Mobile Line Based Data Dissemination Protocol for Wireless Sensor Networks

Dalila Iabbassen
Faculty of Electronic and Computing Department,
University of Sciences and Technology HOUARI BOUMEDIENE (USTHB),
BP 32 El Alia 16111 Bab Ezzouar Algiers Algeria.
Email: iabdalila@gmail.com

Samira Moussaoui
Faculty of Electronic and Computing Department,
University of Sciences and Technology HOUARI BOUMEDIENE (USTHB),
BP 32 El Alia 16111 Bab Ezzouar Algiers Algeria.
Email: moussaoui_samira@yahoo.fr

***Abstract*—Depending on how to collect data and process them, we can distinguish the traditional Client/Server based data dissemination protocols, and the Mobile Agents protocols. In the traditional Client/Server architectures, data sources are transferred to a destination to be processed there, while in Mobile Agents architectures, a specific executable code passes through relevant sources to collect and process data. Mobile Agents can be used to significantly reduce the cost of communication, especially through low bandwidth, by moving the processing function to the data, rather that bringing the data to a central processor. This work proposes to use a Client/Server approach using Mobile Agents to aggregate data in a planar architecture of wireless sensor network.**

***Index Terms*—Wireless Sensor Networks, Energy Saving, Data Dissemination, Data Aggregation, Mobile Agents.**

I. Introduction

WSNs (Wireless Sensor Networks) has attracted a growing number of manufacturers. Indeed, the need for continuous monitoring of a given environment is quite common in various activities of society. Industrial processes, military tracking applications, habitat monitoring, and precision farming are just some examples of a wide and varied range of possible applications of continuous monitoring offered by WSNs. Unfortunately, WSNs are not perfect! Because of their low cost and their deployment in areas sometimes hostile, the motes are quite fragile and vulnerable to various forms of failure: breaking, low energy ... and so on. These problems make WSNs systems with innate fragility, which should be considered as a normal property of the network. This limitation of resources necessitates some form of cooperation on a large scale where interactions between sensors can be extremely complex. This requires the establishment of a protocol at the middleware layer for the dissemination and retrieval of data in an efficient manner.

The purpose of the data dissemination is to send any type of information (data or query) to all nodes affected by this information, while minimizing the number of transmission nodes and the energy cost [1]. Data dissemination is considered as a main phase of energy consumption in the communication of WSNs [2]. Hence the way to eliminate redundant data traffic and reduce communication costs are the main challenges of the data dissemination. On applications of WSNs such as environmental monitoring, data dissemination is very important. Indeed, it is known that the ineffectiveness of the data dissemination causes broadcast storms and blocks all data ommunications in the network [3].

In most WSNs, sensors are deployed in an area to extract environmental data. Once the data collected by multiple sources (multiple sensors in the vicinity of the captured event), they will be transmitted via multiple hops to a single destination (sink). This, coupled with the fact that the information collected by sensors neighbors is often redundant and correlated, and that energy is the most precious resource, requires the use of a data fusion mechanism. Instead of transmitting all data to be processed in a central node, the data is processed locally and only aggregate information is returned to the sink. Data fusion reduces the number of packets to be transmitted via the sensors, and therefore energy consumption and bandwidth. Its advantages are evident, especially in a large-scale network. So, we can distinguish two approaches of data dissemination in WSNs:

Client/Server based data dissemination: The sink sends queries to sensor nodes, each sensor node will then process the request and send the desired data individually to the sink which will process and aggregate it.

Mobile agent based data dissemination: The sink dispatches one or more mobile agents to sensor nodes. This agent will carry the code for data processing. In this way, the data will be aggregated and processed locally at

doi:10.4304/jetwi.5.1.4-11

the sensor nodes, then the agent will collect the data already processed to send it to the sink.

In the collaborative process, the client/server model is the most widely used, where the individual sensors (clients) send the raw or pre-treated data to a treatment center (server) and data integration is performed in the center (usually at the sinks or super nodes). As mentioned above, there are some drawbacks with this model that should be considered especially for WSNs. Indeed, despite the unreliability and the limited bandwidth of wireless links used in sensor networks, in addition to the constraint of energy that is critical in WSNs, the protocols using the client / server paradigm in this type of networks are numerous (DD [4], TTDD [5], LBDD [6]).

DD (Directed Diffusion) [4] is one of the most used approaches in many works on sensor networks. DD is a protocol for data dissemination using multiple paths for routing information. The sink broadcasts an interest in the form of a query in order to interrogate the network on a particular data, and then it determines the best path to use for receiving data.

In TTDD (Two Tier Data Dissemination) [5] each source node creates a virtual grid on the network. This grid is then used by the routing protocol to route queries and data between the source and the mobile sink.

II. Line Based Data Dissemination Protocol

LBDD (*Line-Based Data Dissemination protocol*) [6] defines a virtual structure formed as a band in the middle of the interest area. Nodes positioned inside of the band are called inline-nodes. This band represents an area of appointments for queries and data storage.

This protocol assumes that each node knows its geographical position and the coordinates of the interest area. LBDD uses then a geographic routing.

The operation of LBDD consists of two main steps:

- Dissemination: when a node detects a stimulus, the data is generated and sent to the nearest inline-node;

- Collection: to collect different data, the sink sends a perpendicular query to the band. The first inline-node which receives the query will propagate it in both directions of the band to attain the nodes with the required data. The data will be then sent directly to the sink.

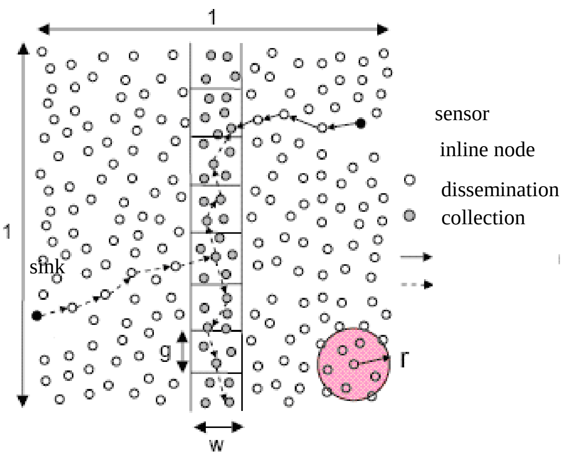

Figure 1. LBDD architecture.

Due to its simplicity and ease of deployment, the Client/Server paradigm is the most used in distributed environments especially in WSNs. However, this solution has been criticized for consuming high bandwidth and energy, and even big latency. For this, there was the appearance of a new paradigm using Mobile Agents.

A MA consists of a process code and a state (variable values, the next instruction ... etc.).. Initially a MA resides on the machine that created it. It is then dispatched to run on a remote host called generally a server. When an MA is dispatched, the entire code of the MA and its execution state is transferred to the host.

The host provides a suitable execution environment for the MA. The MA will use the host resources (CPU, memory ... etc..) to perform its task. After completing his work at the host, the MA will migrate to another machine. Since the execution state is transferred to the host, the MA can resume execution at the point where it stopped at the old host. In this way, the MA will continue its round until the last machine on its route, to finally return to the machine that created it.

The use of MAs in computer networks has advantages but also disadvantages, such as code caching and security in certain scenarios. Nevertheless, they are successfully used in various applications such as parallel programming, data collection, e-commerce and mobile computing. As described in [9], many inherent advantages (such as scalability and awareness of energy) of the MA architecture makes it more suitable for WSNs that the client / server architecture [11]. In fact, mobile agents can be used to significantly reduce the cost of communication, especially through low bandwidth, by moving the processing function to the data rather than bring the data to a central processor [7] [8] [10].

III. Mobile Agent based Wireless Sensor Networks

Based on the principle that the communication cost to send an information using a long message is usually less than sending the same information using several short messages, the MAWSN protocol (Mobile Agent based Wireless Sensor Networks) [7] will perform several concurrent tasks associated with small amounts of data by a single packet carrying multiple queries, and concatenate the results in a single package to reduce the communication cost.

The context of application and design of MAWSN highlights some assumptions:

- The sink knows all sources nodes to be visited by the MA.

- The itinerary of a AM is already designed before dispatching the MA.

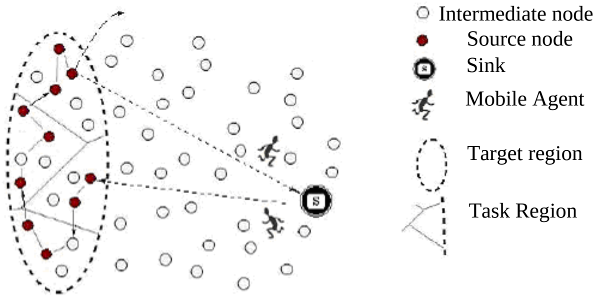

Figure 2. MAWSN architecture.

As In this approach, the sink sends queries to multiple targets simultaneously via a mobile agent. The data in the target region are collected by the mobile agent one by one, and all combined tasks are executed one after the other, so that the entire process will take longer. If the quality of service requirements (eg, latency) is not violated, especially if the target area is far from the sink, the energy gain from this execution combination could be important.

The MAWSN author offered another mobile agent architecture combined with Directed Diffusion, MADD (Mobile Agent based Directed Diffusion) [8] trying to eliminate the maximum of assumptions (choice of target nodes and establishment of the MA visiting sequence) of MAWSN via the combination of Directed Diffusion approach to the mobile agent approach.

By moving the processing code to the data, a MA can avoid the transmission of intermediate data in the network, continue working even in the presence of disconnections in the network, and then complete the entire task faster than the Client/Server traditional solutions.

However, a MA is not always better than a Client/Server solution. For example, if the code of the agent is greater than all the intermediate data, the MA will be less efficient in this case, since it will transfer more bytes over the network than the Client/Server solution.

In addition, if the network is fast, the agent will be less efficient even if the code is smaller. With a fast and reliable network, the interpretation of the agent at the sensors is slower than the transmission of intermediate data to the sink. However, if the network speed and reliability go down, or the data size increases, the case changes considerably.

Decide for a better approach (Client/Server or Mobile Agent) for wireless sensor networks is not a fixed choice, but it depends strongly on the network characteristics and topology and its application.

IV. Mobile Line based Data Dissemination

We propose an approach which uses a Mobile Agent in a mobile appointment area. Once deployed, sensors are expected to operate for extended period of time, and it

is impractical to physically reach all sensors [13]. The advantage in the appointment area is to avoid the blind search of the data and the sink in the interest area. For this, we will assign some nodes located in the appointment area, the role of data storage generated by all the sensors in the network. Thus, each source node that detects an event in the interest area, will generate the corresponding data and send it directly to the appointment area, where it is stored. Subsequently, when a sink node wishes to collect some information, it will generate a query that will also be directed directly to the appointment area, where all data reside. In this way, we avoid the problem of network congestion by eliminating the messages broadcasted to find the data across the network.

In order to collect efficiently the data stored in the appointments area, we will use a mobile agent whose mission is to aggregate the data collected in order to reduce the size of the packet transmitted to the sink and reduce the energy consumption and the bandwidth which are critical resources in wireless sensor networks.

Our approach is used for continuous applications where nodes periodically generate data, and also for event-driven applications where nodes generate data when an event arises in the field of capture. The proposed solution will perform multiple simultaneous tasks associated with small amounts of data carried by a single pack. The result will be concatenated into a single package to give a better aggregation in order to reduce the communication cost [7].

In order to build the appointment area, we will use a structure in the form of a rectangular band placed in the interest area. We assume that the network has a localization system, such as a GPS (Global Positioning System). In this way, each node is supposed to know its location and the coordinates of the interest area, which means that each node can know at any moment if it belongs to the appointment area.

In addition, the nodes must be synchronized in order to make the periodic change of the appointment area.

A. Mobility Of The Appointment Area

In our approach, the interest area is divided into N rectangular bands that do not interfere with each other and each band is also divided into several groups. At a given time, only one band i (i € [1, N]) will represent the appointment area for the data collected by the sensor nodes and sink queries. After the lapse of a period T, the band i+1 will be activated as a current appointment area instead of the band i, and so on. This process is executed periodically to balance the charge among all network nodes and therefore increase the lifetime of the network.

Once the zoning of appointment area is made, the clusterheads in each group must be specified and recognized. The clusterhead of each group must have enough energy to be able to receive the query of the sink and the descriptors (metadata) of data to be collected.

Hello messages of the appointment area will include the rate of energy remaining and the ability of free memory for each inline-node. This information will be used for the selection of clusterheads and nodes hosting the data.

To elect a clusterhead, we rely on the rate of energy remaining for each inline-node. The factor of free memory is not a parameter to be considered for the

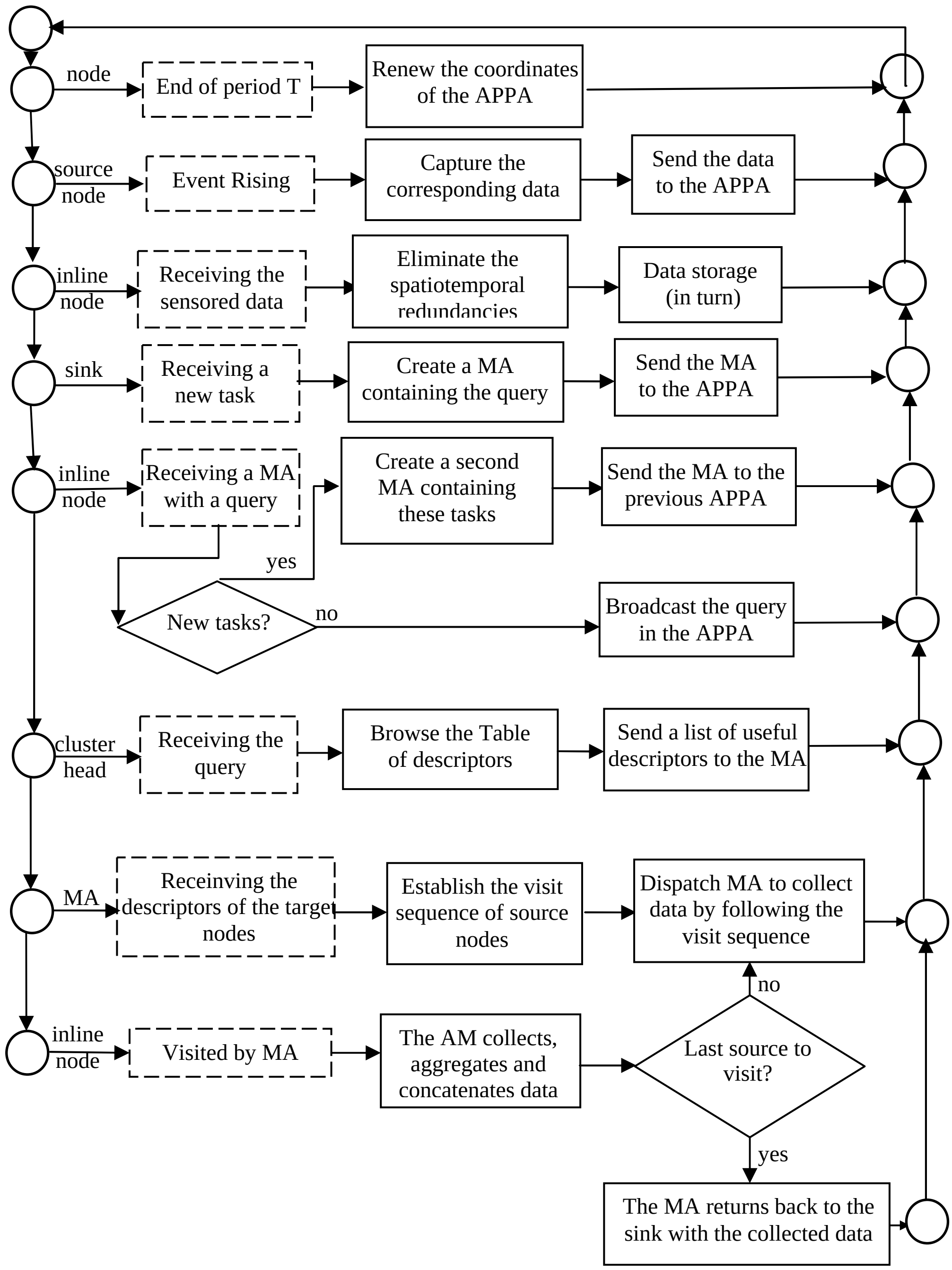

Figure 3. Flowchart of the MLBDD protocol.

election of clusterhead because all inline-nodes have no data stored at the time of changing the appointment area. Thus, when the change of the appointment area is made at a time T, the node having the max of energy in each group (the rate of energy information is transmitted in the Hello packets) will be set as a clusterhaed and will propagate this information in its group.

By using the localization system and the synchronization of nodes, changing the appointment area does not generate any additional messages. Indeed, as the nodes are synchronized, each of them at any time, is autonomously informed on the appointment area to be used for sending data and queries.

B. Dissemination Of Captured Data

When an event arises in the interest area, source nodes having captured it will generate the relevant data and its descriptor (metadata) and send them directly to the actual appointment area. The descriptor of a data contains the information needed to recognize the data by the interests of the sink (data type, time, ID, position..)

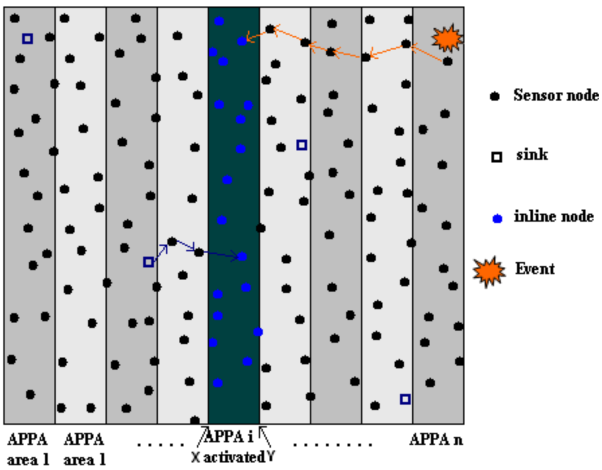

Figure 4. MLBDD architecture

Since the nodes are equipped with a localization system, data packets will be routed to the appointment area using a greedy geographic routing [12]. Each node will route the packet perpendicularly to the rectangular band. This will balance the charge between nodes and avoid congestion.

Data storage: To balance the charge between all the inline-nodes (nodes that belong to the current appointment area), data storage in the appointment area should not focus on specific inline-nodes.

When data arrives at the first inline-node, he will choose a neighbor node belonging to his group, having maximum remaining energy and enough space to house the captured data.

When the host node will receive the data, he will store it in its memory, and send its descriptor to the clusterhead of its group. In this way, each clusterhead has in its memory a table containing all the data descriptors of his group.

Data collection - The Mobile Agent itinerary: When a sink have to collect certain information from the network, it will create a mobile agent and send it directly to the activated appointment area, using the same geographic routing used to disseminate data to the rectangular band. The mobile agent contains the query of the sink and the processing code of the data to be collected (aggregation code). The MA will carry a package containing several queries [7], where each query corresponds to a different application.

When the clusterhead receives a MA, he spread the interest of the sink in both directions of the band to reach all the clusterheads in the appointment area. Each clusterhead receiving the interest will see its data descriptors table to get descriptors that match the interests of the sink and send them to the MA.

After receiving the descriptors, the MA will decide the target nodes to visit for data collection and will establish a logical sequence of visiting these nodes, based on their geographical position as follows (fig.5):

-Divide the rectangular band into two equal bands.

-For each sub-band, classify inline-nodes to be visited based on their geographical position, in order to draw a straight line from one end to another.

-Concatenate the two sub-sequences established for making the full sequence of the inline-nodes route. For routing from one node to another, we will also use the greedy geographic routing.

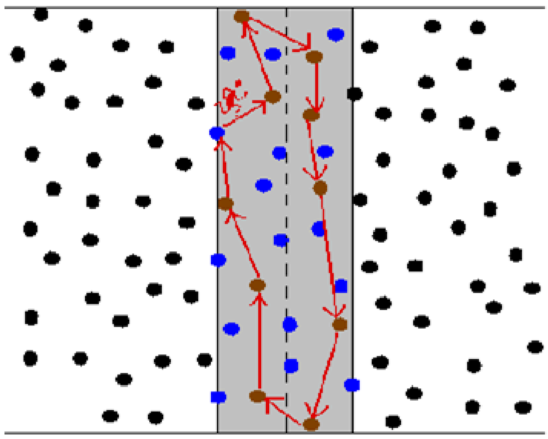

Figure 5. Mobile Agent visit sequence.

Data aggregation: Having established the visits sequence of inline-nodes, the MA will browse the nodes by aggregating data as it collects them [7]:

- Elimination of the application redundancy: With the development of WSNs, "a deployment, multiple applications" is a trend due to the nature of the specific applications of sensor networks. In general, given the constraints of storage capacity, it is impossible to store every application in the local memory on-board sensors. The sink attributes to the MA the processing code (behavior) based on the need of a specific application. The code carried by the MA requires local processing of raw data in the inline-nodes as requested by the application. This behavior allows a reduction of the amount of transmitted data by allowing only relevant information to be extracted and transmitted.

- Elimination of spatial redundancy: The degree of correlation of data collected between sensors is closely related to the distance between the sensors, so it is very probable that the sensors close to each other generate redundancies of collected data. So, the MA aggregates the individual data when it visits each inline-node. Although this aggregation technique is commonly used in protocols for data dissemination based on the clustering or aggregation tree, the aggregation assisted by MA requires no additional cost to build these special structures.

- Data concatenation: Based on the principle that the communication cost for sending a long message is usually less than sending the same amount of data using several short messages, we use a unification technique of packets, which concatenates the data from several small packets into a larger package to reduce communication costs. Due to the concatenation of the data, the communication cost of intermediate sensors can be reduced to increase the lifetime of the network. However, these energy savings can usually be obtained at the expense of a prolonged latency.

C. Perdiodic Change Of The Appointment Area

A sink node can send a query for data collection at any time to the appointment area. Since the appointment area is mobile and changes periodically, the system can fall if the query collection is generated just after the change of the appointment area. That is to say that the collection will be done in the new appointment area (which is the current appointment area), while the data still reside in the last appointment area and have not yet been collected.

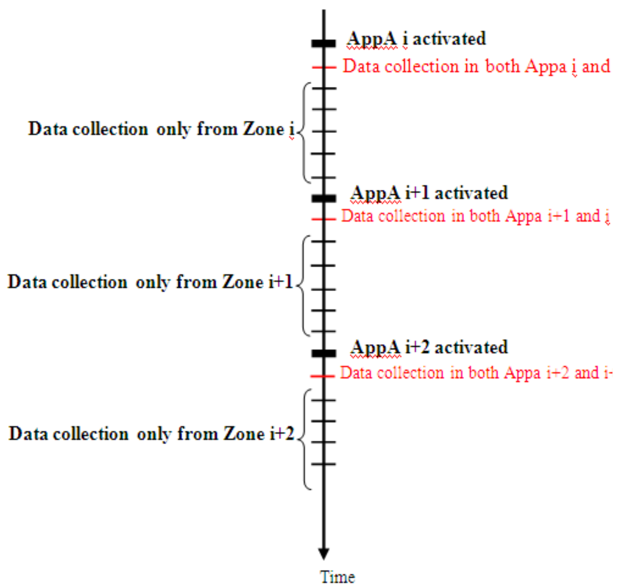

Figure 6. Data collection adapted to the change of the appointment area.

To avoid this, the sink will order the data collection in both appointment areas: the old area and the current area at the same time. This will be done only during the first collection of each data type for each appointment area changing. That is to say that at every periodic change of the appointment area, the collection of each type of data will be done in the two areas (current and last), but just after this first collection, the following data collections will normally be done only in the current appointment area (fig. 6.).

D. Performance Evaluation

To evaluate the performance of the MLBDD solution and compare it with other similar works, we performed simulations. To do this, we used the Glomosim simulator (Global Mobile Information System Simulator).

TABLE I.
SIMULATION PARAMETERS

Area dimensions	200m X 200
Nodes	400
Simulation period	15 minutes
Communication range for each node	37,67 meters
Nodes position	random
Period for the appointment area switching	3 minutes
Period of creating data	1 second
Period of creating sink requests	1 minute
MA eliminating redundancy rate	10%
MA fusion factor	1

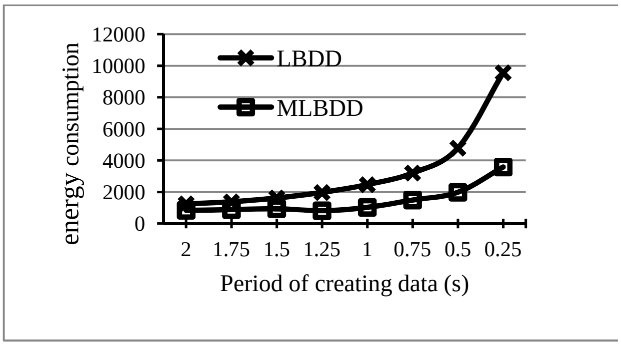

Figure 7. Energy / Data charge.

According to the graph above, we find that the energy consumption increases when we increase the frequency of data creation. This is due to the increased traffic generated by the dissemination and storage of data. However, the consumption of energy in LBDD is always higher than MLBDD; this is because LBDD disseminates data in the whole group in the area of appointments while MLBDD stores each data in a single node belonging to the appointment area.

We can also see in fig.8 that the energy consumption in the two approaches LBDD and MLBDD increases as the number of nodes in the network increases. Because, as more nodes there are as traffic increases between these nodes. However, LBDD energy consumption increased by 154,311 between 200 and 600 nodes, whereas it increased by only 87,371 in MLBDD. This is because LBDD uses data diffusion for storage while MLBDD uses diffuses queries and MA for the collection. And since data packets are usually larger than the controls messages and queries, the total LBDD energy is higher than MLBDD.

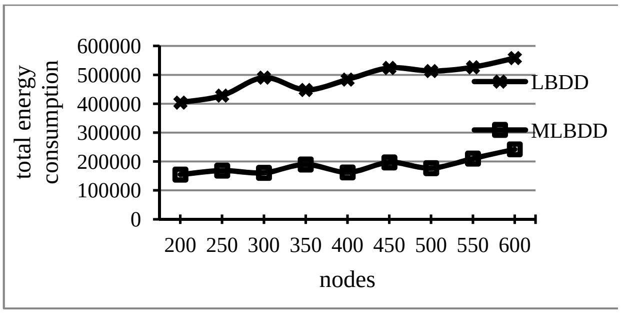

Figure 8. Energy / Scalability

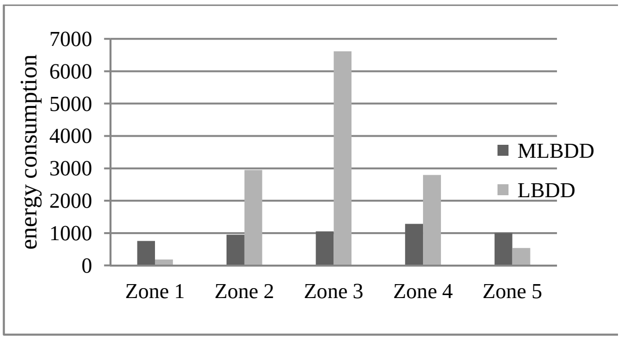

Figure 9. Energy /Zone structure.

The histogram above shows the energy consumption in each band of the interest area. Indeed, LBDD concentrates the charge in the middle area (since it uses a static oppointment area) while MLBDD spread the charge across all areas. This will directly affect the network lifetime. Indeed, the lifetime of the network in LBDD depends directly on the lifetime of the inline-nodes while in MLBDD lifetime does not depend on some special network nodes.

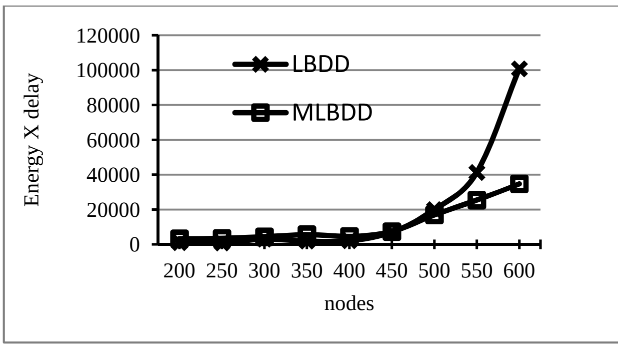

Figure 10. Energy X Delay /Scalability.

The graph above shows that when the number of nodes varies between 200 and 450, the graphs of LBDD and MLBDD are almost superimposed (LBDD is slightly better than MLBDD), while MLBDD is more effective than LBDD when the number of nodes exceeds 450. Therefore, MLBDD goes better than LBDD in scalability.

V. CONCLUSION

As the energy factor is one of the most important constraints that guide the design of protocols in wireless sensor networks, they must incorporate mechanisms that allow users to extend the lifetime of the entire network. For this reason, we developed the MLBDD approach (Mobile Line Based Data Dissemination) whose main objective is to extend the network lifetime, by using an equitable energy dissemination approach which limits the energy cost and communication while maintaining a balance of energy consumption between all nodes. MLBDD uses also a mobile agent [7] to collect data in the appointment area by using aggregate functions to reduce the amount of data transmitted and reduce the communication energy and bandwidth.

The simulation results showed that MLBDD improves greatly the energy consumption and balance the charge among all nodes in the network. Thus, MLBDD is suitable for applications requiring long lifetime where the field deployment is inaccessible as in the environmental monitoring applications ... etc..

REFERENCES

[1] J. Cartigny, F. Ingelrest and D. Simplot-Ryl., "Localized LMST and RNG based minimum energy broadcast protocols in ad hoc networks," *Ad Hoc Networks*, 1--16 (2005).

[2] J. Lu and F. Valois, "On the Data Dissemination in WSNs," In *the Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2007)* 0-7695-2889-9/07 (2007).

[3] S. Ni, Y. Tseng, Y. Chen, and J. Sheu, "The broadcast storm problem in a mobile ad hoc network," Proc. of *the International Conference on Mobile Computing and Networking (MobiCom), Seattle, USA*, 151--162, (1999).

[4] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," In Proceedings *of the 6th Annual ACM/IEEE (MOBICOM '00). 56--67, USA*, (2000).

[5] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, "A two-tier data dissemination model for large-scale wireless sensor networks," *MobiCom 02: Proceedings of the 8th annual international conference on Mobile computing and networking (New York, NY, USA)*, ACM Press. 148–159 (2002).

[6] E. Ben Hamida, A. Ziviani and M. Dias de Amorim, "Dissémination dans les réseaux de capteurs avec puits mobiles," In *9th International Symposium on Telecommunications Algorithmic*. 85--88 (INRIA-00176965, version 1), (2007).

[7] M. Chen, T. Kwon, Y. Yuan and V.C.M. Leung, "Mobile Agent Based Wireless Sensor Networks," *Journal of computers*, vol. 1, NO. 1, (2006).

[8] M. Chen, T. Kwon, Y. Yuan, Y. Choi and V.C.M. Leung, "Mobile agent-based directed diffusion in wireless sensor networks," *EURASIP Journal on Advances in Signal Processing*, Article ID 36871, (2007).

[9] Q. Hairong, Xu. Yingyue and W. Xiaoling, "Mobile-Agent Based Collaborative Signal and Information Processing in Sensor Networks." In *Proceeding of the IEEE*, Vol. 91, NO. 8, pp.1172-1183, (2003).

[10] E. Shakshuki, H. Malik and M.K. Denko, “Software agent-based directed diffusion in wireless sensor network,” *Telecommunication Systems* 38(3-4): pp.161-174. (2008).
[11] E. F. Aiello, G. Fortino and A. Guerrieri, “Using Mobile Agents as Enabling Technology for Wireless Sensor Networks,” *The Second International Conference on Sensor Technologies and Applications*, (2008).
[12] S. Medjiah, T. Ahmed, F. Krief and P. Gélard, “AGEM: Un Protocole de Routage Géographique Angulaire Adaptatif,” *French conference on the protocols engineering. Strasbourg, France*, CFIP (2009).
[13] W. Xiumin, W. Jianping and X. Yinlong, “Data Dissemination in Wireless Sensor Networks with Network Coding,” *EURASIP J. Wireless Comm. and Networking* 2010: (2010).