

# Impact of Data Collecting Techniques on the Performance of a Wireless Sensor Network

Mohammad M. Abdellatif <sup>#1</sup>, José Manuel Oliveira <sup>\*2</sup>, Manuel Ricardo <sup>#3</sup>, Peter Steenkiste <sup>†4</sup>

<sup>#</sup> INESC TEC, Faculdade de Engenharia, Universidade do Porto

<sup>\*</sup> INESC TEC, Faculdade de Economia, Universidade do Porto

Rua Dr. Roberto Frias, 378, 4200-465, Porto, Portugal

{<sup>1</sup>mma, <sup>2</sup>jmo, <sup>3</sup>mricardo}@inescporto.pt

<sup>†</sup> Department of Computer Science, Carnegie Mellon University

Pittsburgh, PA, USA

<sup>4</sup> prs@cs.cmu.edu

**Abstract**—Wireless Sensor Networks (WSNs) consist of small devices with processing, communication and sensing capabilities. These devices interact together to carryout monitoring tasks. An example of such network is a photo-voltaic (PV) power plant where each solar panel has a sensor. The number of interconnected solar panels can become very large, and spread over a large area. Each sensor will sense the output of the panel and send this value to a central node for processing. In this paper we evaluate the performance of a wireless sensor network employing three different data collecting techniques. The study considers different networks, each with a different number of nodes and with different values for the offered load, estimating for each network size and offered load, network throughput, packet loss and end-to-end packet delay. Results show that as the size of the network grows and for higher values of the offered load, the best performance is achieved by using a polling based data collecting technique.

## I. INTRODUCTION

Wireless Sensor Networks (WSNs) are made by small devices with processing, communication and sensing capabilities. These devices interact together to carry out monitoring tasks in a geographical area of interest [1]. Applications of WSN include smart-cities, environmental monitoring, distributed sensing in industrial plants, and health care [2]. In WSN networks, the data rate employed is usually low. However, such networks pose challenging issues related to the communication reliability and to the efficient use of node batteries [3]. Additionally, due to the many-to-one feature of the sensor data collection applications, the existence of wireless interferences and collisions, the large sizes of such networks, the scheduling of data packet transmissions becomes a challenging problem that needs to be carefully addressed. In WSNs, an aspect that needs to be considered is the data collection technique, i.e., how the sensors send their data to the sink. Nodes can send the data to the sink as soon as this information becomes available or they can wait to receive a request from the sink asking each node to start sending their data. As the size of the network grows, the selected data collecting technique will have a huge impact on the performance of the network as a whole. This work studies how data collection technique used to gather

information in a well defined scenario affects the performance of a WSN. The communications scenario is provided by the SELF-PVP project [4]. This project aims to increase the efficiency of a photo voltaic (PV) power plant. It assumes that 200,000 solar panels are deployed in a matrix like deployment spread over a large area (250 hectares). The objective is to put the solar plant operating at an optimum power delivery point. Each solar panel will have a sensor/actuator that will sense local variables, communicate these values with other sensors and compute local errors with the goal of optimizing, the overall performance of the panels' array.

A WSN in this scenario can be viewed as a grid of sensors where each sensor is connected to a solar panel. Panels in the same column will have the same value of the current passing through them. In order to reach the optimum power delivery point, most of the communications will be between nodes in the same column. Based on this we divide the network into small networks of columns. Throughout this paper, we focus on the communication within a column.

In this paper, we evaluate the performance of an IEEE 802.15.4 wireless sensor network employing three different data collecting techniques within a column. In Technique 1, client nodes send their data as soon as it is ready. In Technique 2, the sink sends a broadcast poll message to its neighbors, which is propagated through the network triggering each client node to send its data to the sink. Technique 3 is similar to Technique 2 except that the poll message is sent to half of the network first and then the sink waits to receive the data from this half before sending the poll message to the other half.

The study considers different networks, each with a different number of nodes and with different values of the offered load, estimating for each network size and offered load, the network throughput, the packet loss and the end-to-end packet delay. The evaluation of the performance of the three techniques was carried out using the Contiki COOJA simulator [5], available at [6].

The main contribution of this paper is the evaluation of the three data collection techniques. This evaluation enabled us to conclude that Technique 3 allows a set of nodes in a column to communicate between themselves with low probability of

packet loss and high values of average throughput. Results show that as the size of the network grows, the best performance in terms of packet loss and throughput is achieved using Technique 3. However, this will lead to having a larger overall delay for the data collection process.

The rest of the paper is organized as follows. Section II gives a background on the work related to this research work. Section III presents the topology and the communications architecture proposed for the WSN, as well as the different techniques for data collecting and communication through the WSN. Results are evaluated in Section IV. Finally, paper conclusions and ideas for future work are listed in Section V.

## II. RELATED WORK

In [7], Wang et al. presented an in-depth survey on recent advances in networked wireless sensor data collection. They highlighted the special features of sensor data collection in WSNs, by comparing it with both wired sensor data collection networks and other applications using WSNs. They discussed issues and prior solutions on sensor network deployment and data delivery protocols. Also, they discussed different approaches for control message dissemination, which acts as an indispensable component for network control and management and can greatly affect the overall performance of WSNs for sensor data collections. In our work, we test the performance of a WSN in a strip-based deployment while employing different data techniques that were mentioned by [7].

In [8], Accettura et al. evaluated the performance of Routing Protocol for Low Power and Lossy Networks (RPL) [9], using the Contiki COOJA simulator. They showed that RPL can grant very fast network set-ups and bounded communication delays. Additionally, its effectiveness can be further improved in terms of overhead, which can be very high due to Destination Advertisement Object (DAO) messages, used to handle sink-to-node downward traffic. We use RPL in our work with storing mode and DAO messages in order to allow downward traffic from the sink to the rest of the nodes.

In [10], Wang et al. proposed an RPL based routing protocol design for Advanced Metering Infrastructure (AMI) networks in smart grid. They used expected transmission count (ETX) as the link metric and proposed a low-cost ETX measurement scheme. Also, the authors have proposed a novel, ETX-based rank computation method to be used for the construction and maintenance of the RPL routing tables, which provided high end-to-end reliability for the inward unicast traffic in AMI networks. The simulation results presented in that paper showed that, in the presence of shadow fading, their proposed routing protocol produces satisfactory performances in terms of packet delivery ratio and end-to-end delay. In our work, we assume similar configuration of nodes sensing the voltage output of the panels. However, we use the hop count metric when computing the rank.

## III. PROPOSED ARCHITECTURE

As stated in the introduction, the objective of this work is to put a solar plant operating at an optimum power delivery

point. Each solar panel will have a sensor/actuator made of small and low power wireless sensor node that will sense local variables, communicate these values with other sensors in order to reach the optimum power delivery point of the whole plant.

Because the example network we are using for this study is a photo voltaic plant where panels are spread over a huge area, the network is simulated as a grid system of sensors where each sensor is connected to a solar panel and has constant distances to each of its neighbors. Additionally, since the most of the communications is done within each column in order to reach the optimum power delivery point of this column, we decided to divide the whole network into small networks of columns.

For the routing we propose to use Routing Protocol for Low Power and Lossy Networks (RPL). RPL [9] is a routing protocol that organizes routers along a Destination Oriented Directed Acyclic Graph (DODAG), a category of Directed Acyclic Graph (DAG), rooted at the sink. Using RPL, each column is considered a DAG with its DAG root located in the center. Each sensor reads the output voltage of its respective panel and forwards it to a central node. The proposed WSN nodes topology and communications architecture is shown in Figure 1.

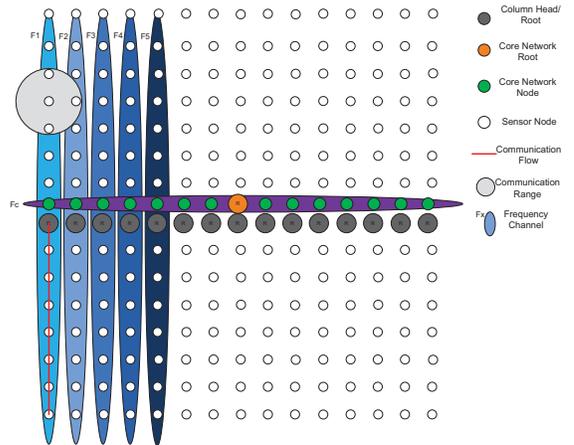


Fig. 1. Network topology and communications architecture.

Such a setup is referred to as a strip-based deployment [7]. It allows for full coverage while each node having a maximum of two neighbors. This allows nodes to have smaller transmission range and so use small transmission powers.

Additionally, in order to reduce the interference between columns, each column will operate in a frequency channel different from the other columns. To enable inter-column communication, a core network will be created operating in a frequency different from all the columns, and when the need for inter-column communication arises, the respective column DAG roots will have to switch to this core frequency and communicate through the core network. However, in this paper only intra-column communication is considered.

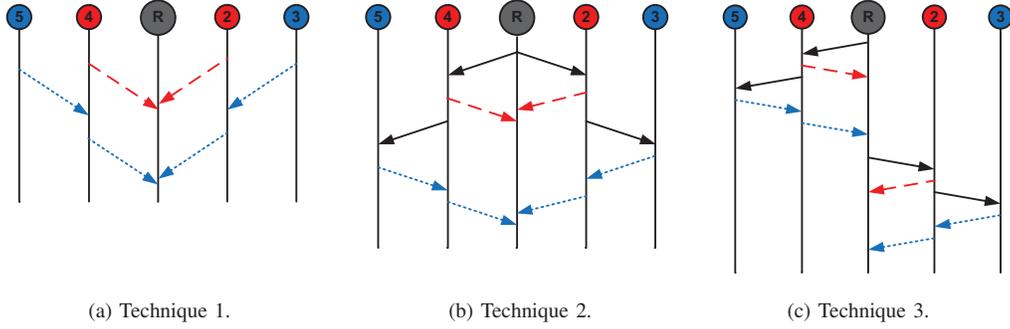


Fig. 2. Proposed data collection and communication techniques.

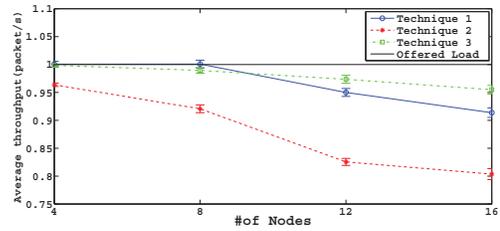
Throughout this work, three techniques were analyzed for data collection and communication between the nodes. Figure 2 shows an example of how these techniques work with 4 client nodes.

In Technique 1, client nodes generate and send data to the sink which is the basic way of communications in WSN. In Technique 2, the sink sends a broadcast poll message to its neighboring nodes and each node replies with its data as soon as it receives this poll before forwarding it to its other neighbors thus propagating the poll message across the column. This is done in order to reduce collisions between the messages. Finally, in Technique 3, the procedure is similar to Technique 2, excepts that the sink first sends the poll to one of its neighbors and waits until it receives the information from all the nodes on this side of the network before sending the poll to the other neighbor in order to receive data from the rest of the nodes on the other side of the network.

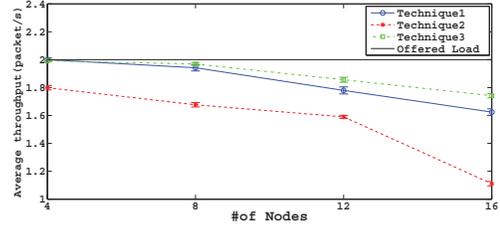
#### IV. RESULTS AND ANALYSIS

Simulations have been done using Contiki2.5 [6] and Sky notes [11]. The three data collection techniques were simulated in Contiki's built-in COOJA simulator using different numbers of nodes in the network (4, 8, 12 and 16) and for different values of the offered load (1, 2 and 4 packet/s), where the offered load is the data rate that each client node is offering to the network and total offered load is the sum of the offered loads from all the client nodes in the network. Inter packet arrival times were made sure to be exponentially distributed in order to simulate a real life scenario. Simulations were repeated 10 times for each case with each run working until the sink receives a total of 10000 packets. Mean values were calculated with a 90% confidence interval. Additionally, Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) with no acknowledgments was used in the simulations as the medium access protocol. Acknowledgments were omitted in order to reduce unnecessary traffic. IEEE 802.15.4 PHY 2.4 GHz [12] was used as the physical layer technology. The channel model used in the simulations is the Unit Disk Graph Medium (UDGM) with free space distance loss and no fading. Nodes were placed apart so that each node have always two neighbors falling within its transmission and

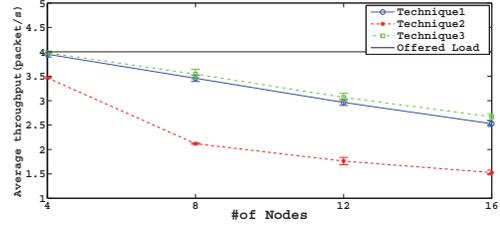
interference ranges.



(a) Average throughput for an offered load of 1 packet/s/node.



(b) Average throughput for an offered load of 2 packet/s/node.



(c) Average throughput for an offered load of 4 packet/s/node.

Fig. 3. Average throughput for different values of offered load.

Throughput is calculated per node as the number of received packets at the sink from such a node divided by the simulation time. Then this value is averaged over the total number of simulated nodes to get the average throughput. Total delay is the time needed to collect the data from all the nodes in the column. The total packet loss is the total number of transmitted packets from all the nodes minus the total number of received packets at the sink divided by the total number of transmitted packets from all the nodes.

Next we show the results gathered from these simulations for the different techniques proposed.

Figure 3(a), shows the average throughput in packet/s for

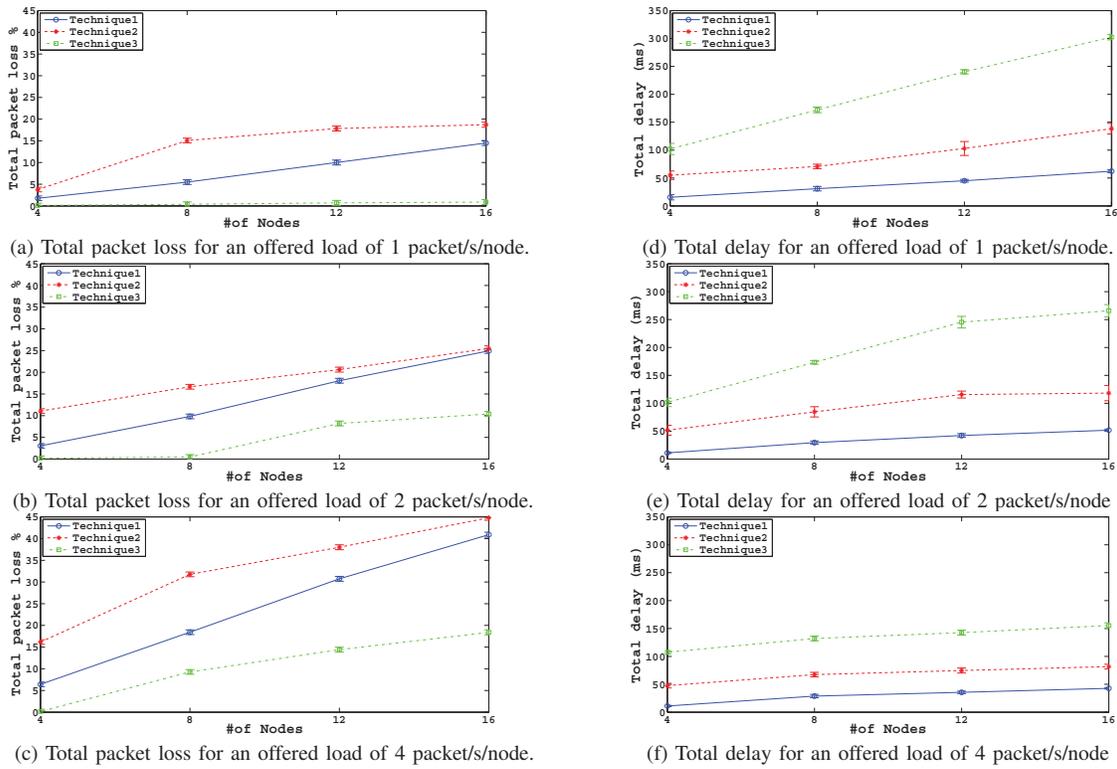


Fig. 4. Packet loss and total delay for different values of offered load.

an offered load of 1 packet/s/node. It is plotted against the number of simulated nodes in each simulation for the three techniques and the offered load. It is clear from the figure that, as the number of nodes increases, the throughput decreases due to the error coming from collisions and from retransmissions. However, Technique 3 operates the closest to the offered load limit in all the cases, while Technique 2 performs worst.

Figure 3(b), shows the average throughput in packet/s for an offered load of 2 packet/s/node. As we can see, curves follow the same trend as for the 1 packet/s/node offered load. However, the throughput decreases much faster as the number of nodes increases. Technique 3 still performs best.

Figure 3(c), shows the average throughput in packets/s for an offered load of 4 packet/s/node. Throughput decreases even faster with the increase of the number of nodes in the network. In this case, Technique 1 and 3 perform very close to each other. This is mainly due to the increased number of collisions that comes from the increased number of transmitted packets. Technique 2 still performs worst. Comparing this figure with the two previous figures, we can see that as the offered load increases the decrease in the throughput becomes faster in all the techniques.

Figure 4(a) shows the total packet loss as a percentage of the total number of packets generated by the network, for an offered load of 1 packet/s/node. We can see that packet loss increases as the number of nodes in the network increases, and with the increase of the offered load. However, Technique

3 performs much better than the other two techniques due to the fact that it tries to avoid collisions by design. Technique 2 suffers more losses because the two nodes neighboring the sink have a high probability of collisions as they send their reply to the sink almost at the same time.

Figure 4(b) shows the total packet loss for an offered load of 2 packet/s/node. As the offered load increases, the loss increases for all the techniques. While techniques 1 and 2 act similar to the previous case, Technique 3 still almost has no loss for low number of nodes, but as the number increases, the loss increases greatly.

Figure 4(c) shows the total packet loss for an offered load of 4 packet/s/node. We can clearly see the gain Technique 3 has over the other techniques in terms of packet loss. While techniques 1 and 2 have loss up to 40%, Technique 3 has only 15% losses.

Figures 4(d), 4(e), and 4(f) show the total delay in milliseconds for offered loads of 1, 2, and 4 packet/s/node, respectively. We can see that the total delay is not affected greatly by the increase in offered load for the values considered. It depends more on the number of nodes in the network. Additionally, Technique 1 clearly has the smallest delay because nodes start sending their data without waiting for a message for the sink, while the other two techniques have more delay, because they are based on a polling mechanism.

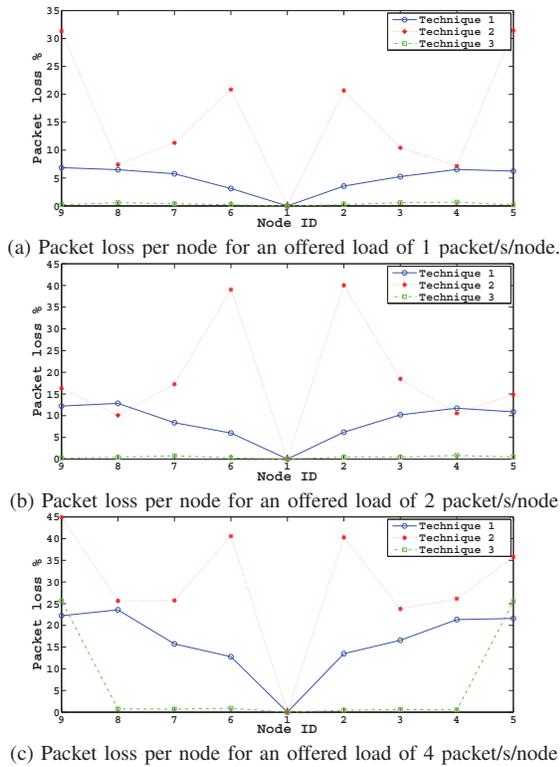


Fig. 5. Node packet loss of a network with 8 client nodes for different values of offered load..

Figures 5(a), 5(b), and 5(c), show the packet loss per node against the node id for a network with 8 client nodes with the sink being node 1, for offered loads of 1, 2, and 4 packet/s/node, respectively. Losses increase as we move away from the sink when using technique 1. In technique 2, nodes neighboring the sink suffers high losses because they are hidden from each other and so their packets collide with higher probability. Technique 3 performs best because it tries to avoid collisions by design. However, as the load increase, further nodes start to suffer high losses.

From what we can see from these results, Technique 3 performs best in terms of throughput and packet losses. However, due to the polling mechanism, data collection takes much longer than Technique 1.

Technique 2 performs worst overall due to the high error from collisions between the two nodes neighboring the sink. They are considered hidden from each other and so their packets collide with high probability. Thus, this technique will be abandoned in further testing.

As the load increases, performance degrades for all the techniques. However, Technique 3 still outperforms the other techniques in terms of throughput and packet loss.

For Technique 3, a conclusion we got during the design is that the offered load and the number of nodes are connected. In order to avoid collisions, the time between polls has to be larger than the time it takes to collect the data. If this is not realized, loss will increase for the further nodes, which will reduce the total throughput of the system.

## V. CONCLUSIONS AND FUTURE WORK

In this paper, we studied the performance of an IEEE 802.15.4 WSN network employing three different data collecting techniques using the Contiki COOJA simulator. The study was done for different network sizes and different values of the offered load, estimating for each network size and offered load, network throughput, packet loss and end-to-end packet delay.

The evaluation of the three techniques enabled us to conclude that Technique 3 allows a set of nodes in a column to communicate between themselves with low probability of packet loss and high values of average throughput. Results showed that as the size of the network grows, Technique 3 performs best in terms of packet loss and throughput. However, this will lead to having a larger overall delay for the data collection process.

The work done in this paper focused on trying to find the best communication technique within the same column with the sink located in the center. The following research aspects we will try to tackle are the inter column communication, as well as the self configuration problem, in order to create a totally scalable and self-configurable network.

## VI. ACKNOWLEDGMENTS

This work is funded by the ERDF through the Programme COMPETE and by the Portuguese Government through FCT-Fundação para a Ciência e Tecnologia, project ref. CMU-PT/SIA/0005/2009 and by the research grant number SFRH / BD / 68759 / 2010.

## REFERENCES

- [1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," in *Computer Networks*, vol. 52, no. 12, 2008, pp. 2292–2330.
- [2] M. Dohler, D. Barthel, R. Maraninchi, L. Mounier, S. Aubert, C. Dugas, A. Buhrig, R. Paugnat, M. Renaudin, A. Duda, M. Heusse, and R. Valois, "The ARESA project: Facilitating research, development and commercialization of WSNs," in *The 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks. SECON '07*, june 2007, pp. 590–599.
- [3] A. Bachir, M. Dohler, T. Watteyne, and K. Leung, "MAC essentials for wireless sensor networks," *IEEE Communications Surveys Tutorials*, vol. 12, no. 2, pp. 222–248, 2010.
- [4] "SELF-PVP," in <http://www.cmuportugal.org/tiercontent.aspx?id=3374>.
- [5] F. Osterlind, A. Dunkels, J. Eriksson, N. Finne, and T. Voigt, "Cross-level sensor network simulation with COOJA," in *Proceedings of the 31st IEEE Conference on Local Computer Networks*, nov. 2006, pp. 641–648.
- [6] "ContikiOS," in <http://www.contiki-os.org/>.
- [7] F. Wang and J. Liu, "Networked wireless sensor data collection: Issues, challenges, and approaches," *IEEE Communications Surveys Tutorials*, vol. 13, no. 4, pp. 673–687, 2011.
- [8] N. Accettura, L. Grieco, G. Boggia, and P. Camarda, "Performance analysis of the RPL routing protocol," in *IEEE International Conference on Mechatronics (ICM)*, april 2011, pp. 767–772.
- [9] T. Winter and P. Thubert, "RPL: IPv6 routing protocol for low power and lossy networks, draft-ietf-roll-rpl-15 (work in progress)," in *IETF ROLL working group*, february 2010.
- [10] D. Wang, Z. Tao, J. Zhang, and A. Abouzeid, "RPL based routing for advanced metering infrastructure in smart grid," in *IEEE International Conference on Communications Workshops (ICC)*, may 2010, pp. 1–6.
- [11] "Tmote Sky," in <http://www.snm.ethz.ch/Projects/TmoteSky>.
- [12] "IEEE802.15.4," in <http://www.ieee802.org/15/pub/TG4.html>.