Efficient Data Collection in Wireless Sensor Networks using Spatial Correlation Algorithm

Mr. Amit S. Bhosale
Department of Computer Technology
Kavikulguru Instijtute of Technology and Science
Ramtek, Nagpur
kitsamit@gmail.com

Mr. Sanjay R Khajure
Department of Computer Technology,
Kavikulguru Institute of Technology and Science
Ramtek, Nagpur
sanjaykhajure@gmail.com

Mr. Manish S. Sharma
Department of Computer Technology
Kavikulguru Institute of Technology and Science
Ramtek, Nagpur
manishh.sharmaa@gmail. com

Abstract**—**Large scale dense Wireless Sensor Networks (WSNs) will be increasingly deployed in different classes of applications for accurate monitoring. Due to the high density of nodes in these networks, it is likely that redundant data will be detected by nearby nodes when sensing an event. Since energy conservation is a key issue in WSNs, data fusion and aggregation should be exploited in order to save energy. In this case, redundant data can be aggregated at intermediate nodes reducing the size and number of exchanged messages and, thus, decreasing communication costs and energy consumption. The Spatial Correlation mechanism takes the advantages of WSN routing techniques to perform energy aware data forwarding to save more energy. Data aggregation algorithm do not take into account the sensor node's energy to choose the representative nodes. In Spatial Correlation mechanism in which nodes that detected the same event are grouped in correlated regions and a representative node is selected at each correlation region. From the results we can conclude that the energy required to send sensed data using Spatial Correlation mechanism is less and on an average 46.25 % (app.) energy saving was achieved and also it reduce the number of nodes to 37.25 % (app.) who used to send sensory data which ultimately reduces control packet overhead.

Keywords: WSN, Data Aggregation, Spatial Correlation, Energy Saving.

*****

I. INTRODUCTION

A WSN consists of scattered autonomous sensors to observe physical or environmental conditions (such as pressure, temperature, and sound etc.) and to send their data through the network to a base station. WSNs are deployed in security applications such as factory monitoring, environmental monitoring, burglar alarms and fire alarms. The sensor nodes for these applications are typically deployed in unsecured locations and are not made tamper-proof due to cost considerations. Hence, an adversary or an attacker could take control of one or more sensor nodes and launch active attacks to threaten correct network operations. Such environments pose a particularly challenging set of constraints for the protocol designer. Sensor network protocols must be highly energy efficient while being able to function securely in the presence of possible malicious nodes within the network [1].

Sensor networks made up of small self-computing devices capable of producing digital representations of real-world environment phenomena. Due to size and battery power limitations, these devices typically have limited storage capacity, limited energy resources, and limited network bandwidth. Data produced by nodes in the network propagates through the network via wireless links. When compared to local processing of data, wireless transmission is extremely expensive. Researchers at the University of California, Santa Barbara estimate that sending a single bit over radio is at least three orders of magnitude more expensive than executing a single instruction. The limited amount of energy, bandwidth, and storage capacity available to sensor nodes calls for specialized optimizations of queries injected into the network.

Initially, in-network aggregation techniques involved different ways to route packets in order to combine data coming from different sources but directed towards the same destination. These protocols were simply routing algorithms which differed from more traditional ad hoc routing protocols in the metric they used to select the routing paths [3]. In¬network data aggregation is a complex problem that involves many layers of the protocol stack and different aspects of protocol design, and a characterization and classification of concepts [2].

The remainder of this paper is organized as follows. Section I introduces about the In-Network Aggregation. Protocols in In-Network data aggregation are categorized in tree based, clusters based and structure less approaches and their details in given in section 3. Section 4 concludes the paper with a comparative summary of the surveyed approaches. Section 5 give the result of the simulation.

A. In-network Aggregation

 ISSN: 2321-8169
418– 423

"In-network aggregation is the global process of gathering and routing information through a multihop network, processing data at intermediate nodes with the objective of reducing resource consumption (in particular energy), thereby increasing network lifetime" [3].

We can distinguish between two approaches:

• With size reduction

It refers to the process of merging and reducing data coming from different sources in order to reduce the information to be sent over the network. As an example, assume that a node receives two packets from two different sources containing the locally measured temperatures. Instead of forwarding the two packets, the sensor may compute the average of the two readings and send it in a single packet.

• Without size reduction

It refers to the process of merging packets coming from different sources into the same packet without data processing: assume to receive two packets carrying different physical quantities, e.g. temperature and humidity. These two values cannot be processed together but they can still be transmitted in a single packet, thereby reducing overhead.

II. EXISTING IN-NETWORK DATA AGGREGATION PROTOCOLS

Data aggregation process is performed by specific routing protocol. The main aim is aggregating data to minimize the energy consumption. So sensor nodes should route packets based on the data packet content and choose the next hoping order to promote in-network aggregation. Basically routing protocol is divided by the network structure, that's why routing protocols in in-network data aggregation are categorized in tree based, clusters based and structure less approaches.

A. Tree Based Protocols

The main aim of tree based protocols is to maintain the energy consumption of the sensor nodes by using multihop communication within the cluster and by performing data aggregation in order to reduce the number of transmitted messages. Tree based approaches includes Direct Diffusion (DD), Shortest Path Tree (SPT and Center at Nearest Protocol (CNS).

1) Direct Diffusion

Intanagonwiwat el. al. [5] have developed Direct Diffusion (DD), which consists of several key elements such as interests, data messages, gradients, and reinforcements as shown in figure 2.1. An interest message is a query which states what a user wants. Each interest contains a narration of a sensing task that is supported by a sensor network for collecting data. Typically, sensor network data are collected information of a physical phenomenon. Such data can be an event, which is a short description of the sensed phenomenon. In directed diffusion, data are named using attribute-value pairs. A sensing task is dispersed throughout the sensor network as an interest for naming data. These scattering sets up gradients within the network designed to "draw" events i.e., data matching the interest. Specifically, a gradient is direction state created on each node that receives an interest. The gradient direction is set to the neighboring node from which the interest is received. Events start flowing toward the originators of interests along multiple gradient paths.

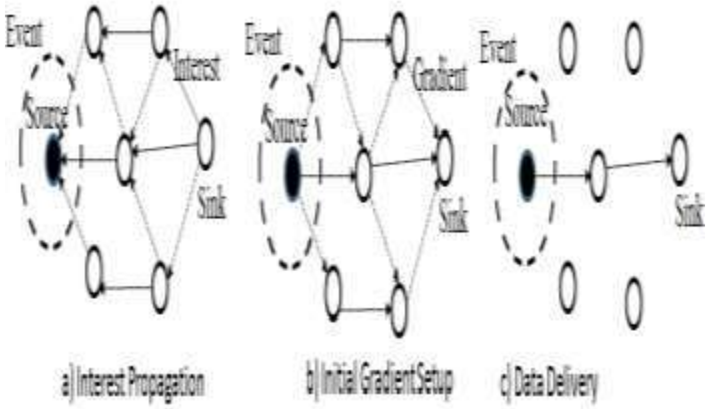

Figure 1. Direct EJiffiision

2) SPT & CNS

B. Krishnamchari et. al. in [6] discussed both these protocols assumes the single network flow model which means that there is only one sink node which collects information from the number of data sources. They use data centric routing schemes which means that during the process of transferring data from sources to the sink the routing nodes can look at the content of data and perform aggregation. In SPT protocol data aggregation scheme all nodes send their information to the sink along the shortest path between the two, and overlapping paths are combined to form the aggregation tree. In CNS protocol all source node sends their data directly to the source which nearest to the sink which sends aggregated data to the sink. The main factor that can affect the process of data aggregation is the placement of source nodes in the network. For this investigation author proposes two models of source placement i.e. the event radius (ER) model and the random source (RS) model. In both models sensor network is generated by scattering n sensor nodes, one of which is sink, in a unit square. In ER model all sources are spread within a range of distance S from random event location and in RS model k nodes are randomly selected to be source. All nodes can communicate with the other nodes in the range of the communication radius R.

B. Cluster based Protocols

Cluster based protocol are similar to the hierarchical organization of the network which includes different protocols like Low Energy Adaptive Clustering Hierarchy (LEACH), Information Based Role Assignment (InFRA), Data Aggregation Aware Routing Protocol (DAARP), and Dynamic Data-Aggregation Aware Routing Protocol (DDAARP), dYnamic and scalablE tree Aware of Spatial Correlation (YEAST).

1) LEACH

Heinzelman, et. al. [4] were first to introduce LEACH. It is a self-organizing, adaptive clustering protocol, which includes dynamic and distributed cluster formation.

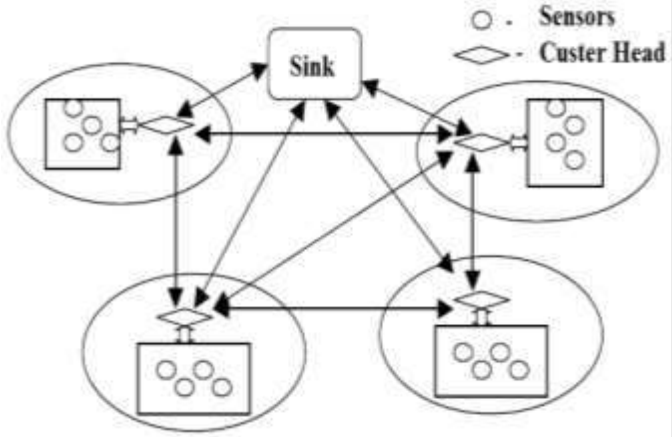

Figure 2. LEACH

The diagrammatic representation of LEACH is as shown in figure 2.2. LEACH randomly selects a sensor node as cluster heads (CHs) and switches this role to evenly allocate the energy load between the sensors in the network. In LEACH, the cluster head (CH) nodes merge data received from nodes inside the cluster, and send an aggregated packet to the sink node in order to decrease the amount of information that must be sent to the base station. In LEACH Time Division Multiple Access (TDMA) schedule is used to send data from node to head cluster. Head cluster aggregated data received from node inside the cluster. Communication is via Direct Sequence Spread Sequence (DSSS) and each cluster uses a unique spreading code to reduce inter cluster interference. Also the data from cluster head is sent to the Base Station with the help of unique spreading code and Carrier Sense Multiple Access (CSMA). In LEACH, a good property is to save energy by using aperiodic communication. After suitable time, a randomized rotation of the role of the CH is conducted so that uniform energy dissipation in the sensor network is obtained. The authors found, based on their simulation model that only 5% of the nodes need to act as cluster heads.

2) InFRA

Nakamura et. al. in [7] discuss the reactive algorithm Information Fusion based Role Assignment (InFRA), in which roles are assigned when any event takes place. Different roles are assign in this protocol like Sink, Collaborator, Coordinator and Relay. In this protocol when multiple nodes detect the same event, they organize themselves into clusters. Then cluster head aggregates data from all cluster members and sends event data towards the sink in multihop fashion. InFRA finds minimum shortest path tree connecting all source nodes to sink such that the intra cluster data aggregation is possible. InFRA provides role migration policy i.e. role of coordinator is transferred from one node to another so that the load of energy consumption is distributed evenly in the nodes inside the cluster. The two types of data aggregation schemes are followed in InFRA i.e. intra-cluster and inter-cluster. In first the data from collaborators are aggregated and in later the data from coordinator nodes are aggregated. A disadvantage about InFRA algorithm is that for every new event that detect by source nodes the information about that event is broadcasted throughout the network to inform other nodes in the network about its occurrence and to update the paths from the already existing cluster heads to the sink node and events are static in nature and of fixed radius. This limits InFRA scalability.

3) DAARP

L. A. Villas et. al. in [8] overcome the disadvantage of InFRA algorithm. Data Aggregation Aware Routing Protocol (DAARP) is new reliable Data Aggregation Aware Routing Protocol for WSN. Similar to InFRA, in this for each event this algorithm performs the clustering of nodes that detected the event in the network and also the election of cluster head. Then cluster head merges inter cluster data and sends the result to the sink node. After the cluster head formation routes are formed by selecting nodes in the network to existing routing path in which the node in existing path act as an aggregation point. This protocol reduces the number of messages during the setup phase of a routing tree and maximizes overlapping routes. It selects routes with the highest aggregation rate and performs reliable data aggregation transmission and uses fewer control packets to build the paths. Different from InFRA, DAARP does not broadcast a message to the whole network whenever a new event occurs. DAARP is not feasible for scenarios with long duration events because the routes are static, which quickly consumes the energy of the nodes that are part of the routing structure.

4) DDAARP

The up gradation of DAARP protocol is discussed by L. A. Villas in [9]. DDAARP builds dynamic routes. In this protocol the routes are configured and processed at the sink node and routes which are created are not dependent on the order of events. Routes established are not kept fixed in DDAARP throughout the duration of events i.e. routes are dynamic in nature. This protocol has low cost in terms of packet control, improves the quality of the routing tree and maximizes information fusion along the routing path. The drawback of this proposal is that packets containing information from nodes tend to increase their size at the information collection stage and this solution becomes impractical for large scale networks. In addition, the sink node needs to have a global knowledge of the network such as node positions, residual energy of nodes and nodes that detected events.

5) YEAST

L. A. Villas et. al. proposes the spatial correlation mechanism in [2]. YEAST is spatial correlation algorithm takes the advantage of best WSN techniques to perform efficient data collection in WSN. It uses spatial correlation model in which a correlation region is defined by the author. In this correlation region the value sensed by the sensor nodes are assumed to be similar and only one sensor node value can represent that value of the whole correlation region. The size of correlation region can be changed dynamically in YEAST. As shown in figure 2.3 the event is divided into cells. Each cell defines a correlated region and only one node within the cell notifies the sensed information. That node is called the leader node or coordinator node. To divide the energy flow evenly within the cell the role of leader transferred from one node to another.

C. Structure less Protocols

This category includes protocols named as Data Aware Anycast (DAA) [11] and Dynamic and Scalable Tree (DST) [12].

1) 2.3.1 DAA

Author K. W. Fan et. al. [11] explains the power of structure free data aggregation in event based application and also it is maintenance free. Two techniquesare propose by author at MAC layer called DAA and at Application layer called Randomized Waiting (RW) for efficient data aggregation. The main challenges evolved in structure free data aggregation are the routing decisions made on the fly as no structure is

available and as the nodes do not know about their neighbors they cannot wait for them before forwarding their own data. The drawback of this protocol is that it does not guarantee aggregation of transmitting packets in-network and with an increase in the size of network it increases the cost of transmitting packets.

2.3.2 DST

Author L. A. Villas et. al. in [12] proposes a novel routing protocol called DST which adopts to any scenario. Different from static routing schemes, in DST, routes which are created during data transmission are not held fixed and also not dependant on the order of events. It aims to build a routing tree

with the shortest routes in Euclidean distance that connects all source nodes to the sink node, maximizing data aggregation while reducing the distance connecting each coordinator node to sink. Routes are based on straight line segments, which are computed by coordinator nodes. Drawback of this protocol is that it do not explore the spatial correlation model. DST loses its performance in situations where nodes detect the same event.

III. COMPARATIVE SUMMARY OF PROTOCOLS

The following table shows the summary of characteristic of protocols of WSN.

TABLE I. COMPARATIVE SUMMARY

Algorithm	**Routing Structur**	**Objective**	**Data Aggregation Nodes**	**Overhead**	**Scalability**	**Spatial Correlatio**
Direct Diffusion	Tree Based	It constructs a spanning tree rooted at the Sink.	Clusterheads and Aggregatior Node	Medium	Low	No
SPT and CNS				High	Medium	No
LEACH	Cluster Based	It forms clustering of nodes and maximizes the overlapping of routes.	Clusterheads and Intermediate Nodes	Medium	Low	No
InFRA				Very High	Low	No
DAARP				Medium	Medium	No
DDAARP				Low	Medium	No
YEAST				Very Low	Very High	Yes
DAA	Structure less	Efficient data aggregation without explicit maintenance of structures.	Clusterheads and Intermediate Nodes	Very Low	Very High	No
DST				Very Low	Very High	No

IV. SIMULATION RESULTS

In this section we evaluate the performance of spatial correlation mechanism with without aggregation techniques.

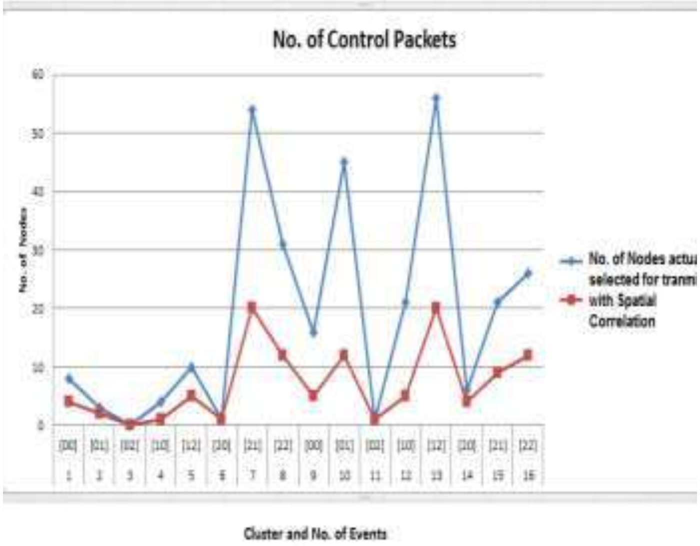

Figure 4. No. of packets send

Figure 4 shows the comparisons of the number of packets send during each event from each cluster. The graph shows that with number of increasing events number of packets transferred are also increases but at the same time there are less transmission of packets when aggregation is applied.

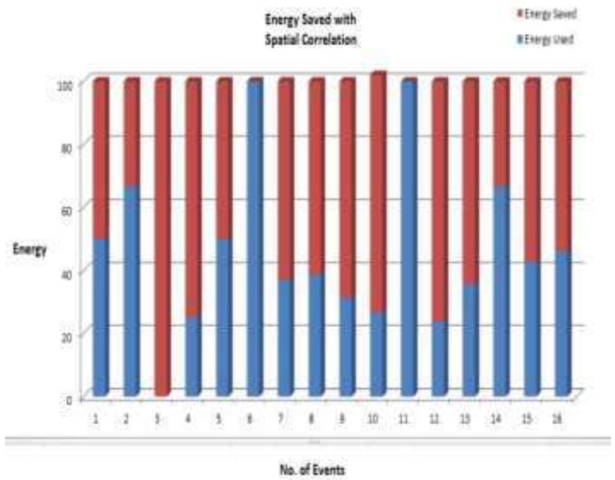

Figure 5. Energy Saving Figure 5. shows the comparison of the energy saving with Spatial Correlation Mechanism. The X axis shows the number of events with cluster and Y axis shows the energy.

V. CONCLUSIONS AND FUTURE SCOPE

Aggregation aware routing algorithms play an important role in event-based WSNs. The Spatial Correlation mechanism takes the advantages of WsN routing technique to perform energy aware data forwarding to save more energy. Data aggregation algorithms typically do not take into account the sensor node's energy to choose the representative nodes. When they do take, they present a high cost in terms of control messages and do not efficiently exploit the spatial correlation.

In the Spatial Correlation mechanism nodes that detected the same event are grouped in correlated regions and a representative node is selected at each correlation region for observing the phenomenon. From the results we can conclude that energy required to send sensed data using Spatial Correlation mechanism is less and on an average 46.25 % (app.) energy saving was achieved and as we also reduce the number of nodes to 37.25% (app.) who used to send sensory data which ultimately reduces control packet overhead.

We consider all sensor node are homogeneous in nature. But many application of sensor nodes contains heterogeneous sensor networks. As a future scope we will try to apply these techniques to some heterogenous sensor network.

REFERENCES

[1] Leandro A. Villas, A. Boukerche, H. Soares Ramos, Horacio A. B., Fernandis Oliveira, R. Borges and A. Loureiro, “DRINA: A Lightweight and Reliable Routing Approach for In-Network Aggregation in Wireless Sensor Network”, IEEE Transactions On Computers, vol. 62, Issue no. 4, pp.676-688, April 2013.

[2] Leandro A. Villas, A. Boukerche, H. A. de Araujo and A. A. Luoureiro, “A Spatial Correlation Aware Algorithm to perform Efficient Data Correlation Aware Algorithm to Perform Efficient Data Collection in Wireless Sensor Network”, Ad Hoc Network, http://www.sciencedirect.com/science.article/pii/s157087051 1001001 892, 2011.

[3] L.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cyirci, “Wireless Sensor Network: A Survey”, Computer Networks, vol. 38, Issue no. 4, pp.393-422, March2002.

[4] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan, “Energy-Efficient Communication Protocol forWireless Microsensor Networks”, Proceedings of the 33rd Hawaii International Conference on System Sciences - 2000.

[5] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, F. Silva, “Directed diffusion for wireless sensor networking”, IEEE/ACM Transaction on Networking, vol. 11 Issue no. 1, pp. 2-16, February 2003.

[6] B. Krishnamachari, D. Estrin, S.B. Wicker, “The impact of data aggregation in wireless sensor networks”, ICDCSW02: Proceedings of the 22nd International Conference on Distributed Computing Systems, IEEE Computer Society, Washington, DC, USA, pp. 575578, 2002.

[7] E.F. Nakamura, H.A.B.F. de Oliveira, L.F. Pontello, A.A.F. Loureiro, “On demand role assignment for event-detection in sensor networks”, ISCC06: Proceedings of the 11th IEEE Symposium on Computers and Communications, IEEE Computer Society, Washington, DC, USA, pp. 941-947, 2006.

[8] L. A. Villas, A. Boukerche, R. B. Araujo, A. A. Loureiro, “A reliable and data aggregation aware routing protocol for wireless sensor networks”, Proceedings of the 12th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, MSWiM09, ACM, New York, NY, USA, pp. 245-252, 2009.

[9] L. Villas, A. Boukerche, R. B. Araujo, “Highly dynamic routing protocol for data aggregation in sensor networks”, in

Proceedings of the IEEE Symposium on Computers and Communication, ISCC10, IEEE Computer Society, Berlin, 2010, pp. 496-502.

[10] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cyirci, “Wireless Sensor Networks: A Survey”, Computer Networks, vol. 38, Issue no. 4, pp. 393-422, March 2002.

[11] K.-W. Fan, S. Liu, P. Sinha, “On the potential of structure-free data aggregation in sensor networks”, INFOCOM 2006: Proceedings of the 25th IEEE International Conference on Computer Communications, pp. 1-12, 2006.

[12] L as, D.L. Guidoni, R.B. Araujo, A. Boukerche, A.A. Loureiro, “A scalable and dynamic data aggregation aware routing protocol for wireless sensor networks”, Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems, MSWIM ’10, ACM, New York, NY, USA, pp. 110-117, 2010.

[13] Jamal N. Al-Karaki, Ahmed E. Kamal, “ Taxonomy of Routing Techniques in WSN”, Handbook of Sensor Network: Compact Wireless and Wired Sensing System, chapter 6, pp. 117-140, Dec 2004.

[14] K. Akkaya, M. Younis, “A survey on routing protocols for wireless sensor networks”, Journal of Ad Hoc Networks, vol.3,Issue no. 3, pp. 325-349, May 2005.

[15] E. F. Nakamura, H.A.B.F. de Oliveira, L.F. Pontello, A.A.F. Loureiro, “On demand role assignment for event-detection in sensor networks”, ISCC06: Proceedings of the 11th IEEE Symposium on Computers and Communications, IEEE Computer Society, Washington, DC, USA, pp. 941-947, 2006.

[16] I. F. Akyildiz, M.C. Vuran, O. B. Akan, “On exploiting spatial and temporal correlation in wireless sensor networks”, Proceedings of WiOpt 2004: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, pp. 71-80, 2004.

[17] C. Liu, K. u, J. Pei, “An energy-efficient data collection framework for wireless sensor networks by exploiting spatiotemporal correlation”, IEEE Transactions on Parallel and Distributed Systems, pp. 1010-1023, 2007.

[18] L. Liu, S. Member, P.S. Yu, “Asap: an adaptive sampling approach to data collection in sensor networks”, IEEE Transactions on Parallel and Distributed Systems , pp.1766-1783, 2007.

[19] L .A. Villas, D.L. Guidoni, R.B. Araujo, A. Boukerche, A.A. Loureiro, “A scalable and dynamic data aggregation aware routing protocol for wireless sensor networks”, Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems, MSWIM ’10, ACM, New York, NY, USA, pp. 110-117, 2010.

[20] S. Madden, M. J. Franklin, J. M. Hellerstein, W. Hong, “Tag: a tiny aggregation service for ad-hoc wireless sensor networks”, in WMCSA02, Proceedings of the Fourth IEEE Workshop on Mobile Computing, DC, USA, 2002, page 490.

[21] L. Villas, A. Boukerche, R. B. Araujo, “Highly dynamic routing protocol for data aggregation in sensor networks”, in Proceedings of the IEEE Symposium on Computers and Communication, ISCC10, IEEE Computer Society, Berlin, 2010, pp. 496-502.