection process in sparse sensor networks with mobile relays (MRs). The model is flexible enough to incorporate different discovery and data transfer protocols. We limited our discussion to a simple discovery algorithm where the MR sends periodic advertisements and the sensors follow an asynchronous scheme based on a low duty cycle. In addition, we considered an ARQ communication protocol with selective retransmission for data transfer. Our findings show that low duty cycles can be actually used for a large class of environmental monitoring applications. Surprisingly, a low duty cycle may not always be the most energy efficient option, depending on a number of different factors such as the speed and the mobility of the MR.

This work could be improved along different directions. First, the model proposed in this paper could be extended to the case of multiple MRs. Second, different discovery and data transfer schemes could be considered. Finally, the findings of our analysis could be used as a basis for the definition of adaptive data collection protocols, which are capable to tailor the operating parameters to the actual conditions (i.e. knowledge of the MR arrivals, buffer constraints, residual energy of sensor nodes etc.). We are currently evaluating these extensions as a future work.

Acknowledgements

Work funded partially by the European Commission under the FP6-2005-NEST-PATH MEMORY project, and partially by the Italian Ministry for Education and Scientific Research (MIUR) under the FIRB ArtDeco project. The authors would like to thank Emmanuele Monaldi, Francesca Mancini and Paolo Caggiari for their help.

References

1. Jain, S., Shah, R., Brunette, W., Borriello, G., Roy, S.: Exploiting mobility for energy efficient data collection in wireless sensor networks. ACM/Springer Mobile Networks and Applications **11**(3) (June 2006) 327–339
2. Chakrabarti, A., Sabharwal, A., Aazhang, B.: Using predictable observer mobility for power efficient design of sensor networks. In: Proc. IPSN 2003. (2003)
3. Shah, R.C., Roy, S., Jain, S., Brunette, W.: Data mules: Modeling a three-tier architecture for sparse sensor networks. In: Proc. IEEE SNPA 2003. (2003)
4. Jea, D., Somasundra, A., Srivastava, M.: Multiple controlled mobile elements (data mules) for data collection in sensor networks. In: Proc. IEEE DCOSS 2005. (2005)
5. Kansal, A., Somasundara, A., Jea, D., Srivastava, M., Estrin, D.: Intelligent fluid infrastructure for embedded networks. In: Proc. ACM Mobisys 2004. (2004)
6. Anastasi, G., Conti, M., Di Francesco, M., Passarella, A.: Energy conservation in wireless sensor networks: A survey. Ad Hoc Networks (in press)
7. Zhao, W., Ammar, M.: Message ferrying: Proactive routing in highly-partitioned wireless ad hoc networks. In: Proc. IEEE FTDCS 2003. (May 2005)
8. Jun, H., Ammar, M., Zegura, E.: Power management in delay tolerant networks: A framework and knowledge-based mechanisms. In: Proc. IEEE SeCon 2005. (2005)
9. Gu, Y., Bozdag, D., Ekici, E.: Mobile element based differentiated message delivery in wireless sensor networks. In: Proc. IEEE WoWMoM 2006. (2006)
10. Gu, Y., Bozdag, D., Ekici, E., Ozguner, F., Lee, C.: Partitioning based mobile element scheduling in wireless sensor networks. In: Proc. IEEE SECON 2005. (2005) 386–395
11. Somasundara, A., Kansal, A., Jea, D., Estrin, D., Srivastava, M.: Controllably mobile infrastructure for low energy embedded networks. IEEE Transactions on Mobile Computing **5**(8) (2006)
12. Anastasi, G., Conti, M., Monaldi, E., Passarella, A.: An adaptive data-transfer protocol for sensor networks with Data Mules. In: Proc. of IEEE WoWMoM 2007. (2007)
13. Anastasi, G., Conti, M., Gregori, E., Spagoni, C., Valente, G.: Motes sensor networks in dynamic scenarios. International Journal of Ubiquitous Computing and Intelligence **1**(1) (April 2007)
14. Anastasi, G., Conti, M., Di Francesco, M.: Data collection in sensor networks with Data Mules: an integrated simulation analysis. In: Proc. of IEEE ISCC 2008. (2008)