Experimental and Analytical Performance Evaluation of a Real Opportunistic Wireless Sensor Network

Jorge M. SOARES[1], Wansheng ZHANG[2], Mirko FRANCESCHINIS[3], Maurizio A. SPIRITO[3], Rui M. ROCHA[1]
[1]*Instituto de Telecomunicações, Instituto Superior Técnico, Technical University of Lisbon*
Av. Prof. Dr. Cavaco Silva, 2744-016 Porto Salvo, Portugal
Tel: +351 214233505, Fax: +351 214233252, Email: {jorgesoares,rui.rocha}@ist.utl.pt
[2]*Politecnico di Torino, Dipartimento di Elettronica*
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
Tel: +39 011 5644001, Fax: +39 011 5644099, Email: wansheng.zhang@polito.it
[3]*Istituto Superiore Mario Boella (ISMB), Pervasive Radio Technologies (PeRT) Lab*
Via Pier Carlo Boggio 61, 10138 Torino, Italy
Tel: +39 011 2276201, Fax: +39 011 2276299, Email: {franceschinis,spirito}@ismb.it

Abstract: The use of opportunistic communications, while increasing in popularity, is still limited, due in part to some uncertainty that still exists regarding its performance in real-world conditions. This paper tries to assess the real performance of an opportunistic routing implementation in a physical setting, by comparing it with its expected performance, determined by a simplified theoretical model. For that purpose, we have deployed an experimental testbed combining static and mobile sensor nodes, and running two different applications in two different platforms in tandem. This allows us to obtain real contact traces from an unmodified application, while at the same time logging the messages transferred between participating nodes. The data collected was later analysed, mainly in what refers to the intra-contact and global communication patterns, as well as the end-to-end delay distributions for each sending node's messages. The results obtained show that the system behaviour can be predicted with high accuracy by our simple model.

Keywords: opportunistic communications, WSN testbeds, contact characterisation, routing, channel quality

1. Introduction

Opportunistic communications are gaining increasing acceptance as a solution to the problem of data collection in wireless sensor networks (WSNs) [1]. The use of short-range, low-power communication devices and the exploitation of node mobility allow the exploitation of network nodes as message carriers, expanding the basic operational range without requiring the use of longer-range radio technologies such as WiMAX, GSM, etc.

The issue of realistic characterisation of contact opportunities among nodes is of critical importance for the use of opportunistic communications, but also of difficult assessment, as reliable characterisation requires physical world experimentation, in addition to simulation.

This paper builds on previous work developed at IST and ISMB in the field of WSNs and opportunistic communications [2]. Taking advantage of our existing applications and hardware, we have designed an experiment that allows us to assess the real-world performance of an opportunistic routing implementation versus its predicted behaviour in the same scenario. This required setting up an experimental testbed, including mobile

elements, that features nodes running each application side-by-side. One of the applications provides lower-level channel quality information (i.e., related to PHY and MAC layers), while the other performs opportunistic data collection. The information logged allows us to derive quantitative metrics describing contact opportunities among nodes and to use them as inputs to calculate the ideal system performance for the scenario at hand; specifically, it allows us to know when testbed elements are in range, even if the opportunistic collection nodes have their radios turned off, thereby providing us with much more accurate contact characterization. Next, we compare it with the performance of the real system in the same exact conditions, which is affected by additional factors.

The remainder of this paper is organised as follows: in Section 2 we introduce the applications used; in Section 3 we present the details of our experimental testbed; in Section 4 we discuss our approach to the analysis of the data collected; in Section 5 we present the results obtained; finally, in Section 6, we extract some conclusions and propose further work.

2. Applications used

The experiments are carried out using two different applications running on two different platforms. The first application is an opportunistic routing solution and is implemented on Sun Microsystems' Small Programmable Object Technology (SPOT) nodes [3]. The second application is used as a real-time channel quality monitoring tool and flashed on Telos nodes [4]. Different nodes were used for each application in an effort to minimize cross-application interference. Furthermore, there are specific behaviours regarding CHARON's control of the radio power that wouldn't be possible when other applications are using it.

2.1 – Sun SPOT application

CHARON is a simple yet efficient solution to the problem of routing messages in sparse mobile WSNs. It aims to minimise the number of message exchanges, while still providing a way for urgent messages to be delivered quickly. It also integrates features such as time synchronisation and radio power management, that are seldom found but of critical importance in the achievement of energy efficiency. It uses a history-based approach for *convergecast* routing and is designed for sparse networks with slow mobility. Consequently, it does not support advanced mechanisms for quick contact detection. For a detailed description of CHARON, we refer the reader to [5].

Two types of nodes are considered: base stations and ordinary nodes. An ordinary node is not only a potential source of data messages, but can also carry those originated by others. Data messages are always destined to a base station, which acts as a sink and is connected to a computer. While a base station is logically unique, multiple physical instances of it can coexist.

In order to be detected and propagate routing information, nodes periodically broadcast beacon messages, also used for synchronisation. When a node receives a beacon broadcast by a better carrier than itself, it transfers all currently held messages in sequence, up to the destination's available buffer capacity. The system has no bundle-layer ACKs, relying instead on MAC-level ones. When a transfer fails, the message is returned to the buffer.

The system uses synchronous power management to extend node lifetime. There is a global, low-precision time reference, which is then used to trigger simultaneous rounds on all nodes. Each round alternates activity and sleep phases, with the round period and the activity time both being customizable. The system uses slow cycling, with each phase lasting several seconds, although the optimal timings are scenario-specific, depending on the radio range and movement speed.

Messages are time stamped at the source, using the global reference, and again on delivery to the base station, allowing us to calculate the end-to-end latency. While in intermediate buffers, messages are kept sorted by generation timestamp, regardless of their origin, guaranteeing that the absolute oldest ones are forwarded first.

2.2 – Telos application

This application allows us to monitor the radio channel quality in real-time. It works with only two nodes, source and destination. The source sends the destination a message every T_{send} ms (set as desired, but with recommended values in the order of tenths of milliseconds, comparable to typical round trip times). A sequence number included in the message header provides the distinction between different messages. The destination only accepts in-sequence messages but sends back an ACK after any message reception.

Part of the message content is stored on a PC: the sequence number, source id, source timestamp and retransmissions count fields. The source timestamp refers to the first message transmission and it is not overwritten in case of subsequent retransmissions. The message retransmissions count, initialised to zero and incremented one by one, can be helpful in deriving information on current radio channel quality.

The destination timestamp and RSSI/LQI values are retrieved locally at the destination and passed on to the PC. As source and destination are not synchronised, the difference between their timestamps only provides a relative indication of the delivery delay. Additional analysis of channel quality can be carried out using the RSSI and LQI values, calculated by the destination on a single message basis.

3. Testbed and scenario

Joint experiments involving both Telos and SPOT nodes, running the previously described applications, have been performed within PERT Lab in ISMB premises. A schematic visualisation of testbed and scenario is provided in Figure 1, where solid and dashed circles respectively represent Telos and SPOT nodes.

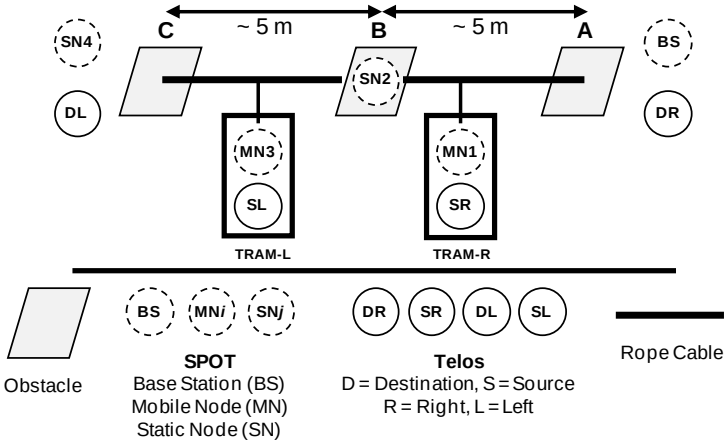

Figure 1: Testbed plan

Two trams move along a cable mounted above the lab. The two trams consist of LEGO Mindstorms NXT robots, each equipped with two touch sensors, one in the front and one in the rear. Tram-R moves between positions A and B, where it stops after a touch sensor hits an obstacle. Similarly, Tram-L moves between positions B and C, where another obstacle is positioned. Both trams are programmed to move in one direction, wait for a few seconds when a barrier is touched, then resume movement in the opposite direction. The trams' speed is assumed to be constant and equal. This is only an approximation, for two reasons:

the slow decrease in movement speed due to the decay in battery level, and the slightly different capacity of each tram's battery. Still, in order to desynchronise the movement of the two trams, the waiting period was set differently for each: $T_{stop\text{-}R}$ = 5 seconds for Tram-R and $T_{stop\text{-}L}$ = 3 seconds for Tram-L.

As shown in the figure, each tram carries two nodes, one SPOT (MN1 and MN3, respectively) and one Telos (SR and SL, respectively), which are thus mobile nodes. The remaining nodes deployed in the scenario are static: two SPOTs (BS, SN4) and two Telos (DR, DL) are located at the ends of the cable, and one SPOT (SN2) is placed approximately in the middle. The physical distance between static nodes of the same kind ensures that they are not able to directly communicate using the carefully tuned transmission power.

The five SPOTs - one base station (BS) and four ordinary nodes (two static, SN2 and SN4, and two mobile, MN1 and MN3) - run the CHARON opportunistic network application and use channel 25 in the 2.4GHz ISM band. Beacons are exchanged every 500 milliseconds during activity phases, whose duration is about 2 seconds. The round period is set to R = 10 seconds. Messages follow a linear path ending in BS, corresponding to left-to-right forwarding in Figure 1. The four Telos are organised in two pairs, the one on the right side tuned on channel 26 and the one on the left side on channel 24, in order to prevent interference between different sets of nodes. The periodicity of message transmissions T_{send} was set to 50 ms.

4. Analysis

Concept. The core idea, and starting point, for our analysis is founded on two key-points. First, the Telos application makes available real-time low-layer information on network channel conditions thanks to the fast dynamics defined by the small time constant T_{send}. Second, the information provided by the Telos application is helpful for the interpretation of the upper-layer SPOT results on CHARON opportunistic routing. Clearly, in order to incorporate the data collected by the two different applications on distinct computers, it is necessary that all events are time stamped using a common time reference. This is done by previously synchronizing, using NTP, all computers responsible for packet logging.

Goal. We build a model able to predict the probability distribution functions (pdf) $D_{BS\text{-}MN1}$, $D_{BS\text{-}SN2}$, $D_{BS\text{-}MN3}$ and $D_{BS\text{-}SN4}$ of delivery delay suffered by messages generated by each SPOT node. We then compare the theoretical pdf with the experimental pdf derived from SPOT logs. To derive the pdf, the model takes as input four quantities, $T_{trip\text{-}R}$, $T_{trip\text{-}L}$, $T_{c\text{-}R}$ and $T_{c\text{-}L}$, which are estimated using Telos and SPOT logs. The model, along with the related assumptions, is described step-by-step in this section.

Model inputs. The twin quantities $T_{trip\text{-}R}$ and $T_{trip\text{-}L}$, measured in seconds, represent the time taken respectively by Tram-R and Tram-L to complete a trip. We consider them constant, even though this is only approximately true because, as mentioned in Section 3, speed is not entirely constant. We are able to empirically calculate the duration of a trip, thus obtaining a sample per trip, by processing data from Telos logs. To this aim, it is enough to identify comparable events that occur regularly at each trip. The value of $T_{trip\text{-}R}$ and $T_{trip\text{-}L}$ is then determined as the average of the samples collected.

The contact time $T_{c\text{-}R}$ (respectively, $T_{c\text{-}L}$) is defined as the time, measured in seconds, spent by MN1 (respectively, MN3) within the radio range of BS and SN2 (respectively, SN2 and SN4), i.e. the fraction of $T_{trip\text{-}R}$ (respectively, $T_{trip\text{-}L}$) during which these pairs of nodes are able to communicate. Here, we are implicitly assuming that all the SPOT nodes have the same radio characteristics. Similarly to $T_{trip\text{-}R}$ and $T_{trip\text{-}L}$, we manage $T_{c\text{-}R}$ and $T_{c\text{-}L}$ as constant quantities, which means that we are considering a time-invariant radio propagation model that only depends on the distance between transmitter and receiver. Thus the following relationship holds: $T_{c\text{-}R} - T_{stop\text{-}R} = T_{c\text{-}L} - T_{stop\text{-}L}$. If $T_{c\text{-}R}$ is known, $T_{c\text{-}L}$ can be derived immediately. In this case, we have to use SPOT logs to estimate $T_{c\text{-}R}$. Indeed, despite Telos

and SPOT nodes using the same radio chip (the CC2420) and having been tuned for the same transmission range, it is never exactly the same. The radio power cycling on SPOT nodes complicates the analysis: in general, MN1 could enter/exit from the coverage area of BS while in sleep mode. It is nevertheless possible to estimate $T_{c\text{-}R}$ by taking a weighted average of the number of interactions between MN1 and BS during a trip and assuming that the times MN1 spends in sleep mode within BS's radio coverage area before the first interaction and after the last one are uniformly distributed between 0 and R.

Model derivation. To derive the theoretical pdf, we make the following simplifying and optimistic assumption: a SPOT node always delivers in a single interaction all of the messages it is carrying - in Section 5 we will see that this is not always the case. In addition, we neglect the message transfer time and the impact of other operations executed during the activity phase. As such, message transfers are instantaneous and end-to-end delays are exact multiples of R. Finally, we consider no limits for storing space, so that messages are never suppressed, and message losses due to failed communication are ignored too.

Determining $D_{BS\text{-}MN1}$ is quite simple since messages generated by MN1 never pass through intermediate carriers. The delivery delay of such a message can be computed deterministically if the position of MN1 when the round starts (and the message is prepared) is known. The position of MN1/Tram-R when randomly sampling the system can be probabilistically determined based on the knowledge of Tram-R mobility pattern. It results in the superposition (i.e., the sum) of a uniform random variable, due to the tram's constant speed, with two Dirac deltas in positions A and B, representing the waiting periods. The deliver delay is null if MN1 is already in contact with BS when the round starts. The farther the tram from entering the contact area when the message is generated, the larger the delay. This follows an uniform distribution except that for the maximum delay which occurs if MN1 has just left BS's contact area when the round starts. Exact characterisation of $D_{BS\text{-}MN1}$ is detailed in Table 1.

Table 1: Characterisation of delivery delays for nodes MN1 and SN2. K is the largest integer such that the $(T_{c\text{-}R}+K\cdot R)/T_{trip\text{-}R} \leq 1$, $K_1 = |(T_{trip\text{-}R}-2\cdot T_{c\text{-}R})/2|^+/R$ *and* $K_2 + 1 = |T_{trip\text{-}R}+(T_{trip\text{-}R}-2\cdot T_{c\text{-}R})/2|^+/R$, *with* $|X|^+$ *representing the smallest integer larger than X.*

$D_{BS\text{-}MN1}$		$D_{BS\text{-}SN2}$	
Delay	Probability	Delay	Probability
0	$T_{c\text{-}R}/T_{trip\text{-}R}$	$K_1\cdot R$	$[K_1\cdot R-(T_{trip\text{-}R}-2\cdot T_{c\text{-}R})/2]/T_{trip\text{-}R}$
$i\cdot R$	$R/T_{trip\text{-}R}$, $i = 1, \ldots, K$	$i\cdot R$	$R/T_{trip\text{-}R}, i = K_1+1, \ldots, K_2$
$(K+1)\cdot R$	$1 - (K\cdot R+T_{c\text{-}R})/T_{trip\text{-}R}$	$(K_2+1)\cdot R$	$1 - [K_2\cdot R-(T_{trip\text{-}R}-2\cdot T_{c\text{-}R})/2]/T_{trip\text{-}R}$

Since messages generated by SN2 are always routed through MN1. Tram-R is the only mobile element involved and the reasoning behind $D_{BS\text{-}SN2}$ characterisation is based on the same concepts as $D_{BS\text{-}MN1}$. The minimum delay is deduced by considering the time Tram-R takes to move between the borders of SN2 and BS radio coverage areas, that is $(T_{trip\text{-}R} - 2\cdot T_{c\text{-}R})/2$. The maximum delay is obtained by considering an additional trip with respect to the minimum delay. Intermediate delays, multiples of R comprised between the minimum and the maximum, occur with uniform probability. Table 1 includes details about $D_{BS\text{-}SN2}$.

In order to characterise $D_{BS\text{-}MN3}$ and $D_{BS\text{-}SN4}$, we preliminarily observe that the position of Tram-R at a random sampling time t can be considered statistically independent from the position of Tram-L at the same time t, regardless of the initial positions, when considering t over an infinite time period. This comes from differently setting $T_{stop\text{-}R}$ and $T_{stop\text{-}L}$ and from having comparable, but not identical, distances AB and BC. The statistical independence is the fundamental assumption to claim that $D_{BS\text{-}MN3} = D_{BS\text{-}SN2} * D_{SN2\text{-}MN3}$ and $D_{BS\text{-}SN4} = D_{BS\text{-}SN2} * D_{SN2\text{-}SN4}$, where "*" denotes a convolution operation and where, implicitly, messages from SN4 and MN3 are assumed to be always routed through SN2. While messages can jump

from MN3 directly to MN1 if the two trams get close enough, the model manages such events as if the message passes transparently through SN2 without additional delay. Finally, it is worth noting that $D_{SN2\text{-}MN3}$ and $D_{SN2\text{-}SN4}$ coincide respectively with $D_{BS\text{-}MN1}$ and $D_{BS\text{-}SN2}$, as reported in Table 1, except that $T_{c\text{-}R}$ and $T_{trip\text{-}R}$ must be replaced with $T_{c\text{-}L}$ and $T_{trip\text{-}L}$.

5. Results

The testbed described in Section 3 was left running for about one hour and a half, a test duration which represents a good trade-off between the need to collect a sufficient amount of data for reliable statistical analysis, and the need to keep the trams movement almost unchanged during the test.

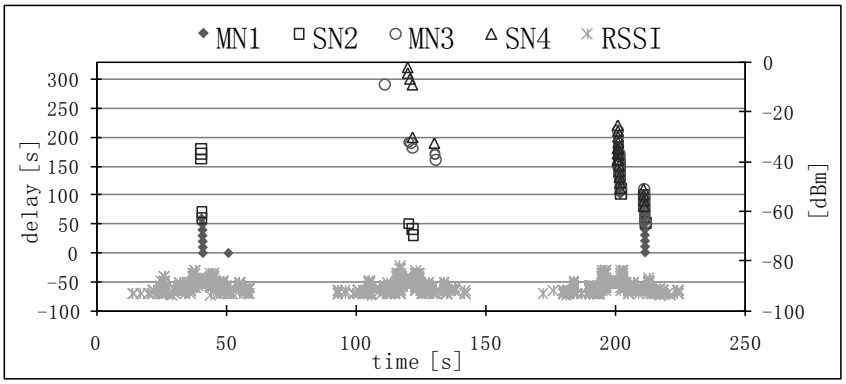

Figure 2: Detail of system events at BS and DR, extracted from the test log, showing delivery delay of messages generated by SPOT nodes and RSSI of messages received by DR.

Figure 2 refers to a four-minute fraction of the test and shows three complete tram trips. The correlation of the message reception events at DR and BS is quite clearly seen. Indeed, delivery of SPOT messages to the BS always occurs during times in which SR and DR are able to persistently communicate and the channel quality seems the best, in terms of RSSI values. Observing the time when the first SPOT messages are transferred in each trip, Figure 2 also suggests that Telos radio coverage area was wider than SPOT's during the test. Moreover, the fact that MN1 always delivers messages in two or three interactions during a trip means that $T_{c\text{-}R}$ was larger than 20 and smaller than 30 seconds. In fact, from SPOT logs we calculated $T_{c\text{-}R}$ = 23.7 seconds, while using Telos logs the average value of $T_{trip\text{-}R}$ and $T_{trip\text{-}L}$ resulted in 78.4 and 67.4 seconds, respectively. Finally, Figure 2 provides additional evidence that, contrary to what is assumed in our model, MN1 is not able to deliver all the messages in only one interaction and, sometimes, the contact ends with MN1 still carrying some messages that will need another trip to be transferred (as happens during the second trip in the plot, when no MN1-generated messages are delivered).

Each of the next four figures, respectively concerning MN1, SN2, MN3 and SN4, shows two plots side-by-side, both representing the pdf of SPOT messages delivery delays. The one on the right is generated from the theoretical model, while the one on the left results from experimental data. Indeed, the model is validated by our experimental data, and the most obvious evidence is that the pdf shapes resulting from the model analysis qualitatively mimic the pdf shapes obtained from the correspondent experimental samples.

In case of MN1, several expected behaviours can be observed, as well as a few minor discrepancies. The null delay is the minimum and also the most frequent, even if the corresponding probabilities are quite different. The rest of the pdf is approximately uniform, with probability equal to about 0.12, in the intervals [10-60] for experimental results and [10-50] for theoretical analysis. Some messages experience delays longer than 60 seconds, which are not predicted by the model. We deduce that such messages are not delivered during the first available contact between MN1 and BS, in contrast with our

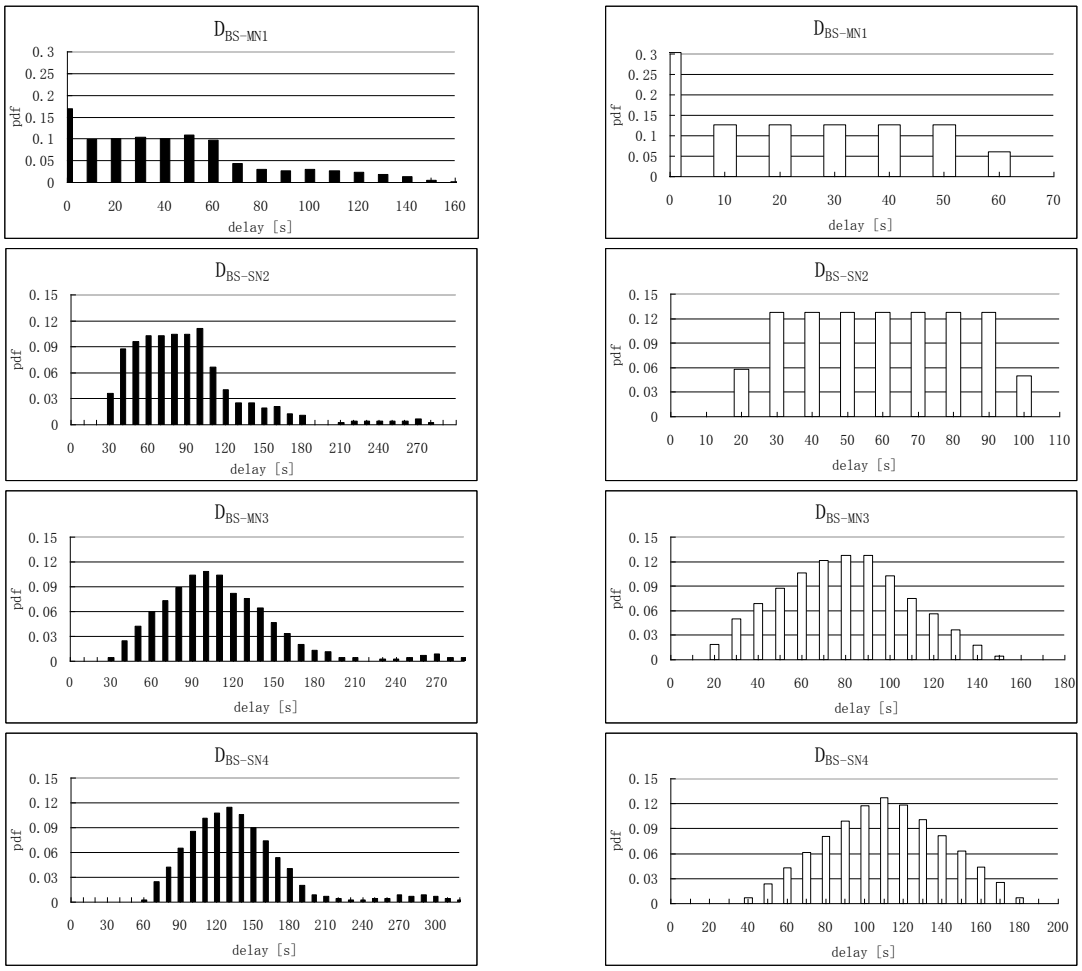

Figure 3: Experimental and theoretical end-to-end delay distributions

assumptions. The impact of such unexpected type of event, which was already underlined previously, can be quantified by the weight of the tail of the distribution. This results in a cumulative probability equal to 0.219 meaning that, on the average, one message over five requires an additional tram trip to be delivered. These discrepancies are nullified by considering such packets as if they were delivered during the first trip. This can be obtained by shifting backwards the tail of the distribution and overlapping it only with the range [0;60] admitted by the model. As expected, in this range the model pdf resembles the modified experimental pdf (not plotted) much better than the original, as quantitatively confirmed by the sum of the least squares which decreases from 0.0224 to 0.0092.

Very similar comments can be reported concerning SN2, whose messages experience minimum delays of 30 seconds, larger than the ones conceived by the model, i.e., 20 seconds. This discrepancy is not surprising since we have already mentioned that the simplifying model assumptions make predictions more optimistic than the real scenario.

The model accuracy in predicting delays is clearly evident for MN3 and SN4. Indeed, the triangular shape predicted by the model as a result of the convolution of (almost) uniform distributions accurately captures the shape of both empirical distributions. We can see that, logically, the difference between minimum delays provided by experimental and analytical curves tends to increase as the distance from BS (in terms of intermediate carriers) increases. The same phenomenon holds for maximum delays. Remembering that messages generated by MN3 and SN4 are routed through SN2, quantitative values are coherent too: minimum experimental delays for MN3 and SN4 are respectively equal to 30 and 60 seconds, while the ones observed for MN1 and SN2 are of 0 and 30 seconds.

To conclude, the weight of the distribution tails heavily decreases with the number of intermediate carriers: 0.219, 0.254, 0.120 and 0.092 for MN1, SN2, MN3 and SN4. This can be explained noting that the tail refers to larger delays that correspond to the simultaneous occurrence of worst-case, and less probable, events along the message path towards BS.

In general, the model captures rather well the system dynamics, and the differences between the expected and observed delays are mostly explained by its simplicity. For instance, we are ignoring the effect of the network load on the delays. While the load is, in fact, quite low, the significant number of messages that can be held in intermediate buffers, leads to many moments in which additional delays are imposed because there is just not enough time to transfer all messages. Another significant error factor is the behaviour of the application itself: we are also not considering possible loss of communication opportunities due to missed beacons, as well as non-deterministic processing delays imposed by background tasks.

6. Conclusions

We have presented a performance assessment of an opportunistic routing solution running in a real-world testbed. By comparing experimental and model results we showed that the model closely matches the working system, generally behaving as expected. At the same time, based on the simplifying model assumptions and on the knowledge of their impact, we were able to explain the minor discrepancies in the system's evolution with respect to our expectations.

As part of our future work, we intend to tackle the following issues:

- The information being collected is still quite limited, thereby limiting our analysis. The inclusion of additional logging capabilities and their extension to all nodes on the network could provide very interesting information.
- Tests have been conducted on a single topology with predictable and periodical mobility. The use of different node arrangements and a scenario with multiple and variable paths would greatly increase the scope of our performance evaluation.
- A more detailed model, dealing not only with the mobility patterns but also with intra-network factors, such as the message load, could provide more accurate predictions.

Acknowledgements

This work was funded by Istituto Superiore Mario Boella, and Instituto de Telecomunicações and Instituto Superior Técnico at the Technical University of Lisbon. It was partly supported by the European Commission through the FP7 Network of Excellence in Wireless Communications NEWCOM++ (contract n. 216715).

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, “A Survey on Sensor Networks”, Published in: IEEE Communications Magazine, vol. 40, no. 8, pp. 102 - 114, August 2002.

[2] L. Pelusi, A. Passarella and M. Conti, "Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad Hoc Networks", Published in: IEEE Communications Magazine, vol.44, no.11, pp.134-141, November 2006.

[3] R. B. Smith, “SPOTWorld and the Sun SPOT”, Published in: Proceedings of the 6th International Conference on Information Processing in Sensor Networks, April 2007.

[4] J. Polastre, R. Szewczyk, and D. Culler, "Telos: Enabling Ultra-Low Power Wireless Research", Published in: Proceedings of the 4th International Symposium on Information Processing in Sensor Networks, April 2004.

[5] J. M. Soares and R. M. Rocha, "CHARON: Routing in Low-Density Opportunistic Wireless Sensor Networks", Published in: Proceedings of the 2nd IFIP Wireless Days Conference, December 2009.