



HAL
open science

An efficient partial data aggregation scheme in WSNs

Imane Horyia Brahmi, Soufiene Djahel, Damien Magoni, John Murphy

► **To cite this version:**

Imane Horyia Brahmi, Soufiene Djahel, Damien Magoni, John Murphy. An efficient partial data aggregation scheme in WSNs. IFIP Wireless Days, Nov 2014, Rio de Janeiro, Brazil. pp.3, <10.1109/WD.2014.7020845>. <hal-01117344>

HAL Id: hal-01117344

<https://hal.science/hal-01117344v1>

Submitted on 26 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

An Efficient Partial Data Aggregation Scheme in WSNs

Imane Horiya Brahmi
Univ. Coll. Dublin, Lero
Dublin, Ireland
horiya-imane.brahmi@ucd.ie

Soufiene Djahel
Univ. Coll. Dublin, Lero
Dublin, Ireland
soufiene.djahel@ucd.ie

Damien Magoni
Univ. Bordeaux, LaBRI
Talence, France
magoni@labri.fr

John Murphy
Univ. Coll. Dublin, Lero
Dublin, Ireland
j.murphy@ucd.ie

Abstract—Highly accurate event detection makes Wireless Sensor Networks popular for real time monitoring. Wireless sensor systems that monitor physical and environmental conditions are expected to be deployed with high density, a situation which leads to spatial correlations and redundancy of collected data. Eliminating these redundancies extends the network lifetime by reducing energy consumption and enhances the velocity of transmitting emergency and periodic sampling. We consider in this paper the scenario where sensors are grouped into clusters. Each Cluster Head (CH) receives sampling from its Cluster Member (CMs), and decides when it should stop sampling and start the transmission of the aggregated packets in order to reduce the transmission delay while ensuring the accuracy of the transmitted data. In this work, we propose a cluster based aggregation scheme which determine the best timing at the CH level for achieving a short delay, and a buffer management strategy for maintaining low energy consumption. Evaluation results based on simulations show that our scheme achieve a good trade-off between energy consumption and end-to-end delay.

Index Terms—Data aggregation, Dynamic waiting time, Periodic packet transmission, Spatial correlation, WSN.

I. INTRODUCTION

Wireless Sensor Networks (WSNs) [1] is an emerging technology for event detection and information gathering. Their main feature is to cooperatively sense specified events of interest in the sensor field, and transmit them toward the sink for processing via multi-hop routing protocol. This type of network has become popular due to its low cost and high accuracy detection. It has diverse applicability that includes areas such as environment, health, agriculture, and military, where different physical conditions are monitored such as temperature, humidity, vibration, air quality, etc.

Sensors are resources constrained with limited processing, storage and battery. A simple solution to this problem could be the replacement of the sensors' battery power. However, this solution is not convenient, mainly when the sensors are used for applications, such as monitoring hostile or non accessible areas like volcanos. In addition, sensor node density in the network may vary spatially and temporally depending on the application requirements [2]. With high density the reported data is likely to be redundant, correlated or copied. We usually associate the primary energy consumption with communication: it is well known that the communication is often the most expensive activity in terms of energy [3],

as the radio communication is the major source of energy consumption [4].

An effective solution of reducing the communication overhead is data aggregation. In-network processing and data aggregation is defined as the process of gathering the data, pre-processing and computing it in the network itself and transmitting the extracted and required data to the sink. Hence, energy is conserved by eliminating redundancy and minimizing the transmission of raw data to the sink [5].

Delaying the transmission of the data packets, in order wait for more packets from neighboring nodes may increase the degree of aggregation [6], [7]. However, in some applications such as monitoring, it is very critical to deliver the readings in a timely manner, especially in monitoring applications. For this purpose, the solutions designed for the WSNs should consider the end-to-end delay in addition to the energy consumption in their conception.

The new generation sensor nodes have been enhanced with significant energy-efficient storage, processing capabilities and data management abilities [8]. Indeed, sensor nodes can be endowed with energy-efficient storage such as new-generation flash memory with several gigabytes of storage and low-power consumption [9], [10]. This emergence of low cost and high-capacity storage and processing prompts us to design a new cluster based aggregation scheme and buffer management strategy that satisfies the trade off between the energy consumption and end-to-end delay, by reducing the number of unnecessary redundant transmissions.

The remainder of this paper is organized as follows. In Section II, we present the related contributions on data aggregation in WSNs and highlight their limitations. Next, we present our scheme in detail in Section III. In Section IV, we present and discuss our results. Finally, we conclude the paper in Section V.

II. RELATED WORK

Two surveys [5],[11] have explored the most significant contributions dealing with data aggregation based on cluster approaches - [12],[13]. In this latter, a Low Energy Adaptive Clustering Hierarchy (LEACH) is proposed. LEACH protocol is based on two phases: a setup phase during which the network is organized into clusters and the cluster heads are chosen, and a steady phase in which the data is aggregated and

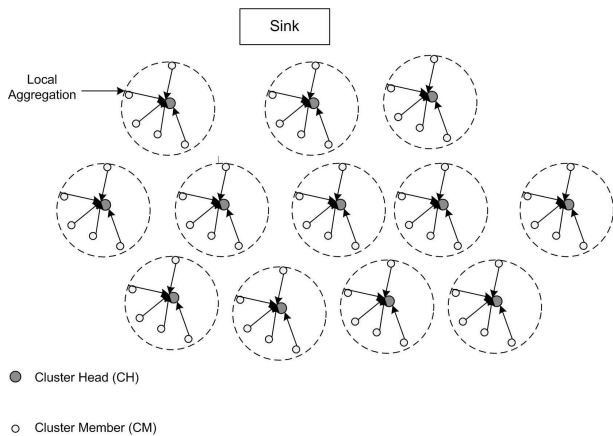


Fig. 1. Clustered WSN topology.

transmitted to the sink. Another protocol extending LEACH protocol is presented in [14], where the cluster head is selected periodically according to the node residual energy.

More recently, the authors of [2] studied the problem of constructing a spatial correlation by proposing dYnamic and scalable tree Aware of Spatial correlaTION (YEAST) algorithm. The sensor nodes that detect the same event are grouped in a correlated region and the CH is selected and rotated in each round. However, this protocol is appropriate in event driven not for a periodic monitoring when the data is transmitted.

III. SYSTEM DESIGN

We consider in our scheme that sensor nodes are organized into clusters as shown in Figure 1 to make the aggregation and the communication easier and more efficient. Each CH gathers readings from its CMs, and aggregates them before sending the resulting packets toward the sink. We define the gathering time as the Local Waiting Time (WT_{Local}). This time is dynamic and determines when the CHs should aggregate the received readings and forward the resulting packets before the end of the monitoring period. These aggregated readings are then sent onward towards the sink via multi-hop routing path. We assume that these sensors are equipped with buffer memory and queue as shown in the Figure 2. The CHs' buffer is used for storing the received packets from their CMs. Each CH waits for a minimum number of samples before it aggregates and forwards the aggregated packets. A copy of the aggregated samples are then stored in the CH buffer. These copies of aggregated samples are used for aggregating future samples received from other sources (i.e. CMs). The forwarding queue is used for storing the aggregated packets which are now ready to be compared against future packets. The samples stored in the buffer have a limited storage time. Therefore, we define a Time out (T_{out}) for the packets stored in the buffer. This T_{out} represents the maximum storage time for each packet in the buffer. Our focus in this work is to design a dynamic local waiting time and a T_{out} mechanism for the packets in the buffer such that they optimize the energy usage, buffer storage

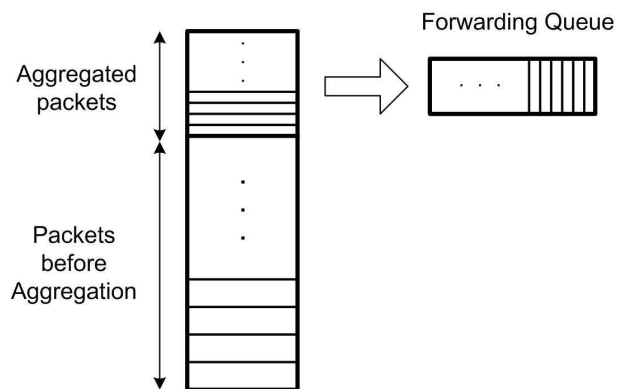


Fig. 2. Representation of a sensor's buffer.

time, and transmission delay of packets, and finally reduce the overall traffic load in the network.

A. Waiting Time Computation

The CMs start periodically measuring given parameters, and forward the reading to their CH in the network. We assume that the CMs start collecting data at the same time (i.e. the beginning of each period). At the end of each monitoring period, the CMs start forwarding their readings (i.e. one reading per CM). The arrival time of the readings at the CH level is random. The aggregation process starts after receiving these first samples.

Our mechanism uses the following key parameters: a minimal number of samples received by the CH, depending on the application, the mean of the received samples values, the calculated standard deviation σ_{cal} and we also define a standard deviation threshold σ_{th} which depend on the application. These parameters are used to select appropriate time to start aggregating the received packets and forward the resulting packets.

When the first samples are received by the CHs, they start the aggregation process, compute the mean and standard deviation (σ_{cal}) of the received values, and compare them with σ_{th} . If the minimal number of required samples is reached, and the threshold condition is satisfied, the samples can be aggregated and sent onward. If a received sample reports a value exceeding a given upper bound of the measured parameter (e.g. a vehicle out of control on the road), this sample is forwarded immediately as an emergency message.

Once the conditions of minimal number of required samples and σ_{cal} is met, the aggregated packets are forwarded and a copy of these aggregated packets is stored in the CH's buffer for limited period of time T_{out} .

B. Sensor Buffer Management Strategy

Since the buffer capacity and energy consumption are limited in WSN nodes, our second objective is to optimize the energy consumption of sensor nodes (CHs) by keeping the packets in the buffer an optimal time period. In this section we define a scheme for sensor buffer management and processing

of the packets, and illustrate the T_{out} calculation, which ensures an optimal trade-off between energy consumption and packet storage duration. Our T_{out} calculation is a function of the following parameters: the sensors nodes popularity, the frequency of receiving the same packets from the same category, their priority level, the residual energy of each CH and the buffer size of the CHs.

$$T_{out} = f(R, F, W, E) \quad (1)$$

Where:

- R: popularity of the CH: computed based on the traffic load transiting through this CH
- F: frequency of receiving the same packets from the same category
- W: packet's priority level
- E: residual energy of the CH node (E_{res} , E_{Max})

We assume that the maximum energy level and the buffer capacity of each node are both known in advance. The popularity of a node is defined by the ratio $R = \frac{N_{\text{ofPacketsTransmitted}}}{N_{\text{ofPacketsReceived}}}$. The closer the nodes are to the sink, the longer they retain their packets in the buffer, since the traffic density is higher in the surrounding of the sink.

We assume that:

$$T_{out} = [0, T_{outMax}] \quad (2)$$

The initial $T_{outMax0}$ depends on the energy and the buffer capacity of the sensor nodes. At $t = 0$ (when the network starts running) the buffer is empty and the energy capacity of the node is at its maximum level, thus we have $T_{outMax0} = T_{outMax}$. In order to give more priority to the packets with higher priority level, every time a packet (P_n) is received by the CHs, they check the category of the packet. For each type (i.e., based on the parameter measured: vehicle speed, temperature, etc.) and category, we adjust the T_{outMax} according to the priority level of the packet received by defining a weighting mechanism. The $T_{outMax0}$ is proportional to the level of priority of the packet. If a packet has a high priority level, then the probability to retain the packet longer in the buffer is higher.

$$T_{outMax0}(Per_i) = T_{outMax0} \times W(Per_i) \quad (3)$$

Where:

$$W(Per_i) = \frac{Priority_i}{Priority_{max}} \quad (4)$$

Such that (Per_i) refers to the type of a periodic packet and $W(Per_i)$ is its corresponding weight. $Priority_i$ indicates the priority level of a given packet type. Once the buffer is full and a new packet arrives, the sensor executes Algorithm 1. When new packets are received, the CHs have to decide which packets are the most essential for future aggregation. For this, the CHs check the frequency of receiving the same packets from the same category, which will indicate any unusual event. The packet with the smallest T_{out} and frequency and the lowest priority will be deleted first. Next, each CH calculates

the new T_{out} based on its remaining battery power. If the same packet is received many times, its T_{out} will be calculated according to its corresponding Average Inter-Arrival time (T_{AIA}), for the purpose of extending its storage time for future aggregation.

Algorithm 1 Sensor Buffer Management Algorithm

```

1: At  $t = T$  (when the buffer is full)
2: if new packet received then
3:   Check if the same packet already exists in the buffer
4:   if packet exists then
5:     Check  $T_{AIA}$ , the packet's category and type
6:     if  $T_{out} \leq T_{AIA}$  then
7:        $T_{out}(P_n) = T_{outMax0}(P_{type})[(\frac{E_{res}(node)}{E_{Max}(node)}) \times R] + T_{AIA}$ 
8:     else
9:        $T_{out}(P_n) = T_{outMax0}(P_{type})[(\frac{E_{res}(node)}{E_{Max}(node)}) \times R]$ 
10:    end if
11:  else
12:    Add the packet to the buffer and calculate its  $T_{out}$ 
13:  end if
14: end if

```

IV. PERFORMANCE EVALUATION

TABLE I
SUMMARY OF SIMULATION PARAMETERS

Parameters	Value
Routing Protocol	AODV
Propagation mode	TwoRayGround
Packet Size	64 Bytes
Number of nodes	100
Inter-node distance	10 meters
Monitoring period interval	15 s, 30 s, 60 s, 120 s
Simulation time	3600 seconds
No. of simulation runs	50
Topology	grid

In this section, we evaluate and discuss the performance of our proposed scheme (Partial Aggregation: P_{Agg}) using the $ns-2.35$ [15] network simulator. We have compared the P_{Agg} to the Full aggregation (F_{Agg}) where the CHs wait to receive packets from all their CMs, and Aggregation Off (Agg_Off) where the received packets are forward immediately upon their reception. Our solution is static cluster based and the cost of deployment and maintenance of the clusters is not considered. The generation of the packets by the CMs and their values are random during each periodic monitoring interval. Once the first packets are received by the CHs, the aggregation process starts by computing the standard deviation (σ_{cal}) of the received packets value and comparing it with σ_{th} , in order to decide when the aggregated packets should be sent. Once the condition is met the CH aggregates the received packets, forwards them and stores a copy in their buffer.

We summarize in Table I the default parameters used in the simulation. The primary metrics evaluated are: (i) the Average End-to-End (E2E) transmission delay of all the packets, which

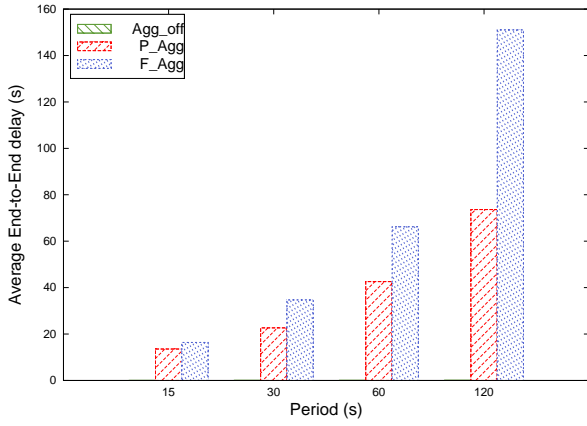


Fig. 3. Average end-to-end delay vs monitoring period (Network Size = 100 nodes).

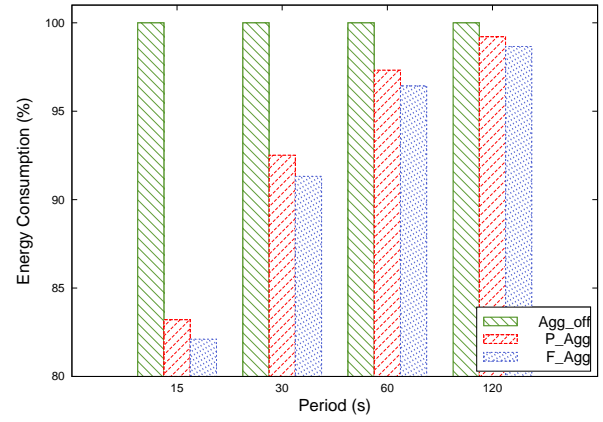


Fig. 4. Energy consumption vs monitoring period (Network Size = 100 nodes).

represents the average time needed for a packet sent by the source to cross the network and reach the sink. (ii) The Energy Consumption (EC) defined by the ratio of the total energy used when the aggregation is enabled to the total energy used when the aggregation is disabled.

The results plotted in Figure 3 compare the average end-to-end delay of the three schemes. In this scenario we have set the network size to 100 nodes and varied the monitoring period. From these results, we can see that the Agg_{off} achieves negligible E2E delay, this is explained by the immediate forwarding of the received packets from the CMs. The early aggregation in the P_{Agg} leads to a lower E2E delay compared to the F_{Agg} , this is due to the long waiting time scheme used in F_{Agg} . We notice that the difference achieved is equal to 20% on average, and increases to approximately 60% with the increase of the monitoring period. From this we can say that the long waiting time is unnecessary in F_{Agg} .

In Figure 4 we notice that the energy consumption achieved by P_{Agg} is as low as that achieved by F_{Agg} . Also the energy consumption increases with the monitoring period interval. The main objective of our work is to achieve a lower end-to-end delay of the packet transmission, while keeping a low energy consumption in the network. We can conclude from our results that our scheme succeeds in this.

V. CONCLUSION

We have proposed in this work a new cluster based aggregation scheme for periodic monitoring to deal with the problem of long waiting times in WSNs at the CH level. We have first defined a solution based on a dynamic waiting time which uses an optimal timing for aggregating and forwarding the packets. Next, we proposed a buffer management strategy for processing the stored packets in an efficient way such that the energy power is saved and the optimally aggregated packets are accurate. The performance evaluation results have proven the efficiency of our scheme in terms of end-to-end delay and energy consumption.

REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002.
- [2] L. A. Villas, A. Boukerche, H. A. De Oliveira, R. B. De Araujo, and A. A. Loureiro, "A spatial correlation aware algorithm to perform efficient data collection in wireless sensor networks," *Ad Hoc Networks*, 2011.
- [3] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, 2000.
- [4] A. Somov, I. Minakov, A. Simalatsar, G. Fontana, and R. Passerone, "A methodology for power consumption evaluation of wireless sensor networks," in *Emerging Technologies & Factory Automation, 2009. ETFA 2009. IEEE Conference on*. IEEE, 2009, pp. 1–8.
- [5] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-network aggregation techniques for wireless sensor networks: a survey," *Wireless Communications, IEEE*, vol. 14, no. 2, pp. 70–87, 2007.
- [6] H. Yousefi, M. H. Yeganeh, N. Alinaghipour, and A. Movaghar, "Structure-free real-time data aggregation in wireless sensor networks," *Computer Communications*, vol. 35, no. 9, pp. 1132–1140, 2012.
- [7] H. Li, K. Lin, and K. Li, "Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks," *Computer Communications*, vol. 34, no. 4, pp. 591–597, 2011.
- [8] Q. Wang, K. Ren, S. Yu, and W. Lou, "Dependable and secure sensor data storage with dynamic integrity assurance," *ACM Transactions on Sensor Networks (TOSN)*, vol. 8, no. 1, p. 9, 2011.
- [9] G. Mathur, P. Desnoyers, P. Chukiu, D. Ganesan, and P. Shenoy, "Ultra-low power data storage for sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 5, no. 4, p. 33, 2009.
- [10] A. Mitra, A. Banerjee, W. Najjar, D. Zeinalipour-Yazti, V. Kalogeraki, and D. Gunopulos, "High performance, low power sensor platforms featuring gigabyte scale storage," in *Third International Workshop on Measurement, Modeling, and Performance Analysis of Wireless Sensor Networks (SenMetrics 2005)*, San Diego, CA, 2005.
- [11] R. Rajagopalan and P. K. Varshney, "Data aggregation techniques in sensor networks: A survey," 2006.
- [12] S. Lindsey and C. S. Raghavendra, "Pegasis: Power-efficient gathering in sensor information systems," in *Aerospace conference proceedings, 2002. IEEE*, vol. 3. IEEE, 2002, pp. 3–1125.
- [13] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*. IEEE, 2000, pp. 10–pp.
- [14] O. Younis and S. Fahmy, "Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *Mobile Computing, IEEE Transactions on*, vol. 3, no. 4, pp. 366–379, 2004.
- [15] "The network simulator ns-2 (<http://www.isi.edu/nsnam/ns/>)."