Global and Local Sensor Clustering using Genetic Algorithms and Fuzzy Logic

Mohammed Omari, Harma Abdelkarim and Benaich Salem
LDDI Laboratory, Computer Sciences Department, University of Adrar, Algeria
omari@univ-adrar.org, har.a.karim@gmail.com, salembenaich@yahoo.fr

***Abstract*—Wireless sensor networks (WSNs) attracted researchers attention in the last decade. In fact, most of the ad hoc routing protocols applied on WSNs were not efficient vis-à-vis the energy consumption at these tiny equipments. In this work, we proposed a new clustering method that has two levels of deciding the cluster head. The first phase (global clustering) is based on genetic algorithms where the clusters as well as the cluster head are defined by the base station. This operation is performed at the stage of deploying the network sensors, or when the need to reorganize the network topology due to low performance. The second phase (local clustering) is performed at each cluster by applying a token-ring like protocol where the cluster head is defined as the token holder. Our protocol aims to enhance the consumed energy by reducing the sensors communication traffic to the base station that is needed for global clustering.**

***Index Terms*—Global clustering, local clustering, wireless sensors, genetic algorithms, fuzzy logic sets, routing.**

I. Introduction

Wireless sensors networks (WSNs) have been developed quickly in the last decade due to the low cost of sensors and adaptability to physical environments [1]. The application of WSN consists of small sensor nodes that are low cost, low power and communicate within short distances [2]. Many clustering and routing protocols were proposed in the literature to assign a specific protocol to such small devices rather than the usual ad hoc protocols. Since energy consumption during communication is a major energy depletion factor, the number of transmissions must be reduced to achieve extended battery life [3][4].

Sensor networks bring up a wide range of new applications such as battlefield surveillance, biological detection, and environmental monitoring [5]. One approach to extend sensor network lifetime is to divide a sensor network into disjoint clusters of sensors where each cluster head must take in consideration all cluster members communications, compress data by deleting redundancy, and report it the to the base station. Many techniques were proposed to solve the clustering problem [6] which has been proved to be very hard (NP-complete).

Cluster based routing protocols are investigated in several research studies [2]. Some algorithms create clusters of uniform size such that the distance between the sensor nodes and the CHs is minimized [5]. With minimal distance, the cost of transmission energy is reduced at the local sensor stage. The LEACH protocol (Low Energy Adaptive Clustering Hierarchy) is widely used in the WSN domain [6]. It describes a hierarchical, self-organized cluster-based approach for sensor domain. In this protocol, the number of clusters is fixed, so that the cluster heads aggregate data from local sensors and transmit them to the base station. In order to avoid collisions, the sensors transmissions are multiplexed using the digital time division multiple access. The LEACH protocol suggests the “election” of a new cluster heads periodically. As an extension to LEACH, a multilevel clustering was proposed [7] based on sensors locations. The hierarchy of sensors is maintained to reduce the energy consumption of multiple hop routing.

The rest of this paper is organized as follows. In Section 2 we present a brief description of the genetic-algorithm version of LEACH. Our proposed model of global and local clustering is shown in Section 3 with a brief description of enhancing our technique using fuzzy logic membership functions. The conclusion and the future work are given in Section 4.

II. Application of Genetic Algorithms to the LEACH Protocol

A. *Brief Description of Genetic Algorithms*

Genetic algorithm (GA) is an adaptive method which is generally employed to solve search and optimization problems [8]. As shown in Fig. 1, the optimization process is done through 4 stages. At the first stage, an initial set of possible solution, or individuals, is randomly selected from the search space. Notice that genetic algorithms state that an encoding method should be considered to convert the possible solutions into chromosomes or a stream of codes. At the second stage, few individuals will be chosen as parents in order to produce the next generation. This selection process is based on a fitness function that measures the goodness of each solution. In fact, the quantitative measuring is related to how close is the possible solution to be optimal. At the third stage, a crossover operation between parent is performed by selection a point or more to “cut” the chromosome (or the code) of each parent and then “glue” parts from each parent to get a new individual, called an offspring. The new generation of offsprings does not assert a better fitness compared to their parent. Yet, the mimic of natural evolution by crossing over good entities may lead to produce some good solutions. The last stage corresponds to mutation phase, where few offsprings are selected to be altered at some part of their code, mimicking again mutation in nature.

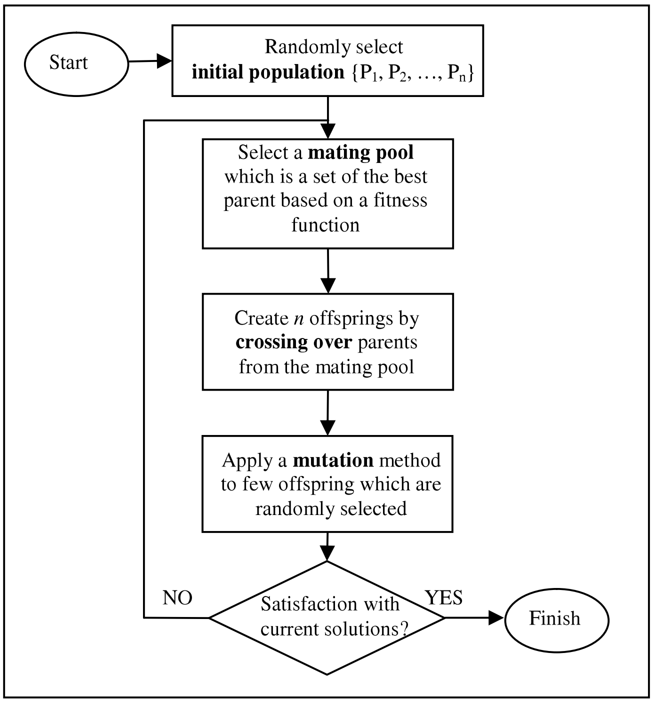

Fig. 1 : Flowchart of genetic algorithms phases

B. *Application of Genetic Algorithms to LEACH*

Liu and Ravishankar introduced a genetic algorithm-based variant of LEACH (LEACH-GA) to determine the optimal value of the user specified probability for various base station placements [9]. As shown in Fig. 2, each chromosome is coded as a series of bits that correspond to the network sensors. A sensor is coded 1 if it is considered as a cluster head, otherwise 0. At the beginning of the preparation phase, each sensor node initially determines whether or not it should be a candidate cluster head, using the following cluster head selection procedure:

1- Every sensor node selects a random number r from the interval [0, 1].
2- If r is smaller than certain threshold then the node is a candidate cluster head (CCH).
3- Then, each node sends its ID, location information, and whether or not it is a CCH to the BS.
4- As the BS receives messages sent by all nodes, it performs GA operations to determine the optimal probability by minimizing the total amount of energy consumption in each round.

The optimal probability is determined by the GA by searching the solution space through an evolutionary optimization process by applying selection, crossover and mutation operators. Once the optimal probability is found, the BS broadcasts its value to all nodes in order to form the corresponding clusters.

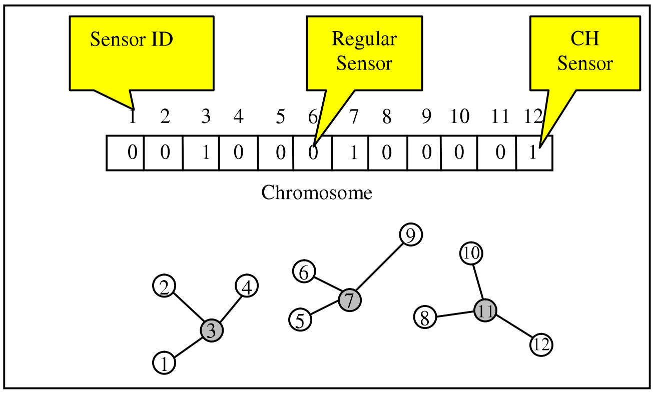

Fig. 2 : Enhanced LEACH clustering based on genetic algorithms

III. APPLICATION OF GLOBAL AND LOCAL CLUSTERING

We noticed that the back-and-forth transmission of sensors information to the base station in order to find best clusters in LEACH or LEACH-GA is very expensive vis-à-vis the sensors energy consumption. In fact, the LEACH protocol states that a reclustering is need periodically following the effects of previous transmissions.

In order to alleviate such a problem, we suggested the distribution of clustering burden over the entire network through two basic tasks: global and local clustering.

A. *Global Clustering*

It resembles the clustering of LEACH and LEACH-GA. Yet, we extended the LEACH protocol by applying fuzzy logic to divide the network into domains (as explained in Sub-section C). A domain is set by a threshold that defines the maximum number of sensor nodes inside each domain. Once the domains are defined based on sensors locations and their current energy, the base station runs a genetic algorithm with chromosomes that correspond to domains instead of the entire network. Hence, the generated chromosomes will be of short size compared to LEACH-GA. Moreover, when the network increases in size, more domains are defined, but the size of the chromosome in each domain will maintain the same maximum size.

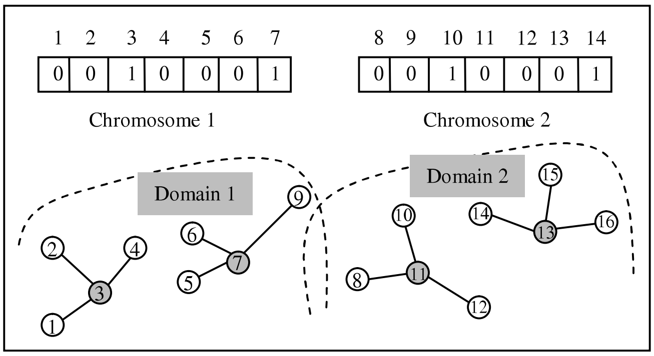

Fig. 3 : Dividing the network into domains to enhance genetic algorithms clustering

In order to enhance the time search, the network is divided into k domains: D_1, D_2, … D_k. Sensors belonging to different domains have low chance to be in the same cluster. For each domain, a GA generates a random set of chromosomes for each domain and starts the GA cycles as shown in Fig. 1 to find optimal clustering. Here is the fitness function used to rank chromosomes:

$$\text{Fitness}(C_k) = \frac{\sum_{j\in CHS} E(j)}{\sum_{i\in RSS} \text{StoH}(i) + \sum_{j\in CHS} \text{HtoB}(j)}$$

Where:

C_k: Chromosome *k*.

E(j): Current energy of a cluster head.

StoH(i): Energy to transmit one packet from a regular sensor *i* to the closest cluster head.

HtoB(j): Energy to transmit one aggregated-data packet to from a cluster head *j* the base station.

RSS: Regular sensors set corresponding to chromosome C_k.

CHS: Cluster heads set corresponding to chromosome C_k.

B. Local Clustering: Token Ring Topology

The base station selects a cluster topology after running a GA algorithm. The base station should decide sort of a ring topology inside each cluster where each sensor knows its successor. Therefore, when the clusters are formed, the initial cluster head must hold a token in order to aggregate data and send it to the base station. Thereafter, it passes the token to the next node to assign it as a new cluster head (Fig. 4).

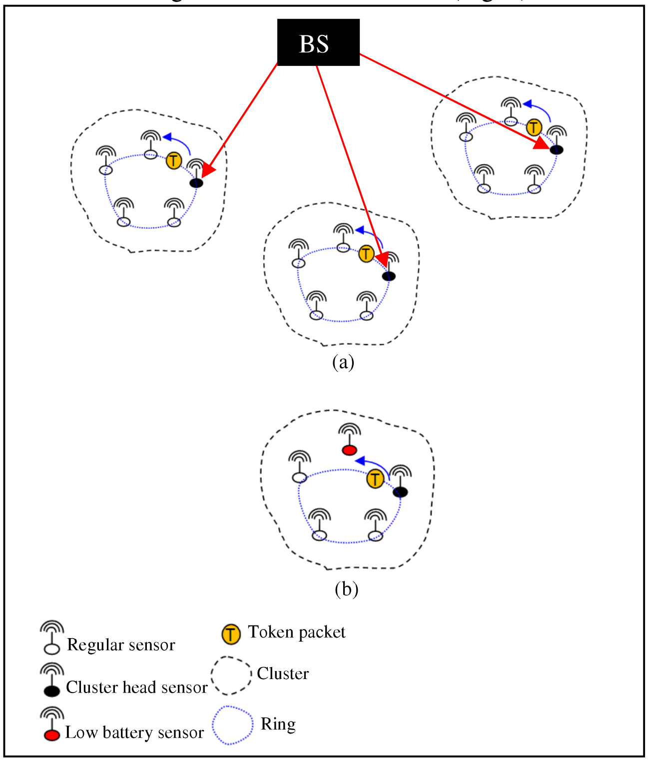

Fig. 4 : Local cluster head switching (a) normal token forwarding (b) escaping low battery nodes

When a node gets isolated after all cluster nodes are of low energy (the token cannot be passed to other node), it transmits its current info to the base station for new topology and clustering assignment. On the other hand, when a node is about to be down, it sends an alerting packet to its predecessor in order to inform it to forward the token in the future to the current next sensor in line.

C. Domains Division using Fuzzy Logic Membership Functions

Membership functions were introduced by Zadeh in the first paper on fuzzy sets (1965) [10].The membership function of a fuzzy set is a generalization of the indicator function in classical sets. In fuzzy logic, it represents the degree of truth that indicates the degree of likelihood to belong to a set or a domain. Our goal is to utilize a membership function that help us constructing domains when a degree of a sensor membership goes beyond 0.5.

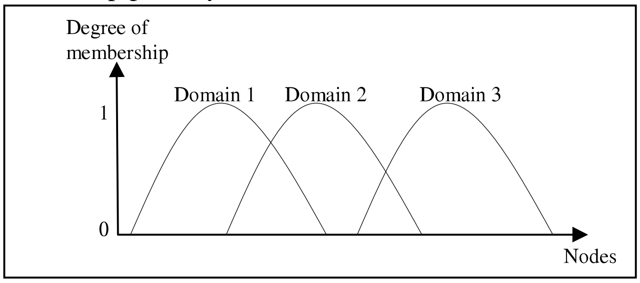

Fig. 5 : Solving the sensor domain memberships using fuzzy logic membership functions

IV. Conclusion and Future Work

Scalable routing protocols that are based on hierarchical clustering are needed to govern WSNs. In fact, most of the routing protocols used in WSNs were built for ad hoc networks where energy is not considered as an important factor. In this paper we introduced a new method that combines both global and local clustering in order to enhance energy consumption at the sensor level. The sensors network is first divided into many domains based on a membership function that clears the fuzziness between domains. Then, the base station run a GA with each domain separately to find better optimized clustering. Thereafter, the cluster head passes a token to its immediate neighbor in the cluster to exchange roles, and thus saving cluster mean energy. As a future work, we will investigate the implementation of a fuzzy logic membership function and enhance its time response through genetic algorithms.

REFERENCES

[1] T. Y Wu, I. J. Liao, W. T. Lee, G. H, Liaw, J. W Ding, C. C. Wu, "Enhancing Indoor Localization Accuracy of Sensor-based by Advance Genetic Algorithms," IWCMC'10, June 28–July 2, 2010.

[2] S. Hussain, A. W. Matin, O. Islam, "Genetic Algorithm for Hierarchical Wireless Sensor Networks", Journal of Networks, vol. 2, no. 5, September 2007.

[3] V. Mhatre, C. Rosenberg, D. Kofman, R. Mazumdar, and N. Shroff, "A minimum cost heterogeneous sensor network with a lifetime constraint," IEEE Transactions on Mobile Computing (TMC), vol. 4, no. 1, pp. 4 – 15, 2005.

[4] N. Trigoni, Y. Yao, A. Demers, J. Gehrke, and R. Rajaramany, "Wavescheduling: Energy-efficient data dissemination for sensor networks," in Proceedings of the International Workshop on Data Management for Sensor Networks (DMSN), in conjunction with the International Conference on Very Large Data Bases (VLDB), August 2004.

[5] Chih-Chung Lai Chuan-Kang Ting Ren-Song Ko, "An Effective Genetic Algorithm for Improving Wireless Sensor Network Lifetime," GECCO '07, July 7-11, 2007. London, England, United Kingdom.

[6] M. Cardei and D. Z. Du. "Improving wireless sensor network lifetime through power aware organization," Wireless Networks, pages 333–340, 2005.

[7] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in Proceedings of the IEEE Conference on Computer Communications (INFOCOM), 2003.

[8] M. Melanie, "An Introduction to Genetic Algorithms", a Bradford Book The MIT Press, 1999.

[9] Jenn-Long Liu and Chinya V. Ravishankar, "LEACH-GA: Genetic Algorithm-Based Energy-Efficient Adaptive Clustering Protocol for Wireless Sensor Networks," International Journal of Machine Learning and Computing, vol.1, no. 1, April 2011.

[10] L. A. Zadeh. "Fuzzy sets and systems," System Theory. Brooklyn, NY, Polytechnic Press, 1965: 29–39.