

Adaptive Data Collection protocol for Extending Lifetime of Periodic Sensor Networks

Ali K. M. Al-Qurabat

Department of Software, College of Information Technology, University of Babylon - Iraq
alik.m.alqurabat@uobabylon.edu.iq

Ali Kadhun Idrees

Department of Computer Science, College of Science for Women, University of Babylon - Iraq
ali.idrees@uobabylon.edu.iq

ARTICLE INFO

Article History:

Received: 20 March 2017

Accepted: 1 April 2017

Published: 10 April 2017

DOI:

10.25212/lfu.qzj.2.2.11

Keywords: *Periodic Sensor Networks, Data Collection, Adaptive Sampling Rate, Network Lifetime.*

ABSTRACT

In Periodic Sensor Network (PSN), the periodic energy efficient collection of a large amount of sensed data by the sensor nodes is considered as one of the main challenges in this type of network. An adaptive sampling approach to periodic data collection is needed for energy optimization due to the limited lifetime nature of the sensor batteries. In this paper, we propose a protocol, called an Adaptive Data Collection protocol (ADaC), which collects periodically sensor readings and prolong the lifetime of a Periodic Sensor Network (PSN). The lifetime of ADaC protocol is divided into cycles. Each cycle is composed of two stages. First, data collection. Second, sampling rate adaptation based on the similarity between periods of one cycle using Euclidean distance measure to adapt its rate of sampling according to the dynamic modification of the monitored environment. ADaC allows each sensor to adapt its sampling rate in accordance with the monitored environment conditions. We conduct extensive simulation experiments on real sensor data by applying OMNeT++ network simulator to explain the effectiveness of the ADaC protocol in comparison with a method without using adaptive sampling.

1. INTRODUCTION

Sensor nodes are spatially deployed in the region of interest in order to monitor the physical or environmental phenomena like temperature, humidity, light, pollution, pressure and sound. They collect the sensed data from the monitored environment, manipulate the data locally, and transmit them to the sink for further analysis. These sensor nodes work in a collaborative manner and constitute a Wireless Sensor Network (WSN) [1, 2, 3, 18, 19]. In sensor node, the radio unit represents the

principal source of energy consumption. Therefore, it is important to remove redundant sensed data before reporting them to the sink to save the energy and improve the lifetime of sensor node [4].

Data collection approaches determine the way of sensor's work in data collection and sending to the base station. There are two models for data collection in WSNs: time driven and event-driven [5, 20]. This work considers time-driven data collection which is named Periodic Sensor Networks (PSNs). In PSN, every sensor node transmits the sensed data of the monitored area to the sink periodically. Several PSNs applications use the periodic way to monitor certain conditions regularly such as pressure, humidity, temperature, etc. Two main challenges in PSN. First, PSN has to provide adequate lifetime in order to satisfy application's needs. Second, data management is more difficult due to the huge amount of collected data by this network. In PSN, the change in the monitored environment can slow down or speed up. The energy consumption can be decreased when the sensor node modifies its sampling rate based on the dynamic modification of the monitored phenomena. Therefore, to prolong the network lifetime, adaptive sampling for periodic data collection is required for energy optimization and data reduction [6].

This paper introduces the following contributions.

- i) A protocol named ADaC is devised to collect the sensor data in an adaptive way. The principal idea of ADaC protocol is to utilize the similarity of collected data and adapts its sampling rate accordingly. ADaC works into cycles. Two stages in each cycle: data collection and sampling rate adaptation using Euclidean distance measure. The sensor node provides a new sampling rate after each cycle based on the similarity between the periods of one cycle.
- ii) A new adaptive sampling rate algorithm based on the Euclidean similarity is suggested. In each cycle, the speed of readings capturing inside the sensor node depends mainly on the previously calculated sampling rate adaptively.
- iii) The simulation results are accomplished by OMNeT++ network simulator to illustrate the effectiveness of the ADaC protocol. The ADaC protocol has been compared to the results of the method without using adaptive sampling.

The rest of this paper is organized as follows: - Next section exhibits literature review. Section 3 explains the description of ADaC protocol. Protocol evaluation is shown in Section 4. Finally, we present conclusion and future works in Section 5.

2. LITERATURE REVIEW

This section reviews some related literature concerning the adaptive data collection in WSNs. Adaptive collection approaches are considered as a good candidate to save energy and extend the network lifetime of PSNs. The major objective of an adaptive collection technique is to make the sensor node be able to change its sampling rate dynamically in accordance with the monitored environment conditions. This can reduce the repetitive gathered data, consume less energy, and decrease the processing load at the base station [6].

Some researchers used prediction as a way to adjust the sampling rate of sensor nodes [7, 8]. An energy saving information gathering scheme is proposed by Liu et al. [7] to predict the sampling rate inside sensor using ARIMA model. In [8], the authors presented an algorithm for adaptive sampling using Box-Jenkins approach to estimate the future sensor readings, depending on the existing readings.

The work proposed in [9] consider adaptive sampling schemes based spatial correlation among the physical sensed data. In [9], the sampling rate is adapted by the base station. Initially, the base station activates a set of sensors to get the sensed data of monitored

environment. The correlation percentage is computed for the received sensed data to increase or decrease the activated sensors. Chatterja and Havinga [10] present a sampling algorithm based temporal correlation among sensed data. In this algorithm, the sampling rate is modified depending on the stability of the monitored environment. The sampling rate increases when the environment conditions are unstable, otherwise the rate decreases. Masoum et al. [11] introduce an energy-saving mechanism for data collection. Their scheme exploits spatio-temporal correlation among sensors and their sensed data to determine the candidate sensors which are responsible for sampling and transmission. The selected sensors are adaptively changed.

In recent years, Laiymani and Makhoul [12] proposed a scheme for adaptive sampling using ANOVA model and Fisher test in PSNs. This algorithm works at the sensor node to adapt its sampling rate. The authors in [13] proposed method to remove the repetition of collected data in PSN called Prefix Frequency Filtering (PFF).

This paper suggests an Adaptive DATA Collection (ADaC) protocol for PSNs. The major goal of ADaC is to remove redundant sensor readings, save energy, and improve the network lifetime. ADaC performs two main phases. First, data collection according to adaptive sampling rate. Second, ADaC allows to each sensor node to adapt its sampling rate for each cycle (cycle = 2 periods) based on the Eculidean similarity. ADaC is simulated on the OMNeT++ network simulator using real data of sensor nodes. The comparison results show that our protocol can provide a better performance and prolong the network lifetime.

3. DESCRIPTIO OF THE ADaC PROTOCOL

ADaC protocol is given in more details in this section. The main objective of this protocol is to enable each sensor to modify its sampling rate adaptively in accordance with the dynamic changing of the monitored environment. Consequently, this reduces the amount of redundant gathered data and minimizes energy consumption (extend the PSN lifetime) whereas the quality of collected data is maintained sufficiently to allow significant analysis. Figure 1 illustrates the flowchart of the proposed ADaC protocol. Table 1 explains some parameters used in this paper.

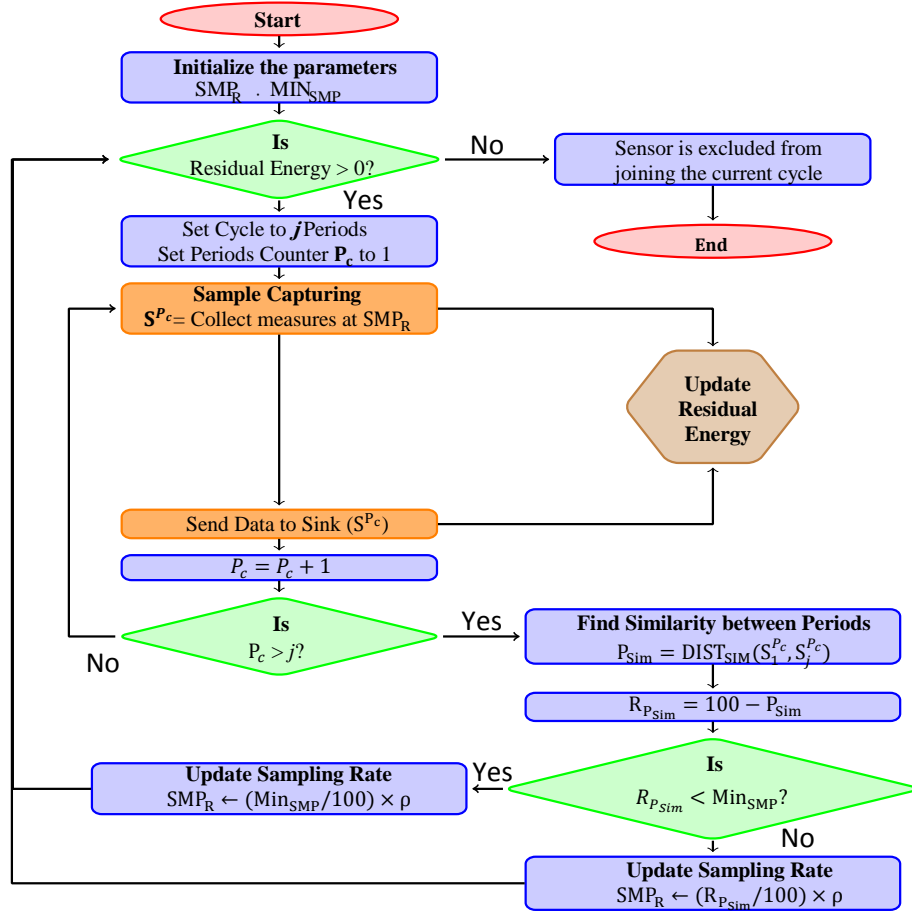


Figure 1. Flowchart of proposed ADaC protocol.

Table 1: Some parameters used in this paper

Parameter	Value
SMP_R	Sampling Rate= ρ
S	Temperature readings series $S = \{s_1, s_2, \dots, s_{\rho-1}, s_{\rho}\}$
MIN_{SMP}	Application criticality (minimum amount of sampling rate)
P_{Sim}	Similarity percentage
R_{Psim}	Reverse of similarity percentage

3.1. Data collection

Each sensor node senses the data reading periodically. These sensed time-arranged data readings set forms time series. Therefore, ADaC protocol treats the sensor readings as a time series. In our work, we named this time series as a temperature readings series.

ADaC protocol is periodic and works into cycles. The cycle includes two periods ($j=2$). The period is partitioned into time slots. Therefore, in this stage, the sensor node n catches one temperature reading S_i each time slot. At the end of each period, the temperature readings series of sensor n is formed such that $S_n = \{s_1, s_2, \dots, s_{\rho-1}, s_{\rho}\}$, where ρ is the total number of temperature readings captured

during the period. The sensor node collects the temperature readings at sampling rate (SMP_R) speed. The SMP_R is initiated to ρ temperature readings per period. The redundant temperature readings captured by the sensor node increase in two states: short time slot and slowly variation of a monitored area of interest.

3.2. Adaptive sampling rate

In this stage, ADaC protocol modifies its sampling rate based on the percentage of similarity between temperature readings of different periods in the cycle. The main purpose of this stage is to calculate the similarity among periods after each finished cycle to acclimate the rate of sampling according to the new similarity rate. Therefore, it uses the Euclidean distance measure to find the amount of similarity between periods of each cycle.

Euclidean distance represents one of the most famous distance measures. ADaC protocol uses similar function to identify the Similarity between two temperature readings series Q and C of the same length ρ using the following formula [14]

$$DIST_{SIM}(Q, C) = \sqrt{\sum_{i=1}^{\rho} (q_i - c_i)^2} \quad (1)$$

After that, in order to measure the similarity percentage (P_{Sim}), it is defined as follow:

$$P_{Sim} = DIST_{SIM}(Q, C) \times 100 \quad (2)$$

Algorithm 1 gives the similarity percentage (P_{Sim}) calculation between two temperature readings series Q and C.

Algorithm 1. Similarity Algorithm

Input: Q, C (two temperature readings series with ρ dimension).

Output: P_{Sim} .

Process:

```

1: sum ← 0
2: for i ← 1 to  $\rho$  do
3:   Sum ← Sum +  $(Q_i - C_i)^2$ 
4: end for
5: if (Sum = 0) then
6:    $DIST_{SIM} \leftarrow 0$ 
7: else
8:    $DIST_{SIM} \leftarrow \sqrt{Sum}$ 
9: end if
10:  $P_{Sim} \leftarrow DIST_{SIM} \times 100$ 
11: return  $P_{Sim}$ 

```

3.3. Verification the similarity of periods

In ADaC protocol, every node able to adapt its rate of sampling according to the amount of similarity among temperature readings series collected during different periods. The aim of computing the similarity between the temperature readings series every cycle is to adapt the rate of sampling based on the new calculated similarity. Therefore, the Euclidean distance is employed to discover the similarity percentage, P_{Sim} among several periods per cycle. On one hand, if P_{Sim} is high, it means the

monitored condition is changed at a slow speed. Therefore, the sensor node will decrease its rate of sampling to the minimum value to prevent collecting redundant readings. On the other hand, if P_{Sim} is low, the sensor node will collect temperature readings at approximately maximum sampling rate so as to prevent losing significant readings. Therefore, to adapt the rate of sampling of sensor node in accordance with the computed similarity among periods, the reverse of similarity percentage for Euclidean coefficient (R_{PSim}) is computed as follows

$$R_{PSim} = 100 - P_{Sim} \quad (3)$$

Consequently, the process of adapting the sampling rate in the sensor node depends on the R_{PSim} , thus the application criticality will be taken into consideration in this process. In ADaC protocol, the criticality of application is expressed as a minimum amount of sampling rate in a period for a sensor node, MIN_{SMP} . MIN_{SMP} takes values in the range 0 to 100 which represent the criticality level either low or high respectively. The sensor node adapts the new sampling rate to the MIN_{SMP} (not to the R_{PSim}) when the recently calculated sampling rate is less than MIN_{SMP} . Depending on the requirements of the application and before the deployment, all the sensor nodes initialize their MIN_{SMP} . It is also possible to change MIN_{SMP} dynamically during the lifetime of the network for the whole sensors or for just a given subgroup of sensors if there are some types of management and control schemes are available. Algorithm 2 illustrates an adaptive sampling rate approach.

Algorithm 2. Adaptive Sampling Rate Algorithm

Input: J (one cycle = j periods), ρ , MIN_{SMP} .
Output: SMP_R (new sampling rate).
Process:

- 1: $SMP_R \leftarrow \rho$ // (initialize sampling rate to ρ readings per period)
- 2: **while** $n_e > 0$ **do**
- 3: **for** $i \leftarrow 1$ **to** j **do**
- 4: Collect temperature readings series (S_i) at SMP_R speed
- 5: Send_ToSink(S_i)
- 6: **end for**
- 7: **for each cycle do**
- 8: $P_{Sim} \leftarrow \text{Similarity}(S_1, S_j)$
- 9: $R_{PSim} \leftarrow 100 - P_{Sim}$
- 10: **if** ($R_{PSim} < MIN_{SMP}$) **then**
- 11: $SMP_R \leftarrow \left(\frac{MIN_{SMP}}{100}\right) \times \rho$
- 12: **else**
- 13: $SMP_R \leftarrow \left(\frac{R_{PSim}}{100}\right) \times \rho$
- 14: **end if**
- 15: **end for**
- 16: **end while**
- 17: **return** SMP_R

4. Protocol evaluation

The evaluation of ADaC protocol is given in more details in this section. The main objective of this protocol is to enable each sensor to modify its sampling rate adaptively in accordance with the dynamic changing of the monitored environment. Consequently, ADaC protocol will remove redundant sensor readings, save energy, and improve the network lifetime.

4.1. Simulation framework

To study and evaluate ADaC protocol, extensive simulations are performed with discrete event simulator OMNet++ [15]. ADaC protocol is distributed at each sensor node and it is based on the dataset of Intel Berkeley Research Lab [16]. PSN in this Lab includes 54 Mica2Dot sensors localized at the lab. The base station is located at the center of the lab. It receives sensed readings from each sensor node by a single hop.

The sensed data of the weather (such as temperature, humidity, and light) are periodically collected by these sensors once each 31 seconds. In our simulation, the sensor nodes use a log file contains about 2.3 million readings collected previously by Mica2Dot sensor nodes in the Lab. This article uses only one measure of sensor node measurements: temperature². There are 7 sensor nodes do not used in our simulation because its data may be missed or truncated. Therefore, the temperature readings of 47 sensor nodes are selected and stored. The results are the average of 47 sensor nodes. Table 2 gives the selected parameters settings.

Table 2: Simulation parameters for PSN initialization

Parameter	Value
PSN size	47 nodes
ρ	20, 50, 100 and 200 readings
MIN_{SMP}	20, 40 and 60
E_{elec}	50 nJ/bit
β_{amp}	100 pJ/bit/m ²

In the experimental simulations, ρ (the total number of temperature readings captured during the period) is set to 20, 50, 100 and 200 readings per period because almost the papers (e.g. [6], [13] and [17]) use these values and we need to compare with these papers in the next step of our work. In addition, some performance metrics are applied to assess the effectiveness of the ADaC protocol such as sampling rate adaptation, number of collected and sent temperature readings by a sensor node, and energy consumption.

ADaC protocol uses the same energy consumption model discussed in [17]. Energy consumed by the sensor node depends only on the periodically collected and sent temperature readings to the base station. The cost of transmission is calculated for a m – bits message and for a distance d as follow

$$E_{TX}(m, d) = E_{elec} \times m + \beta_{amp} \times m \times d^2 \quad (4)$$

The radio expends energy as described in [17], where E_{elec} is the energy consumed for radio electronics and it is equal to 50 nJ/bit, β_{amp} is the energy consumed by the amplifier and it is equal to 100 pJ/bit/m².

The energy consumption required for capturing m –bits by the sensor node is calculated as follow

$$E_{CX}(m, d) = E_{TX}(m, d) \div 7 \quad (5)$$

² The others are done by the same manner.

These experiment simulations consider the length of data reading m equal to 64 (we choose this length because almost the papers (e.g. [6], [13] and [17]) use this length). In the case of transmission, 64 bits are added to m – bits message which corresponds to the frequency of data reading m . Consequently, Energy consumption is defined as the total energy dissipated at each sensor node during the collection and transmission of data readings and formulated as follow

$$E_{Total} = E_{TX}(m, d) + E_{CX}(m, d) \quad (6)$$

4.2. Sampling rate adaptation

Figure 2 shows the adaptation of sampling rate and for two sizes of temperature readings (50 and 100 respectively). The results illustrate the ability of sensor device to modify its rate of sampling dynamically depending on the application criticality level. The risk level MIN_{SMP} can be determined according to the type and requirement of application used to monitor the disaster.

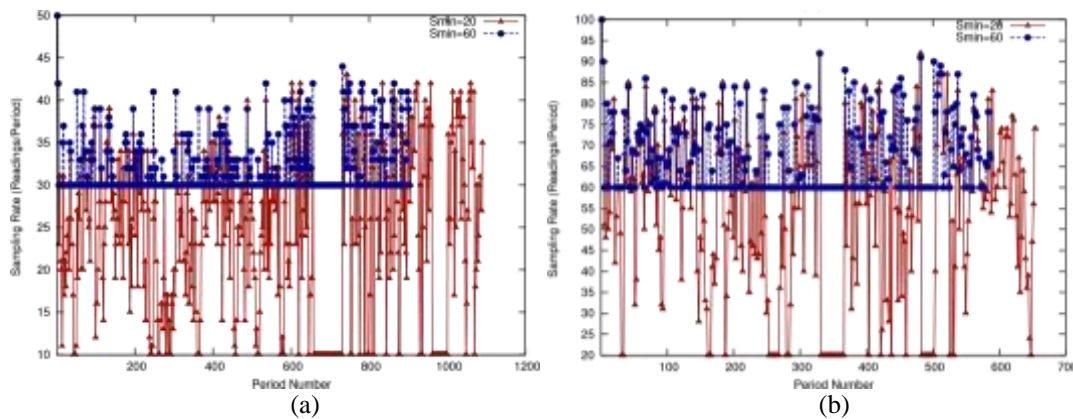


Figure 2. Sampling rate adaptation (A) $\rho = 50$ AND (B) $\rho = 100$.

In this experiment, MIN_{SMP} uses two values: 20 for low risk level disaster and 60 for high risk level disaster. As shown in Figure 2, the adaptation of sampling rate is dynamic and after each cycle based on the application criticality level (i.e., $MIN_{SMP} = 20$ or 60). The results in Figure 2 (a) and (b) validate the good performance of our protocol.

4.3. Percentage of sent readings

When collecting the data readings at each period, ADaC protocol at the sensor node able to decrease the number of sent readings to the base station by using adaptive sampling method. Figure 3 demonstrates the percentage of sent readings by a sensor node to the base station at the end of simulation for ADaC protocol compared with the results of the method without using adaptive sampling technique.

The results illustrate that ADaC protocol at each sensor node decreases from 26% to 69% of the Percentage of sent readings to the base station as comparing to 99% without using adaptive sampling. Therefore, ADaC protocol removes the redundant collected readings successfully and the percentage of sent readings to the base station is reduced. We can also see that the volume of sent readings from the sensor node to the base station

decreases when ρ increases or MIN_{SMP} decreases. This is due to the percentage of sent readings rely on the number of collected readings and the risk level of application.

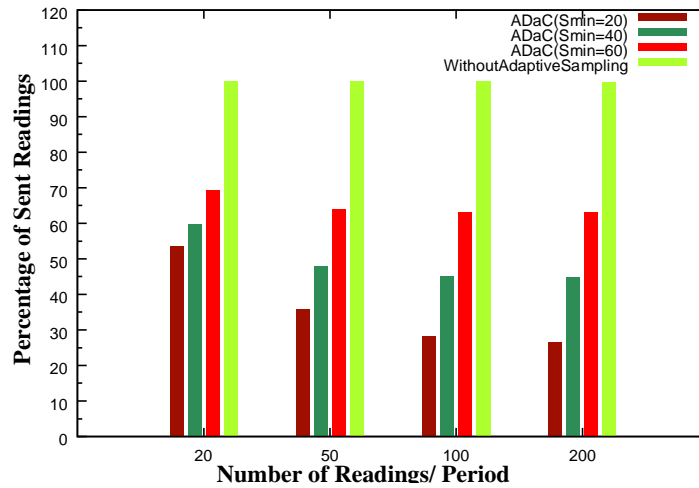


Figure 3. Percentage of sent readings.

4.4. Energy consumption

Figure 4 shows the energy consuming by ADaC protocol at the sensor node compared with the results of the method without using adaptive sampling technique.

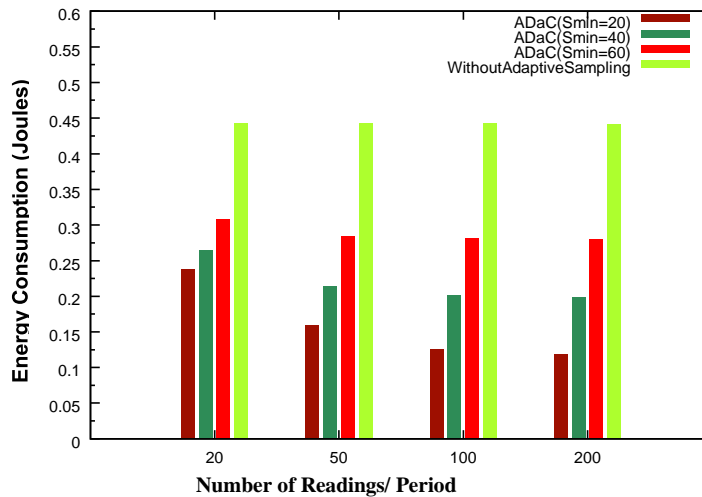


Figure 4. Energy consumption at each sensor node.

As shown in Figure 4, ADaC protocol saves energy because it reduces both collected and sent readings at the sensor node. The consumed energy of a sensor node using ADaC protocol is minimized from 26% to 68% as compared with the results of the method without using adaptive sampling technique. It can be observed that ADaC protocol is effective in terms of reducing energy consumption for the applications with high and low risk level, where more energy is saved when MIN_{SMP} decreases.

5. Conclusion and future works

This paper presents a protocol, called Adaptive Data Collection protocol (ADaC), which collects periodically sensor readings and improves the PSN lifetime. ADaC protocol works into cycles and consists of two phases. First, collecting the data readings. Second, sampling resolution to adapt the rate of sampling at the sensor node in accordance with the dynamic changing of observed environment. ADaC protocol considers the risk level of an application by fixing the minimum sampling rate that permits to sensor node to collect readings at a minimum rate while maintaining a good quality of the collected readings. To assess the effectiveness of ADaC protocol, we compared it with the results of the method without using adaptive sampling technique using several performance metrics like a percentage of sent readings and energy consumption. Simulation results show the efficiency of ADaC protocol to conserve the energy at the sensor nodes thus prolong the PSN lifetime.

In future, we plan to improve our work to consider the sensing overlap among sensor nodes at the aggregator level to optimize both the aggregated readings and lifetime while maintaining a good accuracy. In addition, we plan to compare our proposed method with an existing method in the literature [20].

6. References

- [1] I.F. Akyildiz, and M.C. Vuran, *Wireless sensor networks*, Vol. 4. John Wiley & Sons, 2010.
- [2] M. Abdelaal, O. Theel, C. Kuka, P. Zhang, Y. Gao, V. Bashlovkina, D. Nicklas, and M. Fränze, “Improving energy efficiency in QoS-constrained wireless sensor networks,” *International Journal of Distributed Sensor Networks*, 2016.
- [3] A.K. Idrees, K. Deschinkel, M. Salomon, and R. Couturier, “Distributed lifetime coverage optimization protocol in wireless sensor networks,” *The Journal of Supercomputing*, vol. 71, no.12, pp.4578-4593, 2015.
- [4] X. Tang and J. Xu, “Adaptive data collection strategies for lifetime constrained wireless sensor networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 6, pp. 721–734, 2008.
- [5] G. Campobello, A. Segreto, and S. Serrano, “Data gathering techniques for wireless sensor networks: A comparison,” *International Journal of Distributed Sensor Networks*, vol. 2016, 2016.
- [6] A. Makhoul, H. Harb, and D. Laiymani, “Residual energy-based adaptive data collection approach for periodic sensor networks,” *Ad Hoc Networks*, vol. 35, pp. 149–160, 2015.
- [7] C. Liu, K. Wu, and M. Tsao, “Energy efficient information collection with the arima model in wireless sensor networks,” in *Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE*, vol. 5, pp. 5–pp., 2005.
- [8] Y. W. Law, S. Chatterjea, J. Jin, T. Hanselmann, and M. Palaniswami, “Energy-efficient data acquisition by adaptive sampling for wireless sensor networks,” in *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*. ACM, pp. 1146–1151, 2009.
- [9] R. Willett, A. Martin, and R. Nowak, “Backcasting: adaptive sampling for sensor networks,” in *Proceedings of the 3rd international symposium on Information processing in sensor networks*. ACM, pp. 124–133, 2004.
- [10] S. Chatterjea and P. Havinga, “An adaptive and autonomous sensor sampling frequency control scheme for energy-efficient data acquisition in wireless sensor networks,” in

-
- International Conference on Distributed Computing in Sensor Systems*. Springer, pp. 60–78, 2008.
- [11] A. Masoum, N. Meratnia, and P. J. Havinga, “An energy-efficient adaptive sampling scheme for wireless sensor networks,” in *Intelligent Sensors, Sensor Networks and Information Processing, 2013 IEEE Eighth International Conference on*. IEEE, pp. 231–236, 2013.
- [12] D. Laiymani and A. Makhoul, “Adaptive data collection approach for periodic sensor networks,” in *9th International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, pp. 1448–1453, 2013.
- [13] J. M. Bahi, A. Makhoul, and M. Medlej, “A two tiers data aggregation scheme for periodic sensor networks.” *Adhoc & Sensor Wireless Networks*, vol. 21, no. 1, 2014.
- [14] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, “A symbolic representation of time series, with implications for streaming algorithms,” in *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*. ACM, pp. 2–11, 2003.
- [15] A. Varga’s, “Omnet++ discrete event simulator,” [Online] 2003, <https://omnetpp.org/> (Accessed: 10 January 2017).
- [16] S. Madden, “Intel berkeley research lab,” [Online] 2004, <http://db.csail.mit.edu/labdata/labdata.html>. (Accessed: 25 August 2016).
- [17] H. Harb, A. Makhoul, A. Jaber, R. Tawil, and O. Bazzi, “Adaptive data collection approach based on sets similarity function for saving energy in periodic sensor networks,” *International Journal of Information Technology and Management*, vol. 15, no. 4, pp. 346-363, 2016.
- [18] Idrees AK, Deschinkel K, Salomon M, Couturier R. Coverage and lifetime optimization in heterogeneous energy wireless sensor networks. ICN 2014. 2014 Feb 23:60.
- [19] Idrees AK, Deschinkel K, Salomon M, Couturier R., “Perimeter-based coverage optimization to improve lifetime in wireless sensor networks,” *Engineering Optimization*, vol. 48, no. 11, pp. 1951-72, 2016.
- [20] Harb, H., Makhoul, A., Jaber, A., Tawil, R. and Bazzi, O., “Adaptive data collection approach based on sets similarity function for saving energy in periodic sensor networks,” *International Journal of Information Technology and Management*, vol. 15, no. 4, pp.346-363, 2016.