Single-Node Cluster Reduction in WSN and Energy-Efficiency during Cluster Formation

Chérif Diallo, Michel Marot, Monique Becker

SAMOVAR CNRS Research Lab – UMR 5157; Dept Réseaux et Services de Télécommunications (RST)
Institut TELECOM; TELECOM SudParis; 9, Rue Charles Fourier – 91011 Evry CEDEX, France
Email:{cherif.diallo, michel.marot, monique.becker}@telecom-sudparis.eu

***Abstract*—A large Ad Hoc network can be represented as several sets of clusters. Each cluster contains one or more nodes and has its clusterhead (or caryomme) chosen following an election based on an appropriate criterion. The clustering in wireless networking adds scalability, reduces the computation complexity of routing protocols, allows data aggregation and then enhances the network performance. Several studies have focused on clustering algorithms, providing some mechanisms well suited to wireless sensor networks (WSN). The MaxMin clustering algorithm proposed by [1] was generalized, corrected and validated in [2],[3] which added the use of criteria functions to choose appropriate clusterheads. We take part in this work by examining comparative results of different criteria that reveal single-node cluster phenomena. A single-node cluster is a cluster of which the only member is its clusterhead itself. In WSN clustering, the density of single-node clusters is a performance criterion. We use MaxMin in a cold chain monitoring application which shows that single-node clusters have negative impacts on the energy consumption. In this paper we compare several criteria including the "remaining energy" and the "link quality indicator (LQI)". We propose an effective manner of using the LQI which is among the best performant criteria. We also propose a simple mechanism to reduce single-node cluster phenomena which highly enhances the energy efficiency.**

***Index Terms*—Wireless Sensors Networks (WSN); Clustering; MaxMin; Single-node Cluster; Link Quality Indicator (LQI).**

I. INTRODUCTION

In a cold chain monitoring application, due to the size of a warehouse which hosts large numbers of pallets, provided each with a temperature sensor, the WSN can reach several hundreds of nodes which collaborate for sending alarms towards the base station (BS). This application specifically collects rare events (alarms) to ensure the proper monitoring of the system. If the temperature is over a threshold, an alarm will be generated; this "interesting event" is then sent towards the BS. In such a context, network clustering techniques add scalability feature and then reduce the computation complexity of data gathering and routing protocols. MaxMin [1] is a popular clustering algorithm which has been subjected to numerous publications. Originally, it used the node address (node ID) as the criterion for selecting caryommes (clusterheads). This algorithm was generalized, corrected and validated in [2],[3] which proposed a theoretical study of the multihop clustering in WSN; [2],[3] added the use of criteria functions to select caryommes. In this paper, we take part in this work by studying more general caryomme selection criteria, such as: the Remaining energy level, the Degree of connectivity, the Proximity with respect to the base station, the Link Quality Indicator (LQI), and a hybrid criterion composed of any pairs of these above criteria. The use of the LQI as defined in the ZigBee standard [4],[5] is less efficient. So we propose an effective manner of the LQI use which reduces the energy consumption. In network clustering, a single-node cluster is a cluster of which the only member is the clusterhead itself. By running the MaxMin algorithm for a simple WSN example, we notice a significant density of single-node clusters. It is interesting to study this phenomenon since a less effective criterion leads to a higher density of single-node clusters and vice versa. Obviously, a non-clustered network can be considered as a clustered one having as many single-node clusters as the network size, hence the importance of having a low density of single-node clusters. Two main steps compose clustering schemes: the caryomme selection followed by the cluster construction. After the caryomme selection, the canonical method allows a step by step construction of the clusters in which the non-clustered nodes select their caryomme upon the reception of its announcement in the d-neighborhood. The canonical method is used by Delye in [2]. In this paper, we optimize the canonical method by proposing a simple mechanism which reduces the density of single-node clusters. Applying the MaxMin algorithm to a cold chain monitoring application, we show how high densities of single-node clusters could have negative impacts on the network performance especially on the energy efficiency. In this application MaxMin is used to select the caryommes which manage their respective clusters upon a TDMA based organization. Regular sensors send alarms to their respective caryommes which aggregate them and then forward data towards the BS using the "Link Reliability based Routing Protocol" (L2RP) we have proposed in [6]. L2RP is run with the weighted round robin load balancing mechanism using the "MinLQI" metric. The rest of this paper is organized as follows. After an overview of the related works in the next part, the next one gives a brief summary of the generalized MaxMin algorithm proposed by [2] and [3]. In the fourth part, a simple WSN example allows us to better understand the concrete single-node cluster phenomenon produced by MaxMin. Then, we present the studied clusterhead selection criteria (in the fifth section) before presenting the proposed single-node cluster reduction mechanism in the sixth section.

The last section presents simulation results pertaining to a cold chain monitoring application (presented in the seventh section).

II. State of the art

This section takes advantage of valuable works done by Marot in [3] which give a complete overview of recent advances in WSN clustering protocols.
In some clustering protocols nodes become caryommes after a randomized timer (CLUBS in [7] and RCC in [8]). However, the criterion which commonly determines the choice of the caryomme is the node address ([9], [10], ACE-C and ACE-L in [11]), the remaining energy level (LMSSC in [12]) or the degree of connectivity (such as MECH in [13]). It can also be based upon a weight basis of several of these above criteria (cf. DWEHC in [14], HEED in [15], WCA in [16]). In LEACH [17], clusterheads are selected under a criterion built on a probability function. In [18], Klaoudatou and al. were interested in a WSN application designed for medical surveillance, in which they select the closest node to the BS as caryomme. Multihop clustering algorithms address two main challenges: first, the question arises how to optimally choose caryommes and, secondly, how to construct the parental relationship between regular nodes and their caryommes in such a way that any common sensor can reach its clusterhead within k hops, that is to say, how to construct an optimally independent k-dominating set. Unfortunately, finding such a set is an NP-complete problem [1], so some heuristics have been proposed. In [19], Dai and Wu proposed three algorithms to build k-dominating sets which are also k-connected. One method is to compare sensor values (criteria) such as node ID, "remaining energy level", weight, etc. (cf. KHOPCA in [20], CABCF in [21], MaxMin in [1]). It is possible that two nodes have the same criterion value. Then, in [22] and [2], the authors propose to consider the pair composed by the degree of connectivity of the node and its address. Some works ([15],[23],[24],[25],[26]) introduced the notion of single-node clusters as a performance criterion for evaluating clustering protocols, whithout showing why single-node clusters are not a desirable feature in WSN. HEED [15] estimates the percentage of non single-node clusters to prove the protocol effectiveness. Furthermore, in the state of the art, it lacks for a comparative study of different criteria that can be used by the MaxMin algorithm. Also, the related works do not take into account a possible use of the link quality indicator (LQI) as a criterion for selecting caryommes. [2] and [3] offer a valuable theoretical generalization of MaxMin by presenting results on the degree of connectivity, without addressing the issue of single-node clusters produced by MaxMin. In this paper, we study the single-node cluster phenomenon produced by MaxMin and compare the impact of the following clusterhead selection criteria: "remaining energy level", "degree of connectivity", "proximity with respect to the base station", "link quality indicator (LQI)". We propose an effective manner of using the LQI which is among the best performant criteria. We also propose a simple mechanism for reducing single-node clusters. An application designed for a cold chain monitoring system shows the negative impacts of single-node clusters on the network performance.

III. The generalized form of MaxMin Algorithm

This algorithm is proposed by [2],[3] as a generalization of the earlier MaxMin algorithm proposed by [1]. It takes place in 2d+1 rounds. The first round consists of information exchanges to initialize the algorithm. The following d-rounds are the floodmax phase, which is followed by the floodmin phase composed of last d rounds.
The WSN can be modeled as a graph $G = (V, E)$, where two nodes are connected by an edge if they can communicate with each other. Let $x \in E$ be a node in the WSN. $N_1(x)$ is the neighbourhood of the node x. Let ν be a bijective function defined in E which is a totally ordered set.
Initial phase: $k = 0$,

$$\forall\, x \in E,\; W_0 = \nu(x),\; S_0(x) = x \tag{1}$$

Floodmax phase: $k \in [|1, d|]$,
Assuming that $\forall x \in E$, $W_{k-1}(x)$ and $S_{k-1}(x)$ are known in a previous step. Let $y_k(x)$ be the unique node in $N_1(x)$ defined by:

$$\forall\, y \in N_1(x) \setminus \{y_k(x)\}\,,\; W_{k-1}(y_k(x)) > W_{k-1}(y) \tag{2}$$

W_k and S_k are calculated as follows:

$$\forall\, x \in E,\; W_k(x) = W_{k-1}(y_k(x)),\; S_k(x) = y_k(x) \tag{3}$$

Floodmin phase: $k \in [|d+1, 2d|]$,
Assuming that $\forall x \in E$, $W_{k-1}(x)$ and $S_{k-1}(x)$ are known in a previous step. Let $y_k(x)$ be the unique node in $N_1(x)$ defined by:

$$\forall\, y \in N_1(x) \setminus \{y_k(x)\}\,,\; W_{k-1}(y_k(x)) < W_{k-1}(y) \tag{4}$$

W_k and S_k are calculated as follows:

$$\forall\, x \in E,\; W_k(x) = W_{k-1}(y_k(x)),\; S_k(x) = y_k(x) \tag{5}$$

The set S of clusterheads is defined by :

$$S = \{\, x \in E,\; W_{2d}(x) = \nu(x)\} \tag{6}$$

IV. Single-Node Cluster Phenomenon

We run the MaxMin algorithm on a WSN example, where the remaining energy of sensors is used as the criterion for selecting caryommes. So in this example the ν function is defined as follows:

$$\forall\, x \in\; E,\;\; \nu(x) = (f(x), id(x)) \tag{7}$$

Where $id(x)$ returns the address of the node x. The total ordering in E is defined as follows: $\forall x \in E, \nu(x) > \nu(y) \Longleftrightarrow (f(x) > f(y))\; or\; (f(x) = f(y)\; and\; id(x) > id(y))$
In this example, we consider the WSN in Fig. 1 where the neighbourhood relationships between the 8 sensors are defined by the hedges of the graph.

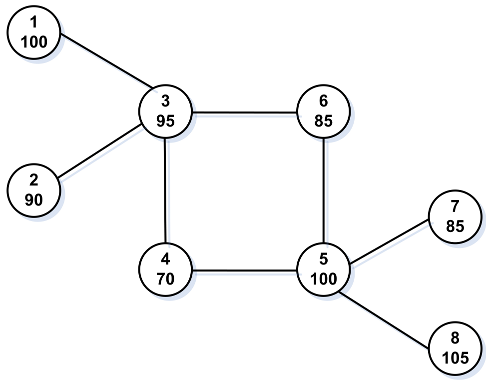

Fig. 1. The WSN used in the MaxMin examples (Tables I,II)

TABLE I
RUNNING MAXMIN ALGORITHM (D=1) FOR THE WSN IN FIG. 1

node id	1	2	3	4	5	6	7	8
W_0	**100**	90	**95**	70	**100**	85	85	**105**
	1	2	**3**	4	**5**	6	7	**8**
S_0	1	2	3	4	5	6	7	8
W_1	100	95	100	100	105	100	100	105
	1	3	1	5	8	5	5	8
S_1	1	3	1	5	8	5	5	8
W_2	**100**	95	**95**	100	**100**	100	100	**105**
	1	3	**3**	1	**5**	1	5	**8**
S_2	1	2	2	3	4	3	7	5

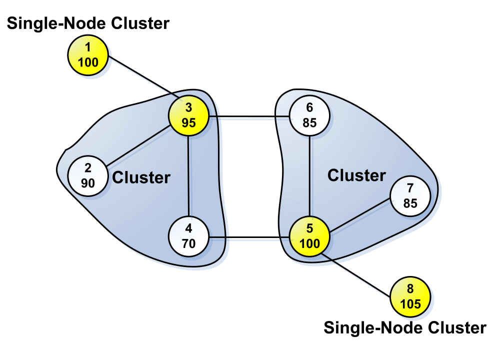

Fig. 2. Clusters Produced by the MaxMin Algorithm (d = 1)

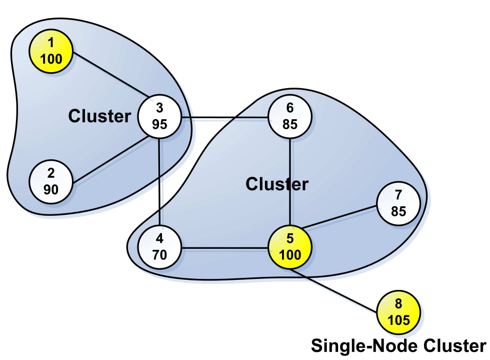

Fig. 3. Clusters Produced by the MaxMin Algorithm (d = 2)

TABLE II
RUNNING MAXMIN ALGORITHM (D=2) FOR THE WSN IN FIG. 1

node id	1	2	3	4	5	6	7	8
W_0	**100**	90	95	70	**100**	85	85	**105**
	1	2	3	4	**5**	6	7	**8**
S_0	1	2	3	4	5	6	7	8
W_1	100	95	100	100	105	100	100	105
	1	3	1	5	8	5	5	8
S_1	1	3	1	5	8	5	5	8
W_2	100	100	100	105	105	105	105	105
	1	1	5	8	8	8	8	8
S_2	3	3	6	5	8	5	5	8
W_3	100	100	100	100	105	100	105	105
	1	1	1	5	8	5	8	8
S_3	1	2	1	3	4	3	5	5
W_4	**100**	100	100	100	**100**	100	105	**105**
	1	1	1	1	**5**	1	8	**8**
S_4	1	2	1	3	4	3	5	5

This MaxMin example, ran for d = 1, produces 4 caryommes (1,3,5,8). The clusters represented by caryommes 1 and 8 are single-node clusters (Fig. 2).

The same WSN computed in MaxMin algorithm, run for d = 2, produces 3 caryommes (1,5,8). The cluster represented by the clusterhead 8 is a single-node cluster (Fig. 3). As we can see on the example above, the single-node cluster phenomenon is a real fact. This phenomenon is especially important, because the density of single-node clusters (50% and 33%) produced by MaxMin is not negligible. We are now interested in studying this phenomenon by comparing the criteria presented in the following section

V. CLUSTERHEAD SELECTION CRITERIA

A. *Remaining energy level*

The remaining energy of a sensor could be a criterion for selecting caryommes since a node with a better battery life seems to be a better candidate for the cluster management and the data aggregation. Conversely, if a sensor with low power is selected as a clusterhead, this can lead to packet losses because it might not have enough batteries to aggregate the received events; it also might not be able to send the aggregated data towards the BS. In this paper, we consider that each node knows its energy level.

B. *Sensor proximity with respect to the base station (BS)*

We consider a WSN deployed with a base station where each node knows its exact location and that of the BS. As the main goal of the application is to send events towards the BS, it seems natural to look at the criterion defined as follows:

$$C_i = 1/d(S_i, BS) \tag{8}$$

Where $d(S_i, BS)$ is the distance seperating the sensor S_i from the BS. We choose the inverse of the distance to promote the election of the closest sensor to the BS.

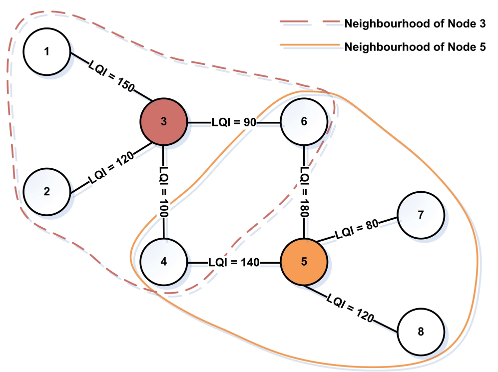

Fig. 4. Example of a WSN with LQI values of the links

C. Degree of connectivity

The degree of connectivity of a node, i.e. the number of its neighbors, is also a criterion that seems interesting to study. Intuitively, the more neighbors a sensor has, the more it seems to be an appropriate candidate as clusterhead, since a sensor with a low degree of connectivity might have little information, from its neighbourhood, to aggregate and to forward to the BS. In the initial phase, each sensor is involved in the neighbourhood information exchanges (hello protocol), which allows it to determine its degree of connectivity and the BS location.

D. Link Quality Indicator (LQI)

In the Zigbee standard [4],[5], the LQI measurement is defined as a characterization of the strength and/or quality reception of a packet. The use of the LQI result by the network or the application layers is not specified in [4],[5]. The LQI measurement is performed for each received packet, and the result is reported to the MAC sublayer as an integer ranging from 0 to 255. The minimum and maximum LQI values (0 and 255) are associated with the lowest and the highest quality IEEE 802.15.4 reception detectable by the receiver, and the LQI values in between are distributed between these two limits [4], [5].

For moteiv′s Tmote Sky [27] sensors equipped with chipcon′s CC2420 [28], the LQI values range from 50 to 110. Even so, we stick with the ZigBee standard [4],[5]. Then, we use standard values (i.e. 0, 255), instead of those of CC2420. In this paper, we define three LQI based clusterhead selection criteria: AvgLQI, MaxLQI and MinLQI. The AvgLQI is the average calculated from the LQI values of all the links between the node and its neighbors. AvgLQI values give a characterization of sensors throughout their respective coverage quality. This criterion might be useful in the context of the WSN deployed in a warehouse which hosts a large number of pallets, one upon the other. Such an environment is proned to high unreliability of wireless links. The MaxLQI criterion is the maximum LQI value which matches to the standard definition of the LQI [4],[5] used in the MultiHopLQI routing algorithm [29],[30]. As for the MinLQI, it is the minimum value beyond the given LQI threshold. For example (Fig. 4), assuming that the LQI threshold for an acceptable link quality is 100, the MinLQI for node 5 is 120 (LQI of link 5-8) instead of 80 (LQI of link 5-7). Thus, Table III gives LQI values for the WSN in Fig. 4.

TABLE III
LQI CRITERIA VALUES RELATED TO THE WSN IN FIG. 4

Sensor ID	1	2	3	4	5	6	7	8
AvgLQI	150	120	115	120	130	135	80	120
MaxLQI	150	120	150	140	180	180	80	120
MinLQI	150	120	100	100	120	180	80	120

E. Composite criteria (Hybrid)

In this paper, we define the composite criteria (hybrid) as follows:

$$Hybrid(LQI, C_i) = \alpha * LQI + (1 - \alpha) * Sc(C_i) \quad (9)$$

$$Hybrid(C_i, C_j) = \alpha * Sc(C_i) + (1 - \alpha) * Sc(C_j) \quad (10)$$

where $Sc(C_i)$ is a scale function which returns remaining energy values comparable to LQI values. It helps avoiding the composite criteria to be strongly influenced by the C_i component in equation (9):

$$Sc(C_i) = \beta + \frac{\Psi * \log(1 + (C_i - C_{i,min}))}{\log(1 + C_{i,max})} \quad (11)$$

Where C_i is a criterion, $C_{i,min}$ (resp. $C_{i,max}$) is the minimum (resp. maximum) value of C_i. If C_i is the remaining energy of the node, $C_{i,min}$ represents the value under which, the sensor is considered dead (battery depletion); while $C_{i,max}$ is the initial amount of energy provided with each sensor. $\beta = 50$, $\psi = 255$.
Like the LQI criteria definitions, we can also define AvgHybrid, MaxHybrid and MinHybrid criteria depending on whether, we are respectively considering AvgLQI, MaxLQI and MinLQI as defined in Table III.

VI. SINGLE-NODE CLUSTER REDUCTION (SNCR)

After the selection of caryommes using the generalized form of the MaxMin algorithm, clusters are then built, according to the following mechanism. This mechanism is compared, later in simulations, with the canonical method presented in [2],[3].

- At the end of the MaxMin floodmin phase, each caryomme initializes a timer W_T inversely proportional to its degree of connectivity.
- At the expiration of the timer W_T, the first caryomme which has the highest degree of connectivity, informs its neighbourhood that it is a selected clusterhead by sending a "CH-INFORM-MSG" packet.
- The "CH-INFORM-MSG" message contains the caryomme ID $id(CH_i)$ and has a time-to-live equal to d (the same d as in the MaxMin algorithm). It is retransmitted to all nodes within d-hops from the originating clusterhead.
- Upon reception of a "CH-INFORM-MSG" message, each neighbor which has not yet chosen a clusterhead, chooses

the sender as caryomme, decrements the TTL and then forwards the "CH-INFORM-MSG" message to its neighbors.

- Upon reception of a "CH-INFORM-MSG" message, by another clusterhead, it creates a list "SRC-INFORM-MSG" of senders which contains node IDs $id(S_i)$ of sensors from which the message is received. This node ID is not necessarily the one of the originating caryomme.
- The caryommes send their "CH-INFORM-MSG" messages in descending order of their respective degree of connectivity, until all caryommes have announced their state.
- A clusterhead of single-node cluster recognizes itself by the fact that its "CH-INFORM-MSG" message is not retransmitted by any of its neighbors. Such a clusterhead inspects its "SRC-INFORM-MSG" list. If its list is not empty, it chooses the first node of its list as caryomme. If its "SRC-INFORM-MSG" list is empty, then the clusterhead has no neighbor. This denotes a single-node cluster arising from the topology deployment of the WSN.

The waiting time W_T is calculated according to the following formula:

$$W_T(CH_i) = \tau + \frac{\zeta}{1 + \log(1 + \delta_i + \frac{id(CH_i)}{\Gamma} * \delta_i)} \qquad (12)$$

Where τ and ζ are small nonzero positive constants, δ_i is the degree of connectivity of the clusterhead CH_i and $id(CH_i)$ returns its address. Γ is a constant which is larger than the network size ($\Gamma = 10^6$, for example). This timer function avoids collisions between caryommes having the same degree of connectivity. If $\delta_i = 0$, then the clusterhead CH_i has no neighbor. This also denotes a single-node cluster arising from the topology deployment of the WSN.

VII. Cold chain monitoring application

A. Network organization and deployment

The application is designed for a cold chain monitoring purpose. Its goal is to monitor a warehouse by logging alarms originating from sensors. Alarms are generated when the sensed temperature exceeds a given threshold. After a first phase consisting of hello exchanges, the MaxMin clusterization algorithm is run. Then, each caryomme manages a TDMA organization (Fig. 6) by assigning one slot time ($T_{Slot}(S_i)$) to each one of its cluster members. To save energy, sensors switch in "sensing mode" and turn off their respective radios while leaving sensor modules in the active mode in order to continue collecting events. Then, in the data collection phase ($T_{Data} = 1s$) sensors wake up in turn upon their respective time slots in which each sensor sends its alarms to its respective caryomme. Since, the caryommes do not necessary form a connected backbone, all sensors wake up during the routing phase (Fig. 6) in which each caryomme aggregates the received alarms and then sends towards the BS. In the routing phase, only caryommes are sources of data packets. Other regular sensors are only participating in the routing effort by retransmitting received data towards the BS. The routing

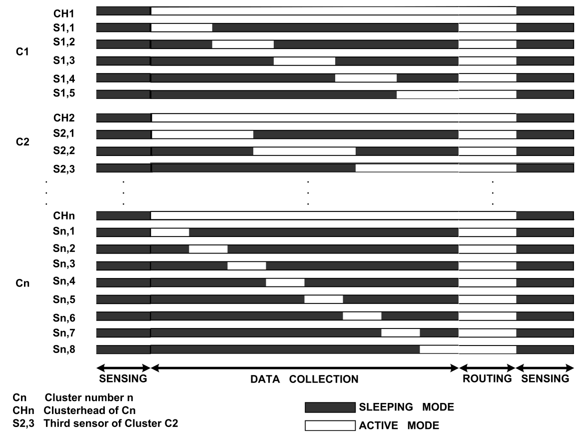

Fig. 5. Active/Sleep mode organization of the WSN

protocol used is the "Link Reliability based Routing Protocol" (L2RP) we have proposed in [6]. It is run with the weighted round robin load balancing mechanism using the "MinLQI" metric. The routing phase ($T_{Routing} = 1min$) is followed by a long sensing one ($T_{Sleep} = 8min59s$). With these time values, the total duration of a complete cycle (Sensing, Data collection, Routing) is $T_{Cycle} = 10min$. The assigned time slots to each regular clustered sensor are computed as follows:

$$T_{Slot}(S_i) = \frac{T_{Data}}{\eta_i} \qquad (13)$$

Where η_i is the number of regular sensors which are in the same cluster as S_i.

In the simulation model N sensors are randomly deployed over an area of length L=100m, and width l=100m. The base station is located at the (0,0) location. Each node generates alarms, which are sensed data over than the temperature threshold $Temp_{min}$, following the Poisson process of parameter $\lambda = 1$. The transmission range of each sensor (including the BS) is R = 20m. Each node knows the value of its remaining energy level, its location and that of the BS. At the beginning of the network deployment, the BS broadcasts a message containing its location. This information is then retransmit to all sensors in the network. In this phase, the degree of connectivity and the initial LQI values are calculated.

B. Energy consumption model

As in [31] and [32], let $E_{Tx}(k, d)$ the energy consumed to transmit a k bits message over a distance d [17]:

$$E_{Tx}(k, d) = E_{elec} * k + \varepsilon_{amp} * k * d^2 \qquad (14)$$

Let E_{Rx} the energy consumed to receive a k bits message:

$$E_{Rx}(k, d) = E_{Rx-elec}(k) = E_{elec} * k \qquad (15)$$

$$E_{elec} = 50nJ/bit \; and \; \varepsilon = 100pJ/bit/m^2 \qquad (16)$$

The energy consumed by a sensor S_i in Active/Sleep modes is calculated following the model proposed by [33]:

$$E_{Radio}(S_i) = P_{Active} * T_{Active} + P_{Sleep} * T_{Sleep} \qquad (17)$$

As in [33], $P_{Active} = 1040mW$ and $P_{Sleep} = 200mW$.

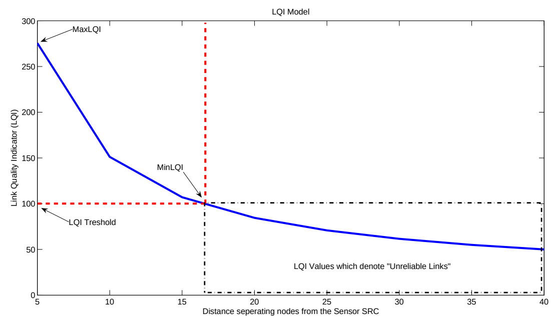

Fig. 6. LQI model: Links between S_{SRC} and its neighbors (Table IV)

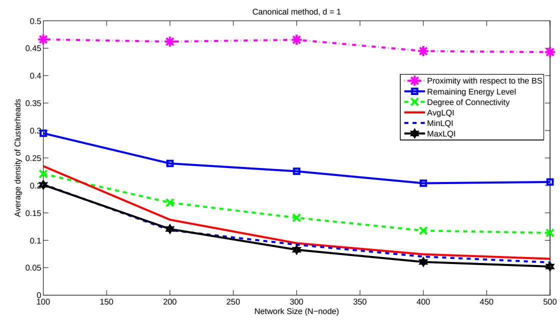

Fig. 7. Average density of clusterheads (canonical method)

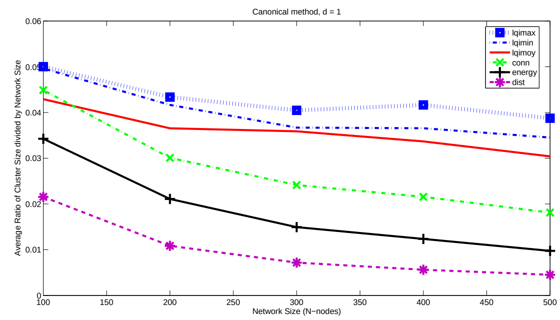

Fig. 8. Average cluster sizes divided by network size (canonical)

C. LQI model

After the WSN deployment in the warehouse, the BS initially broadcasts a message containing its location. This information is then retransmitted to all sensors in the network. In this phase, each node knows its degree of connectivity. Then, initial LQI values are calculated by using the $LQI(S_i, S_j)$ function defined below (similarly to the scale function Sc defined in the composite criterion in equation (11)):

$$LQI(S_i, S_j) = \beta + \frac{\Psi * \log(1 + (\gamma_j^i - \gamma_{min}^i))}{\log(1 + \gamma_{max}^i)} \quad (18)$$

$$\gamma_j^i = 1/d(i, j) \quad (19)$$

$$\gamma_{min}^i = Min_j\{\gamma_j^i\} \quad (20)$$

$$\gamma_{max}^i = Max_j\{\gamma_j^i\} \quad (21)$$

Where $\beta = 50$, $\psi = 255$ and $d(i, j)$ is the distance seperating S_j from S_i. The choice of this model is guided by experimental results shown in [34] and [35] which stated that the LQI decreases when the distance between nodes increases in Zigbee-based WSN. As we can see, $LQI(S_i, S_j) \neq LQI(S_j, S_i)$. Hence, the model allows to take into account asymetrical aspects of wireless links. In Fig. 7, we plot an example of LQI values produced by this model, when a node S_{SRC} has neighbors as shown in Table IV.

TABLE IV
EXAMPLE: DISTANCE SEPERATING NEIGHBORS S_j FROM S_{SRC}

Node ID (S_j)	1	2	3	4	5	6	7	8
D_j^{SRC}	5	10	15	20	25	30	35	40

VIII. Simulation Results

Simulations are run for a network size ranging from 100 to 500 nodes. The performance results presented here are obtained by averaging the results for 100 different simulations for each scenario, except for the scenario of the caryomme location (Fig. 19, 20 and 21) for which 80 different simulations were run. For each simulation, a new random node layout is used. In all simulation results presented below, $\alpha = 0.5$ for the composite criteria as defined in equations (9) and (10). Then, the MinHybrid criterion (resp. MaxHybrid) is composed of 50% of the "Remaining energy" criterion and of 50% of the "MinLQI" (resp. MaxLQI) criterion.

A. Single-Node Cluster Reduction (SNCR)

The Fig. 8 shows the average density of clusterheads produced by the MaxMin algorithm combined with the canonical method of cluster construction. The "Proximity with respect to the BS" criterion has an average density (around 45%) relatively high compared to other criteria. It is followed by the "Remaining Energy Level" criterion (between 25% and 30%), and then by the "Degree of Connectivity" criterion. The AvgLQI, MinLQI and MaxLQI criteria produce lower densities of clusterheads which decrease when the network density increases.

The Fig. 9 shows the ratio of the average cluster size divided by the number of sensors in the network. The clusters are formed by the canonical method. This result confirms the previous one, because the criteria that produced the most of caryommes (Fig. 8) are those which have the smallest average ratio of cluster size. So it is consistent to have in the decreasing order of the average ratio of cluster size: MaxLQI, MinLQI, AvgLQI, degree of connectivity, and "Remaining Energy Level", "Proximity with respect to the BS".

The Fig. 10 shows the average density of single-node clusters produced by MaxMin combined with the canonical method of cluster formings. The "Proximity with respect to the BS" (from 62% to 75%)and "Remaining Energy Level" (between 30% and 60%) criteria produces high densities of single-node clusters which increase when the network density

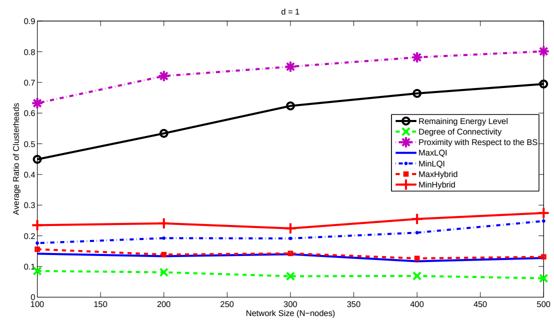

Fig. 9. Average density of single-node clusters (canonical method)

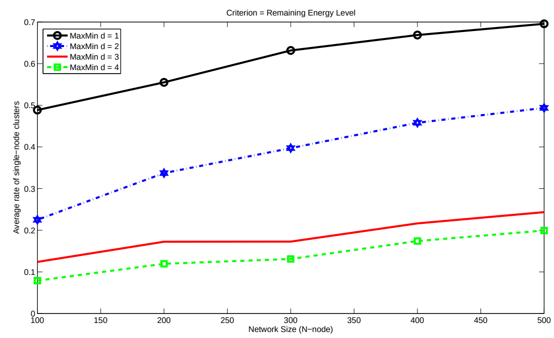

Fig. 10. Average density of single-node clusters (canonical method)

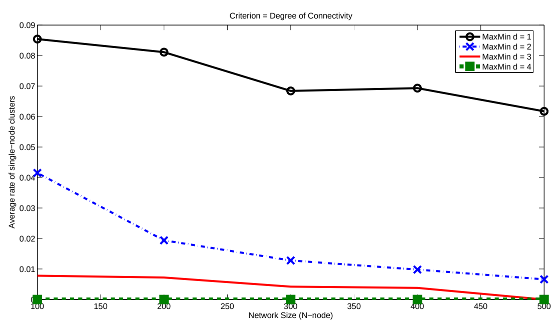

Fig. 11. Average density of single-node clusters (canonical method)

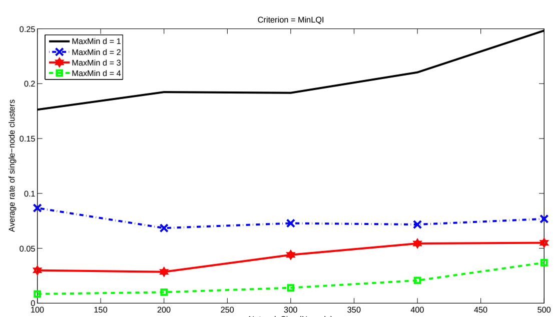

Fig. 12. Average density of single-node clusters (canonical method)

increases. This explains the previous result (Fig. 8) in which

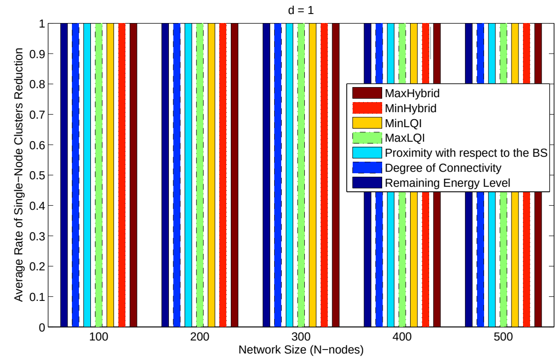

Fig. 13. Average density single-node cluster reduction (SNCR)

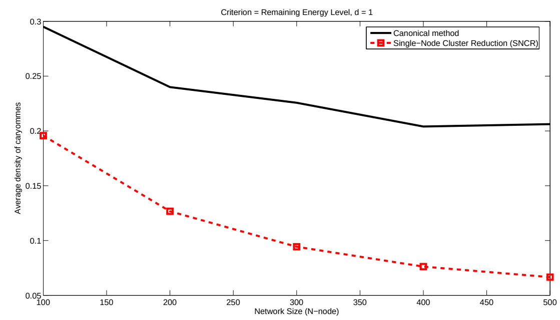

Fig. 14. Average ratio of clusterheads (Remaining Energy, d=1)

these two criteria produced more clusters than other criteria. The densities for the Hybrid and LQI criteria remain low at around 20%. The "degree of connectivity" criterion has the lowest density of single-node clusters (under 10%). The Fig. 11 for "Remaining Energy Level" (resp. Fig. 13 and Fig. 12 for MinLQI and the "degree of connectivity") shows that the average density of single-node clusters is decreasing when the MaxMin d parameter increases. For the "degree of connectivity" criterion, there is no single-node clusters as soon as $d = 4$ (Fig. 12).

The Fig. 14 shows the single-node cluster reduction density produced by the proposed single-node cluster reduction (SNCR) mechanism. For all studied criteria, the reduction percentage is 100%. This means that all the single-node clusters, produced by MaxMin run with the canonical method, have been eliminated by this reduction mechanism, whatever the criterion under consideration. Accordingly, in a wireless sensor network where there is zero non-connected node (i.e each sensor has at least one neighbor), this mechanism totally eliminates the single-node cluster phenomenon.

The figures (Fig. 15,16,17 and 18) shows the average ratio of clusterheads for each criterion by comparing the results obtained with the canonical method of cluster constructions and with the single-node cluster reduction mechanism (SNCR). SNCR produces less number of clusterheads than the canonical method. The gap, due to single-node clusters, is largest for the "Proximity with respect to the BS" criterion (20% to

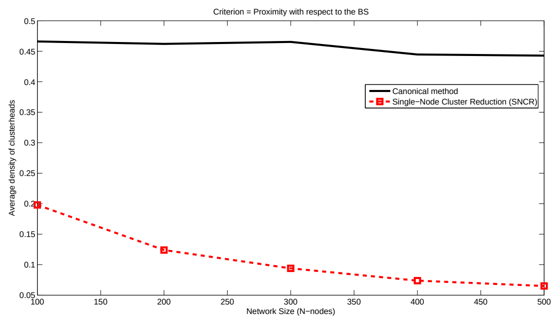

Fig. 15. Average ratio of clusterheads (Proximity BS, d=1)

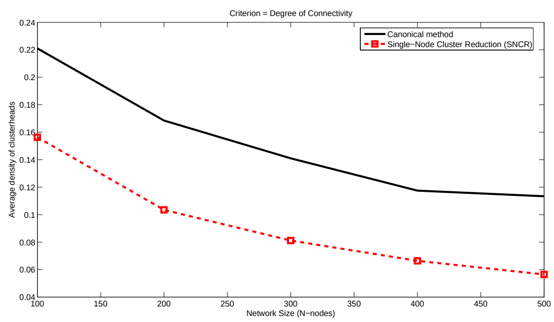

Fig. 16. Average ratio of clusterheads (Degree of connectivity, d=1)

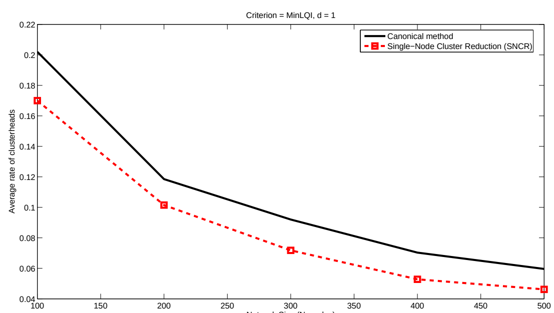

Fig. 17. The average ratio of clusterheads (MinLQI, d=1)

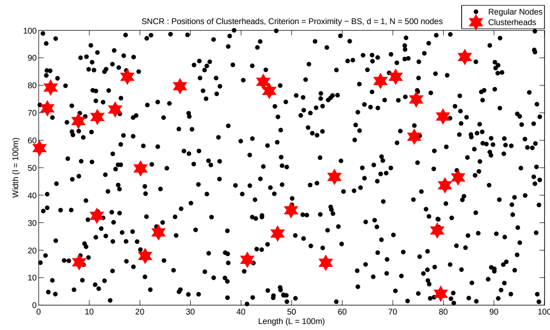

Fig. 18. Positions of caryommes (SNCR, Proximity-BS, d=1)

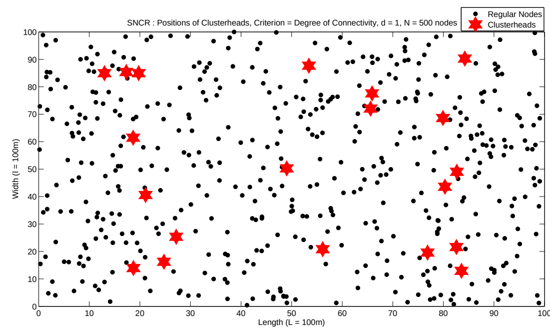

Fig. 19. Positions of caryommes (SNCR, Degree of Connectivity, d=1)

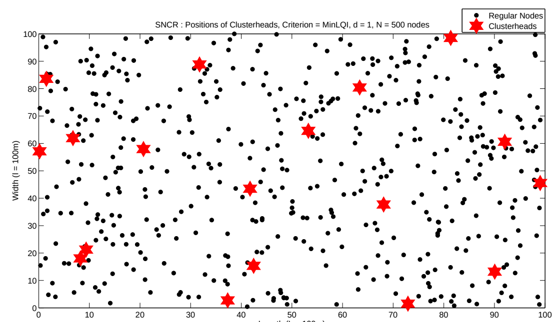

Fig. 20. Positions of caryommes (SNCR, MinLQI, d=1)

45%, Fig. 16). The difference is also relatively substantial for the "Remaining Energy Level" criterion (from 10 to 20%). The results for other criteria are: degree of connectivity (6%, constant gap), MinLQI (decreasing differences from 6 to 2% when the network size increases)

B. Positions of caryommes

The Figures 19, 20 and 21 display the clusterheads selected with MaxMin combined with the single-node cluster reduction (SNCR) mechanism. These results show that the locations of caryommes are not optimal when the "Proximity with respect to the BS" (Fig. 19) and the "degree of connectivity" (Fig. 20) are used as criterion. As for MinLQI (Fig. 21), clusterheads are sufficiently outspread which denotes better locations for clusterheads. The locations of caryommes generated by the degree of connectivity criterion are not optimal because if a node has a high degree of connectivity, then its closest neighbors also have a high degree of connectivity. So this criterion promotes the creation of neighboring nodes as clusterheads. Likely, for the "Proximity with respect to the BS", if a node is close to the BS, its nearest neighbors are also close to the BS. Conversely, choosing the MinLQI criterion promotes the election of sensors enough apart from each other. This leads to a better geographical distribution of caryommes. If caryommes are not sufficiently separated from each other, this affects the energy efficiency of the network. Indeed, when a "regular node" communicates with its own caryomme, the other neighboring caryommes hear the communication which is not intended to them. Therefore the energy consumption increases. Furthermore the risk of collision also increases because two neighboring nodes which have distinct clusterheads

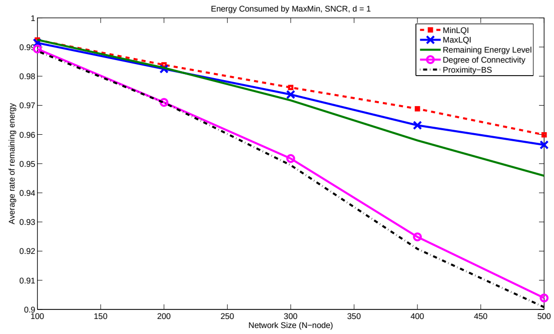

Fig. 21. Average rate of remaining energy (SNCR, d=1)

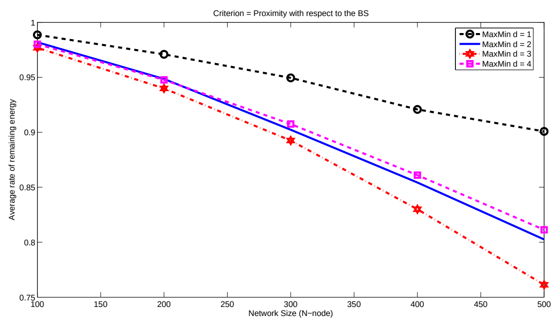

Fig. 22. Average rate of remaining energy (Proximity-BS)

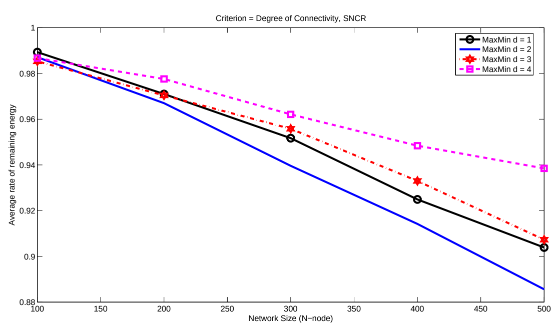

Fig. 23. Average rate of remaining energy (Degree of Connectivity)

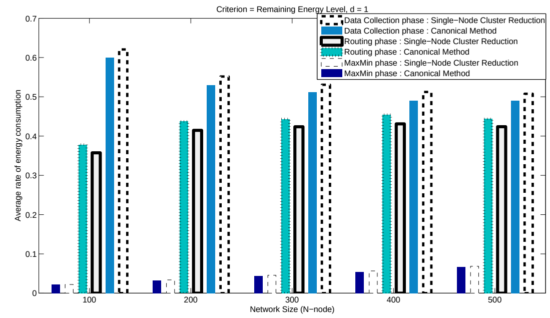

Fig. 24. Average rate of energy consumed by phase (Energy, d=1)

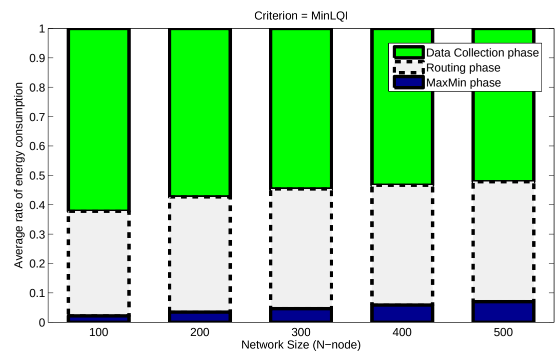

Fig. 25. Average rate of the consumed energy by phase (SNCR, d=1)

could try to communicate simultaneously with their respective clusterheads which are also in the same radio range.

C. Energy consumption

The application consists of three main phases: the MaxMin phase, the data collection phase and the routing phase (Fig. 6). The MaxMin phase is composed of clustering ones (initial, floodmax and floodmin phases), followed by the step of cluster formation (canonical vs. SNCR).

The Figures 22, 23 and 24 show the average rate of the remaining energy after running the MaxMin algorithm: floodmax, floodmin and cluster formation using SNCR. The MaxMin energy consumption depends on the criterion used to select caryommes. MaxMin consumes less energy with MinLQI. The "Proximity with respect to the BS" and the "degree of connectivity" criteria consume much more energy. The Fig. 23 denotes that increasing the value of d (d=1,2,3), has the effects to increase the MaxMin energy consumption for the "Proximity with respect to the BS" criterion for which the energy consumption begins to decrease since $d = 4$. Conversely, for other criteria (Fig. 24 shows the "degree of connectivity" criterion), increasing the value of d has the effects to decrease the energy consumption (since $d = 1$) of the MaxMin algorithm. Increasing the d parameter increases the number of rounds for floodmax and floodmin. Thus, the energy consumption also increases in these phases when the depletion of the caryomme density is not significant (Fig. 23). However, when the d parameter increases, the number of caryommes (sometimes considerably) decreases and therefore the number of "CH-INFORM-MSG" messages. The overall energy consumption decreases when the depletion of the number of caryommes is important (Fig. 24).

The Figures 25 and 26 show the average energy consumed in each phase, considering the "Remaining energy level" and MinLQI criteria. The Figures 27 and 28 are related to the energy consumption of the "Proximity with respect to the BS" criterion. We compare the case where clusters are formed by the canonical method with the single-node cluster reduction mechanism (SNCR). These results show that the MaxMin phase is the one that consumes less energy. Moreover, its rate of energy consumption is relatively low (from 2 to 7%) compared with energy expenditure in the data collection phase where each sensor sends alarms towards its caryomme (from

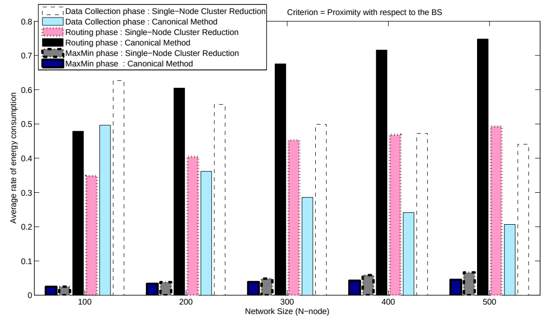

Fig. 26. Average rate of the consumed energy by phase (Proximity-BS, d=1)

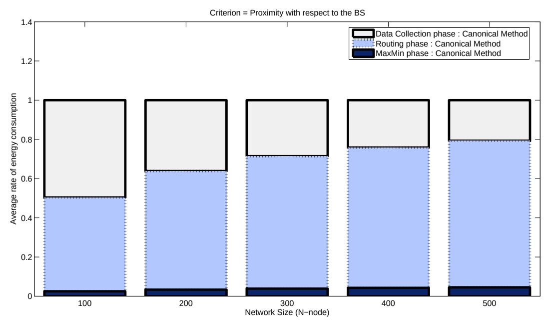

Fig. 27. Average rate of the consumed energy by phase (Proximity-BS, d=1)

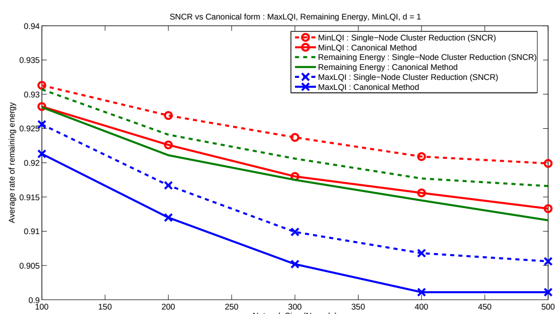

Fig. 28. Average rate of the remaining energy (d=1)

38 to 45%). The routing phase, in which alarms are aggregated and sent from each caryomme to the base station, consumes more energy (approximately from 60 to 48%).

The Fig. 29 shows the average rate of the remaining energy in the network after a complete cycle (node deployments, MaxMin clustering, data collection, routing and sensing phase Fig. 6). The MaxLQI, MinLQI and remaining energy criteria are compared in the two mechanisms of cluster formations (canonical and SNCR). The SNCR mechanism is more energy efficient than the canonical method. This is explained by previous results in which SNCR totally eliminates single-node clusters and then produces less number of clusterheads than canonical method. Due to the high density of single-node clusters, the large amount of energy spent in the routing phase causes a less energy efficient clustering scheme when

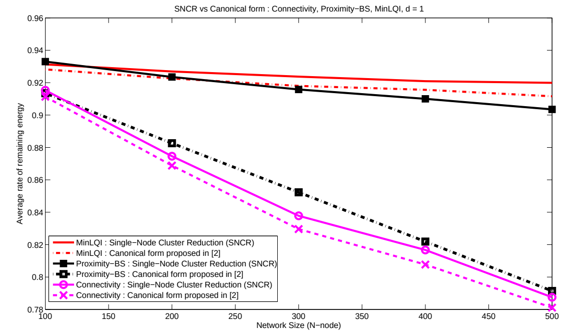

Fig. 29. Average rate of the remaining energy (d=1)

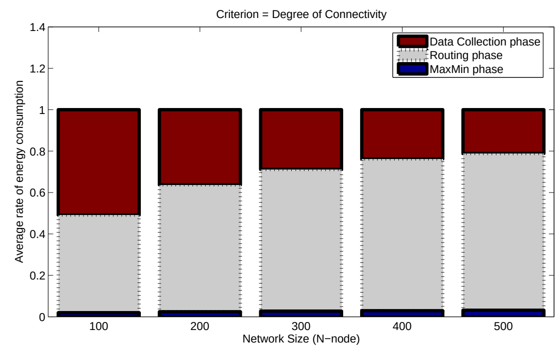

Fig. 30. Average rate of the consumed energy by phase (SNCR, d=1)

canonical method is used. This result denotes that a high density of single-node clusters has negative effects on energy consumption. Moreover, MinLQI is more energy-efficient than MaxLQI which matches the Zigbee standard definition.

The Fig. 30 displays the average rate of the remaining energy in the network after a complete cycle for the following criteria: "proximity with respect to the BS", "degree of connectivity" and "MinLQI". These criteria are compared in both canonical and SNCR mechanisms. MinLQI is more energy efficient than other criteria. The "degree of connectivity" criterion has the worse energy efficiency, its average rate of the remaining energy significantly decreases. The Fig. 30 shows that this criterion is also less energy efficient with clusters produced by the canonical method. The figures Fig. 31 and Fig. 32 (for SNCR) denote that the "degree of connectivity" criterion consumes much more energy in the routing phase compared with other critera (compare also Fig. 26 with Fig. 31). This explain why this criterion is less energy efficient than others: the relative locations of selected caryommes are not optimal. In such a situation this criterion leads to high energy consumption even if it produces low number of single-node clusters (Fig. 10).

IX. Conclusion

In this report, we studied the single-node cluster phenomenon pertaining to the MaxMin clustering algorithm. We compared several clusterhead selection criteria and then proposed a simple single-node cluster reduction (SNCR) mecha-

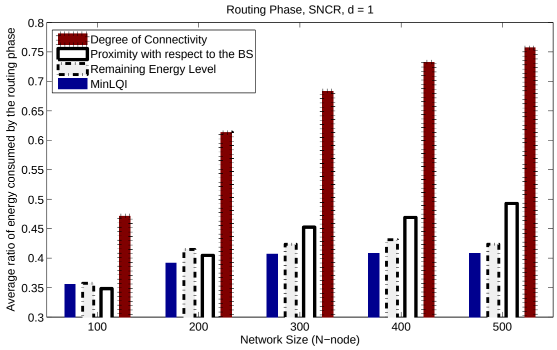

Fig. 31. Average ratio of the energy consumed by the Routing phase

nism. We also proposed a performant manner of using LQI-based criteria (MaxLQI, MinLQI). The density of single-node clusters is relatively high for the "Proximity with respect to the BS" criterion and for the "Remaining energy level". The "degree of connectivity" criterion has the lowest average density of single-node clusters. However, even if the phenomenon exists it is less important for LQI and Hybrid criteria. The "Proximity with respect to the BS" criterion is less performing than the "Remaining energy" which provides intermediate performance. The "degree of connectivity" is the worse energy-efficient criterion because the locations of selected caryommes are not optimal. By setting a given LQI threshold, i.e a value of acceptable LQI, and considering the lowest LQI value beyond this threshold, we obtain the optimal MinLQI criterion which highly enhances the energy-efficiency. Thus MinLQI is better than MaxLQI which matches the Zigbee standard definition. Single-node clusters have the drawback of increasing the energy consumption. The proposed single-node cluster reduction mechanism eliminates all connected single-node clusters in the WSN, and then improve the energy efficiency of the network. This work shows that, although it is important to be performant in selecting clusterheads, the step of forming clusters is also crucial. In this step reducing single-node clusters should be the primary objective in order to achieve energy efficiency.

REFERENCES

[1] A.D. Amis, R. Prakash, T.H.P Vuong, and D.T. Huynh. Max-min d-cluster formation in wireless ad hoc networks. In *9th Annual Joint Conference of the IEEE Computer and Communications Societies*, 2000.

[2] A. Delye. Etude théorique des clusters multi-sauts dans les réseaux de capteurs sans fils. In *Thèse de doctorat*, Université Paris VI, 2007.

[3] M. Marot. Conception et analyse de méthodes d'agrégation et multiniveaux dans les réseaux. In *Habilitation à Diriger des Recherches (HDR)*, Université Pierre et Marie Curie (Paris VI), 2009.

[4] IEEE Std 802.15.4-2006. Wireless medium access control (mac) and physical layer (phy) specifications for low-rate wireless personal area networks (wpans). In *IEEE Computer Society*, 2006.

[5] Zigbee Alliance. Zigbee specification v1. Jun. 2005.

[6] C. Diallo, M. Marot, and M. Becker. Link quality and local load balancing routing mechanisms in wireless sensor networks. In *Rapport de Recherche no 10002-RST*, Institut TELECOM; TELECOM SudParis, Jan. 2010.

[7] R. Nagpal and D. Coore. An algorithm for group formation in an amorphous computer. In *Proc. of the 10th International Conference on Parallel and Distributed Systems, PDCS'98*, 1998.

[8] K. Xu and M. Gerla. A heterogeneous routing protocol based on a new stable clustering scheme. In *Proc. MILCOM 2002, Vol. 2*, pages 838–843, Oct. 2002.

[9] M. Rajiullah and S. Shimamoto. An energy-aware periodical data gathering protocol using deterministic clustering in wireless sensor networks (wsn). In *IEEE Wireless Communications and Networking Conference, WCNC 2007*, pages 3014–3018, Mar. 2007.

[10] D.-H. Nam and H.-K. Min. An efficient ad-hoc routing using a hybrid clustering method in a wireless sensor network. In *3rd IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'07)*, 2007.

[11] C.-M. Liu, C.-H. Lee, and L.-C.Wang. Distributed clustering algorithms for data-gathering in wireless mobile sensor networks. *Journal of Parallel and Distributed Computing, ISSN 0743-7315, Vol. 67, Issue 11*, pages 1187–1200.

[12] P. Tillapart, S. Thammarojsakul, T. Thumthawatworn, and P. Santiprabhob. An approach to hybrid clustering and routing in wireless sensor networks. In *IEEE Aerospace Conference*, pages 1–8, Mar. 2005.

[13] R. Chang and C. Kuo. An energy efficient routing mechanism for wireless sensor networks. In *Proc. of the 20th international Conference on Advanced Information Networking and Applications, Vol. 2*, Apr. 2006.

[14] P. Ding, J. Holliday, and A. Celik. Distributed energy efficient hierarchical clustering for wireless sensor networks. In *Proc. of the IEEE International Conference on Distributed Computing in Sensor Systems, DCOSS'05*, Marina Del Rey, CA, Jun. 2005.

[15] O. Younis and S. Fahmy. Distributed clustering in ad-hoc sensor networks : a hybrid, energy-efficient approach. In *Proc. of IEEE Infocom 2004*, Hong Kong, Mar. 2004.

[16] M. Chatterjee, S. K. Das, and D. Turgut. Wca : A weighted clustering algorithm for mobile ad hoc networks. *Journal of Cluster Computing, Special Issue on Mobile Ad hoc Networks, Vol. 5*, page 193Ű204, 2001.

[17] W. B. Heinzelman, A. Chandrakasan, and H. Balakrishnan. An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications, Vol. 1(4)*, pages 660–670, Oct. 2002.

[18] E. Klaoudatou, E. Konstantinou, G. Kambourakis, and S. Gritzalis. Clustering oriented architectures in medical sensor environments. In *3rd International Conference on Availability, Reliability and Security, ARES 08*, pages 929–934, Mar. 2008.

[19] F. Dai and J. Wu. On constructing k-connected k-dominating set in wireless networks. In *Proc. of the 19th IEEE international Parallel and Distributed Processing Symposium, IPDPS'05, Vol. 01*, Washington, DC, USA, Apr. 2005.

[20] M.R. Brust, H. Frey, and S. Rothkugel. Dynamic multi-hop clustering for mobile hybrid wireless networks. In *ACM Proc. of the 2nd international Conference on Ubiquitous information Management and Communication, ICUIMC'08, Vol. 01*, Suwon, Korea, Jan. 2008.

[21] Y. Liu, J. Gao, L. Zhu, and Y. Zhang. A clustering algorithm based on communication facility in wsn. In *WRI International Conference on Communications and Mobile Computing, CMC '09, Vol. 2*, pages 76–80, Jan. 2009.

[22] F. G. Nocetti, J. S. Gonzalez, and I. Stojmenovic. Connectivity based k-hop clustering in wireless networks. *Telecommunication Systems, Vol. 22, Numbers 1-4*, pages 205–220, Jan. 2003.

[23] M. Qin and R. Zimmermann. An energy-efficient voting-based clustering algorithm for sensor networks. In *Proc. of the 6th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks (SNPD/SAWN)*, 2005.

[24] J.Y. Cheng, S.Ruan, R.Cheng, and T.Hsu. Padcp: Power-aware dynamic clustering protocol for wireless sensor network. In *IFIP International Conference on Wireless and Optical Communications Networks (WOCN)*, Apr. 2006.

[25] K. Sun, P. Peng, P. Ning, and C. Wang. Secure distributed cluster formation in wireless sensor networks. In *Proc. of the 22nd Annual Computer Security Applications Conference*, pages 131–140, Dec. 2006.

[26] K. Iwanicki and M. v. Steen. Multi-hop cluster hierarchy maintenance in wireless sensor networks: A case for gossip-based protocols. In *6th European Conference on Wireless Sensor Networks, EWSN 2009*, Cork, Ireland, 2009.

[27] Tmote Sky. http://www.moteiv.com/products/docs/tmotesky-datasheet.pdf.

[28] CC2420 Radio. http://www.chipcon.com.

[29] A. Gupta, C. Diallo, M. Marot, and M. Becker. Understanding topology challenges in the implementation of wireless sensor network for cold chain. In *Proc. IEEE Radio and Wireless Symposium, RWS 2010*, New Orleans, LA, USA, 2010.

[30] J. Polastre, J. Hui, J. Z. P. Levis, D. Culler, S. Shenker, and I. Stoica. A unifying link abstraction for wireless sensor networks. In *SenSys*, 2005.

[31] C. Diallo, A. Gupta, M. Becker, and M. Marot. Energy aware database updating protocols for autoconfigurable sensor networks. In *GlobeNet 2009, the 8th international conference on Networks, ICN'09*, Cancun, Mexico, Mar. 2009.

[32] C. Diallo, A. Gupta, M. Marot, and M. Becker. Virtual base station election for wireless sensor networks. In *ACM Notere 2008, the 8th international conference on New Technologies in Distributed Systems, Vol. 2*, Lyon, France, Jun. 2008.

[33] E. Shih, S.-H. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. Chandrakasan. Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks. In *ACM MobiCom*, Rome, Jul. 2001.

[34] J. Blumenthal, R. Grossmann, F. Golatowski, and D. Timmermann. Weighted centroid localization in zigbee-based sensor networks. In *IEEE International Symposium on Intelligent Signal Processing, WISP*, 2007.

[35] M. Becker, A.L Beylot, R. Dhaou, A. Gupta, R. Kacimi, and M. Marot. Experimental study: Link quality and deployment issues in wireless sensor networks. In *Proc. NETWORKING 2009, LNCS 5550*, pages 14–25, NETWORKING, Aachen, Germany, 2009.