EAI Endorsed Transactions
on Future Internet
Research Article ICST.ORG

Weighted Route Selection in Cluster-Based Protocol for Wireless Sensor Networks

Rivo S. A. Randriatsiferana[1,*], Herimpitia T. C. Antilahy[1], Frederic Alicalapa[1] and Richard Lorion[1]

[1]University of La Reunion, LE^{2}P Laboratory, Saint Denis, Reunion, France

Abstract

The efficiency of wireless sensor network is highly based on the algorithm complexity of the communication protocol which must minimize the energy dissipated by the sensor nodes. In order to pursue energy consumption minimization, this paper presents Weighted Route Selection (WeRoS) in Cluster-Based method, which uses multihop communication for data forwarding from cluster-head (CH) to the base station as a unique path. WeRoS takes also into account the remaining energy of a node and its coefficient of variation during CHs selection process. A novel-scheduling algorithm is also proposed, in which the CHs can self-organize themselves during the data transmission phase. This algorithm is executed in the cluster formation phase, in order to avoid another phase of route discovery and to conserve energy, while considering the node memory constraint. Complexity analysis and simulation results show that WeRoS can efficiently balance the energy consumption and extends the network lifetime.

Keywords: binary greedy forwarding, cluster-based, energy-balance, network lifetime, self-organization, wireless sensor

Received on 20 June, 2015, accepted on 21 September, 2015, published on 18 May, 2016

doi:10.4108/eai.18-5-2016.151249

1. Introduction

Through technological advances in recent decades, society rapidly comes closer to an ubiquitous information systems era, composed of many digital distributed systems in our physical environment. The main idea is based on embedded computing tools in objects of everyday life. In other words, this is an era of technological convergence in which a set of disparate equipments discreetly communicates through a network of heterogeneous networks.

Distributed measurement systems contribute to this evolution through the development of wireless sensor networks (WSN). The miniaturization of microelectronics, micro-mechanical and communication systems has enabled the realization of wireless communicating sensors, allowing rapid deployments. The operational nature and expected features essentially depend upon constraints: energy and its correlated management [4]. A WSN is a set of nodes communicating over wireless links. It allows observing a given phenomenon in a geographical area. Actually, WSN has important applications in various fields with latest technologies [3].

Controlling the energy consumption of WSN in order to maximize a network lifetime remains the most crucial challenge. This challenge must be considered in the current context: nodes are small embedded components with limited resources, mainly memory capacity, data processing, communication range and non-rechargeable energy [3, 16]. Additionally, the communication part is the major energy consumer in these nodes [14]. Most of these applications area, such as battlefields, volcanoes and forests monitoring, industrial and clinical fields, requires that the WSN has a certain degree of autonomy: it has to run correctly without assistance, when human interventions and supervisions are difficult or impossible [11]. Therefore, sensor networks operation in a self-configured ad-hoc network and standalone mode became indispensable. This requires a well-organized network architecture in which the "next-hop" data transmission range toward the Base Station (BS) should be minimal as possible, with nodes autonomous decisions.

Architecture hierarchy has already been demonstrated as one of the best method to organize the network into a connected hierarchy, load balancing, increasing the network lifetime. It considerably allows light memory constraint due to the reduction of the routing table of each node when network scalability increases. Moreover, it avoids frequent broadcasts that can overload nodes [10, 11]. In fact, this architecture

*Corresponding author. Email: rivo.randriatsiferana@univ-reunion.fr

consists of partitioning the network as virtual clusters. In order to effectively ensure the load balance between nodes, some of them are sometimes elected as leader, which are usually called cluster-heads (CHs). In this case, the CH ensures that the data collection and aggregation from member nodes is operational. Afterwards, the aggregated data will be sent directly to the BS or through other CHs. Therefore, they generally consume more energy than the member nodes in the fact they perform these additional tasks. The CH selection method becomes a main key in the clustering process [11]. It has to ensure load balancing and a good geographical CH distribution in the network. A common solution to balance this energy consumption between all nodes of the network is to periodically elect new CHs. Thus the role as CH should be distributed over time (rounds) to all nodes in the network.

Basically, the process of clustering is divided into two main phases: the cluster construction phase and the data communication phase. The cluster construction phase consists of CH node selection method, such as probabilistic or non-probabilistic, and members node assignment to a cluster. While the data communication phase focuses not only on reliable data transmission to the BS, but also on finding optimal routes to enhance energy efficiency. Two possibilities can be used for the data communication process: one-hop and multi-hop communications [2, 9, 11].

One-hop communication assumes that all nodes can communicate directly with the BS. Although, this assumption could be appropriate for a small network but it penalizes distant nodes. The furthest CHs (from the BS) will die more quickly according to the longest transmission distance [8]. This can result in areas that will not be covered during network operation [11]. Also, it is important to adopt a multi-hop communication between distant CHs and the BS, as much as possible, to improve the network coverage. However, multi-hop communications generally route data using fixed paths and overuse the nodes closest to the BS, making them also die quickly and resulting in the existence of energy holes [17, 18]. Therefore, energy hole avoidance mechanism that ensures the data transmission, and adequate CH nodes selection that ensures energy balanced across the network, should be established.

In this paper, we propose a Weighted Route Selection in Cluster-Based Protocol for WSNs, denoted WeRoS, to mitigate the mentioned problems. WeRoS is an energy-efficient protocol that combines the CH selection rotation and a multi-hop data communication in order to extend the network lifetime. In fact, during the clustering process implementation, WeRoS adopts the probabilistic, adaptive and distributed approach. The CH selection depends on the remaining energy of the nodes and its coefficient of variation. These parameters are introduced to elect CHs by maximizing the remaining energy as well as minimizing its variance. This is useful to predict node state and to ensure balanced energy consumption over all nodes in the network. However, given the probabilistic and distributed feature, the main challenge is to ensure the existence of an optimal number of CHs at each round. WeRoS will focus on that point. In the data transmission phase, a self-scheduling algorithm named "Binary Greedy Forwarding" is proposed. It consists of designing an adjacency table between CHs in order to establish communication between CHs and BS. So that CHs can be self-organized during the data transmission phase. The data routing is based on the construction of "elementary path" which passes only one time through a CH to address the energy hole problem. To implement these features, we introduce a weight at each node. It depends on the remaining energy and the distance between the elected CH and the BS.

The paper is organized as follows: in section 2 related works are presented. Section 3 describes the proposed WeRoS protocol and its algorithm. In Section 4, simulation context and results are given and discussed. And finally in Section 5, conclusions are exposed.

2. Related works

Low Energy Adaptive Clustering Hierarchy (LEACH) [7] protocol is an elegant distributed algorithm and is the most popular cluster-based routing protocol that was proposed for reducing energy consumption. The clustering algorithm is a time process, for which time is expressed in rounds. Each round is divided into two main phases, the set-up and the steady-state phases (Fig.1). In fact, set-up phase is dedicated to perform clusters formation in which CHs selection takes place, while the steady-state phase is reserved to perform data transmission. Basically, the role of CH is rotated among the all the nodes, based on a probability of becoming a CH per round. A non-CH node joins its cluster by choosing the CH that can be reached with the least energy consumption using one-hop communication. At the data transmission step, one-hop communication is also used by each CH to send aggregated data to the BS. Therefore, the distant CH nodes from the BS die quickly due to the long transmission range. Moreover, remaining energy at each node is not taking into account in the CH selection process. It means that all nodes have the same probability to become CH, regardless of their remaining energy, resulting in premature death of nodes with lower energy. Authors propose a centralized version of LEACH, called LEACH-C [7], that takes into account the remaining energy, and the total energy in the network, which can be implemented by using energy information at each node. This requires additional communications between the nodes and the BS. This is a problem.

Several distributed protocols have also been published such as Advanced-LEACH [2] and Deterministic CH Selection (DCHS) [6]. They also address the energy problem by taking into account the remaining energy as the main metric in the CHs selection operation, but in a distributed process. Recently, e-LEACH [15] was proposed to improve the CHs selection by introducing an energy variation coefficient. The reduction of energy consumption phenomena has evolved, and the CHs layout is well distributed by setting as a main goal the following criteria: all nodes in the network should consume, approximately, the same amount of average energy at each round. However, the problem of unbalanced energy consumption still exists due to a direct communication between CHs and the BS. Therefore, a cluster-based routing protocol with equalized energy consumption has to be found.

A well-known evolution of LEACH is Hybrid Energy-Efficient Distributed (HEED) [19]. The HEED clustering protocol uses a hybrid criterion for CH selection. It considers the residual energy of each node to elect tentative CHs, and the final CHs are elected according to the intra-cluster communication cost. Its main targets are to achieve "even-distributed" CHs over networks and to reduce long distance transmission. But it cannot guarantee an optimal elected set of CHs. And the weakness of using "Tentative Status" algorithm is the communication strategy of isolated nodes. Indeed, they are forced to be CHs and to directly transmit their data to the BS. The iterative nature of the CHs election is one more weakness due to many communications exchanges.

Other approaches were proposed to improve the performance of clustering. EEUC [5], addresses the hotspot problem: the CHs closest to the BS are burdened with heavy relay traffics and tend to die early. EEUC partitions the nodes into clusters of unequal size, and clusters closer to the BS have smaller size compared to those further from the BS. EECS [5], another extended version of LEACH, realizes a localized selection of CHs and a near uniform distribution of them. In cluster formation phase, a non-CH node chooses its corresponding CH by considering not only saving its own energy, but also balancing the load of CHs. EECS requires the global distance between all CHs and the BS and all nodes have to contribute to become CH. Those requirements substantially increase the overhead complexity of control messages.

3. Weighted Route Selection in Cluster-Based Protocol

In this section, we present WeRoS protocol, which is proposed not only to enhance the network lifetime but also to avoid energy hole problem by balancing the nodes energy consumption. The communication mechanism of WeRoS protocol includes cluster formation phase and the data transmission phase. The organization of cluster construction phase is based on the clustering method of LEACH protocol [7]. Nevertheless, WeRoS provides an optimized threshold function to elect the CH nodes by ensuring their existence at each round, in which, the data communication phase is made up of the "intra-cluster" and "inter-cluster" communication (Fig.1).

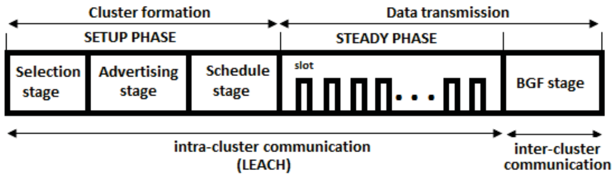

Figure 1. Communication mechanism of WeRoS protocol.

In the main, WeRoS protocol is based on the following mechanisms:

1. Intra-cluster communication consists of:

   - The formation of clusters is periodic and under control of CHs.

   - Data transmission between a CH and its node members is done in one-hop.

2. Inter-cluster communication consists of:

   - Create an adjacency table between CHs by defining a scheduling function weight.

   - Communication between CHs and the BS is done in multi-hop path if the base station cannot be reached directly.

Additionally, the following hypotheses are introduced:

1. all the nodes embed a voltage battery monitor electronic circuit. And the voltage value can be used for local calculus.

2. All the nodes also embed RSSI (receiver signal strength indication) circuit, to estimate the distance to any compatible emitter. Our approach considers static nodes.

3. Energy initial state is the same for all the nodes: for the simulation step, using TelosB nodes, two 3.3 mAh, 1.5 V batteries will be considered to provide the supply voltage for each node.

3.1. Energy model

The radio hardware model considered is to be the same as shown in [7]. In which, the energy expended for transmission is based on radio-electronics and power amplifier characteristics. Moreover, the power amplifier model is based on the distance, by assuming d^2 for the free-space (ϵ_{fs}) power loss and d^4 for multipath (ϵ_{mp}) fading channel. The energy consumed for transmission (E^{TX}), of an L-bit message over distance d is:

$$E^{TX}(L,d) = \begin{cases} E_{tx} \cdot L + \epsilon_{fs} \cdot L \cdot d^2, & if \;\; d \leq d_o \\ E_{tx} \cdot L + \epsilon_{mp} \cdot L \cdot d^4, & if \;\; d > d_o \end{cases} \quad (1)$$

With: $d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$. E_{tx} is the transmit energy consumed per bit. Whereas, the energy consumed to receive the L-bit message, can be found by:

$$E^{RX}(L) = E_{Rx-elec}(L) = E_{rx} \cdot L \quad (2)$$

where E_{rx} is the receiving energy consumed by the radio transceiver. The energy consumed in data aggregation at the CH from itself and m neighbor nodes is represented as:

$$E^{DA} = (m+1) \cdot L \cdot E_{da} \quad (3)$$

where E_{da} is the energy required to aggregate one-bit data. Through this model, the energy dissipated by the CH in a round is given by:

$$E_{CH} = (\frac{n}{k} - 1)L \cdot E_{rx} + \frac{n}{k} L \cdot E_{da} + L \cdot E_{tx} + \epsilon_{mp} \cdot L \cdot d^4_{toBS} \quad (4)$$

where d_{toBS} is the distance between CH and the BS. The energy dissipated in each non-CH node is expressed by:

$$E_{nonCH} = E_{tx} \cdot L + \epsilon_{fs} \cdot L \cdot d^2_{toCH} \quad (5)$$

3.2. Intra-cluster communication

The organization of the intra-cluster communication can be organized as follows: when the CH was selected, the CSMA protocol is used to broadcast advertisement messages "AdvMsg" to nodes in the network (Fig.1 "Advertising Stage"). All nodes estimate their distances to CHs based on RSSI. Then, the nodes send a "Join-Request" message to the nearest CH again using the CSMA protocol. Once the CH has received all the information messages, it uses TDMA to schedule the data reception that is based on the number of member nodes in the cluster. Afterwards, it sends notification to its member nodes (Fig.1 "scheduling stage").

Cluster-head selection. Given the role supported by the CHs, it obviously consumes more energy than the member nodes. It makes sense that selecting as CH the node having the most reserve energy and rotate the role of CH periodically between nodes significantly extend the network lifetime [6]. However, the problem of unbalanced energy consumption persists because of the random deployment of the nodes and non-uniform node distribution. Therefore, WeRoS incorporates the residual energy value and its variation coefficient as main criterion to select CH. The coefficient of variation represents the ratio of the standard deviation to the arithmetic mean. This is a useful for statistical measurement of the dispersion, and also for comparing the degree of variation of the remaining energy of all nodes, even if the means are radically different from each other [12, 13]. And then, the efficiency of this dispersion measurement, $E^i_{FF}(r)$, can be defined as the square of the variation coefficient, i.e:

$$E^i_{FF}(r) = (\frac{\sigma^i(r)}{\mu^i(r)})^2 = \frac{\mathcal{V}^i(r)}{\mu^{i2}(r)} \quad (6)$$

$\sigma^i(r)$, $\mu^i(r)$ and $\mathcal{V}^i(r)$ denote the standard deviation, the arithmetic mean and the variance of remaining energy of the node i at round r, respectively.

At first round, the variance of remaining energy $\nu^i(r)$ value equal to zero, then the efficiency $E^i_{FF}(r)$ goes to 0. However, so that we can see the impact of the variance and efficiency, we intend to be biased the remaining energy with it. Therefore, this efficiency should be initialized to 1 to ensure the convergence of our algorithm. Thus, a convex function $\gamma^i(r)$ is proposed, by considering the target, which tries to maximize the average of remaining energy while minimize the variance of remaining energy of node. Let's define $\gamma^i(r)$ as follow:

$$\gamma^i(r) = \frac{\mu^{i2}(r)}{\mu^{i2}(r) + \zeta \mathcal{V}^i(r)} \quad (7)$$

Where ζ is a tunable constant which is introduced to guarantee the same order of magnitude for the mean and the variance. Its value depends on the simulation parameters.

For choosing a node with highest value of remaining energy and lowest variance of energy consumption to become CHs, the remaining energy can be weighted by the $\gamma^i(r)$ function. Thus, the threshold expression $p^i(r)$ for CH selection proposed in [7] is modified as follow:

$$p^i(r) = \begin{cases} \frac{k}{N-k*(r \bmod \frac{N}{k})} \frac{\gamma^i(r)*E^i(r)}{E_o}, & C_i(r) = 1 \\ 0, & C_i(r) = 0 \end{cases} \quad (8)$$

$E^i(r)$ denotes the remaining energy of node i, $C_i(r)$ is an indicator function, with $C_i(r) = 1$ if node has not been a CH, and $C_i(r) = 0$ otherwise.
Via this criteria selection, the node will be selected as CH in the current round, which ensures that the data should be collected to a node having a higher energy reserve value and lower variance.

By considering that all nodes have almost the same amount of initial energy. Hence, at beginning of the network operation (i.e, $r = 1$), $\mathcal{V}^i(r) = 0$, so, $\gamma^i(r) \to 1$. Therefore, the threshold expression is the same as mentioned in [6]. As the network runs continuously, the energy reserve of nodes progressively deviates from equality. Thus, the $\gamma^i(r)$ expression considerably plays an important role.

To determine the mean and the variance values of remaining energy, all nodes must keep their previous remaining energy from first round to the current round (e.g.: $R = 1$ to $R = 450$ rounds). That puts a high burden on nodes to their limited processing and memory capacities and this iteration may create latency. Thus, it leads to a node's dysfunction in this case, and it will also disturb the whole network operation. Therefore, it is preferable to use another method to calculate $\mu^i(r)$ and $\mathcal{V}^i(r)$ in order to reduce the cost of the computation load. By using a moving average method deduction, we can recursively define $\mu^i(r)$ as follows:

$$\mu^i(r) = \frac{(r-1)\cdot\mu^i(r-1) + E^i(r)}{r} \tag{9}$$

and $\mathcal{V}^i(r)$ it can be written:

$$\mathcal{V}^i(r) = \frac{(r-1)\cdot\mathcal{V}^i(r-1) + E^{i2}(r)}{r} - \mu^{i2}(r) \tag{10}$$

In this recursive average approach, the result obtained is exactly as an arithmetic mean that is a benefit. In addition, in a practical manner, a node just memorizes its previous average remaining energy.

Optimal number of Cluster-Head. After a long running time, nodes in the network have little energy left. Then, the expression factor $\frac{\gamma^i(r)*E_i(r)}{E_o}$, in equation (8), becomes smaller. This phenomenon will lead to the issue that the CHs number is too small in $(r \mod (N/k))$ rounds, or, there is no CH in a round. The factor $\frac{\gamma^i(r)*E^i(r)}{E_o}$ should be controlled by applying the threshold $p^i(r)$ in order to guarantee the existence of k clusters number in each round.

Following the analysis in [7], author has proved that the probability function generated by $P^i(r) = \frac{k}{N-k*(r \,mod\, \frac{N}{k})}$ can ensure that there is k optimal number of CHs in each round, if this probability threshold is less than a number choose randomly in the range [0 1]. But in our approach, equation (8) will break this balance. However, the existence of k cluster should be ensured. To do that, WeRoS proposes the lemma 3.1:

Lemma 3.1. The expected optimal number of CHs k is guarantee in the I interval $[0 \;\; \frac{\gamma_i*\overline{E}_{NCH}(r)}{E_o}]$.

Where $\overline{E}_{NCH}(r)$: Average of remaining energy of nodes that have not been CH in the previous round.

Proof. The expected number optimal k of CH can be defined as [7]:

$$E[\#CH] = \sum_{i=1}^{N} P^i(r) * C_i(r) = k$$

Where $P^i(r)$ is the node i probability computing by itself in order to react as CH or not, in the current round [7]. In our case, $P^i(r)$ can be obtained by normalizing $p^i(r)$ in the I interval:

$$\begin{aligned} P^i(r) &= \frac{p^i(r) - 0}{\frac{\gamma^i*\overline{E}_{NCH}(r)}{E_o} - 0} \\ &= \frac{k}{N-k*(r \,mod\, \frac{N}{k})} \frac{\gamma^i(r)*E^i(r)}{E_o} \frac{E_o}{\gamma^i * \overline{E}_{NCH}(r)} \\ &= \frac{k}{N-k*(r \,mod\, \frac{N}{k})} \frac{E^i(r)}{\overline{E}_{NCH}(r)} \end{aligned}$$

Thus,

$$E[\#CH] = \sum_{i=1}^{N} \frac{k}{N-k*(r \,mod\, \frac{N}{k})} \frac{E^i(r)}{\overline{E}_{NCH}(r)} * C_i(r) \tag{11}$$

- Considering $\sum_{i=1}^{N} \frac{E^i(r)}{\overline{E}_{NCH}(r)} * C_i(r)$, $C_i(r) = 1$, if the node is eligible to become CH and $C_i(r) = 0$ otherwise. Hence, this term is the total energy remaining of the nodes has not been already elected as CHs divide by its average. Then, it represents the total number of nodes that have not been selected, which are eligible to be a CH in the current round. Thus,

$$\sum_{i=1}^{N} \frac{E^i(r)}{\overline{E}_{NCH}(r)} * C_i(r) = N - k*(r \,mod\, \frac{N}{k}) \tag{12}$$

The equation (11) can be rewritten as:

$$E[\#CH] = \frac{k}{N-k*(r \,mod\, \frac{N}{k})} \sum_{i=1}^{N} \frac{E^i(r)}{\overline{E}_{NCH}(r)} * C_i(r)$$

By considering (12), thus

$$\begin{aligned} &= \frac{k}{N-k*(r \,mod\, \frac{N}{k})} * (N - k*(r \,mod\, \frac{N}{k})) \\ &= k \end{aligned}$$

As a result, the average number of CHs k per round is guaranteed. □

In the steady phase (Fig.1), the nodes start to transmit the data payload which is composed of the remaining energy information and the sensed data to their CH. And then the CH sends the aggregated data and the remaining energy to the BS, during the inter-clustering communication phase (Fig.1). After that, the BS computes the average energy $\overline{E}_{NCH}(r)$, and then broadcasts this information to the entire network. Furthermore, this broadcast event is the beginning of rounds synchronization.

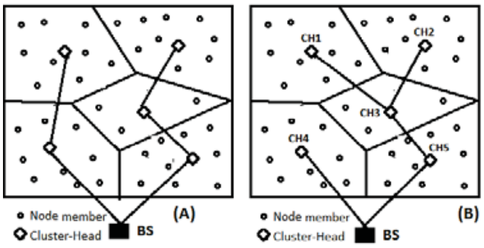

Figure 2. WeRoS Network Topology

	CH1	CH2	CH3	CH4	CH5	BS
CH1	0	0	1	0	0	0
CH2	0	0	1	0	0	0
CH3	0	0	0	0/1	1	0
CH4	0	0	0	0	0	1
CH5	0	0	0	0	0	1
BS	0	0	0	0	0	0

Figure 3. Boolean matrix of link existence.

3.3. Inter-cluster communication

The inter-cluster communication step consists of the data delivery between the furthest CHs to the BS. In which, the main target is to find an optimal path to balance the energy consumption. In fact, to solve the multi-hop communication paradigm as overloading the CHs closest to the BS, WeRoS proposes an novel algorithm which construct a single path from a CH toward the BS (Fig.2a). Furthermore, WeRoS also proposed that the CHs, which have already participated in the data transmission, should go to sleep mode, to conserve energy. Accordingly, this can generate isolated CH. Therefore, if the BS is not reached directly, the data will be lost. Thereby WeRoS presents the Binary Greedy Forwarding algorithm (BGF), which tends to overcome this problem (Fig.2b). The BGF algorithm consists in the design of a scheduling algorithm between the CHs and the BS so that the CHs self-organize themselves during the data transmission phase, in the inter-cluster communication step. Indeed, the CHs just keep in their routing table the successor ID (SID) and the number of their predecessor (NP). The NP impacts the fact the CH should participate or not in the data transmission. Afterwards, if the CH has no predecessor, it may decide autonomously to communicate to his successor.

Overview Binary Greedy Forwarding algorithm. BGF algorithm runs in the data transmission phase (Fig.1:Inter-cluster communication). Before developing this algorithm, let's firstly consider the graph (Fig.2b), by referring to its adjacency matrix (Table.3). It is called Boolean Matrix, which indicates the existence of a link between two CHs. BGF main goal is not only to build an elementary path for each CH, but also to order the CHs actions from furthest CH to BS (Fig.2a). However, there is a critical case, for example, CH_4 may be not accessible by CH_1. In fact, CH_1 must go through CH_3 (Fig.2b). Furthermore, the node CH_3 must choose a path between of CH_4 or CH_5. For easier analysis, the binary table is adopted which can be defined as follows: we assign the value 1 if the link (CH_i, CH_j) exists, and 0 otherwise.

In light of this table, the columns represent predecessor table of CHs, while the lines represent its table successor. First, reasoning on the predecessor table: columns only containing 0 (no link), it means that the node has no predecessor (NP=0). Therefore, nodes can react as a source without waiting for any predecessor (CH_1 and $CH2$). Let suppose that CH_5 is considered as CH_3 successor, then CH_4 can also react as a source and obviously, the BS is its successor. On the other side, for CH_3 and CH_5 their predecessor table is marked by 1, so they have to wait for their corresponding predecessors. Then, if we now examine the successor table (lines): rows only containing 0 (no link, BS excluded), it means that the node has no successor. Therefore, the nodes can react as sink (CH_4 and CH_5). They have to wait for their corresponding predecessors before sending their data to the BS. The following algorithm provides an overview of establishing a link between CHs.

Algorithm 1: BGF algorithm

```
Input: NP[i]: NP of CH_i et Ω: set of CHs
Result: Ordered construction of the link of CHs.
for i ← 1 to Card(Ω) do
    if NP[i] = 0 then
        for j ← i + 1 to Card(Ω) do
            if m_ij = 1 then
                Establish the link between(CH_i,CH_j);
                NP[j] = NP[j] − 1;
            end
        end
    end
end
```

In practical, the nodes only need a table containing its successor ID (SID) and the number of its predecessors (NP). This method carries on benefits on memory constraint and the iterations number of computation. For example, CH_3 only keep its SID and its NP, respectively $NP = 2$ and $SID = 5$. This node must wait for data from CH_1 and CH_2.

Table construction of "Binary Greedy Forwarding". The main goal is to form a binary matrix that not only would order the CHs but also help to easily identify their direct neighbor. A weight is introduced in order to arrange the CH nodes so that the distant CHs to the BS are able to react as a primary source without its intervention. For that purpose, all CH nodes evaluate the following weight value ψ_i:

$$\psi^i(r) = (1 - \frac{R_{Th}}{D^i(rssi)})\frac{E^i(r)}{E_o} \tag{13}$$

Where $E^i(r)$, E_0: remaining energy and initial energy value of the CH i, respectively. $D^i(rssi)$: distance to the BS estimation according to RSSI value, and R_{Th} is its communication range threshold. In relation (13), the remaining energy of the node is biased by their distance from the base station. Therefore, the weight value depends strongly on the distance. The remaining energy is required when the node should select a route among its neighbors. Based on this relationship, two cases could be envisaged:

Case 1. If $D_i(rssi) \leq R_{Th}$, then the value of $\psi_i(r)$ becomes negative or zero. Thus, the BS is in the coverage radius of the cluster. In fact, the CH can communicate directly with the BS unless there is no predecessor.

Case 2. If $D_i(rssi) > R_{Th}$, then the value of $\psi_i(r)$ is certainly positive. Furthermore, if the node is further away from the base station, the weight value increases. Thus, the CH shall communicate to multi-hop.

Indeed, this expression allows the CH to pick one of his neighbors who has a high weight. This choice is to build an elementary path and also to avoid overloading the CH closer to the BS. Let assume, Ω is the set of CH neighbors of the cluster CH_i and Ψ as the set of these weights. The binary matrix m_{ij} can be defined as follows:

$$m_{ij} = \begin{cases} 1, & \text{if } \psi_j = \max\limits_{k \in \Psi_i}(\Psi_k) \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

In other words,

$$route(CH_i, CH_j) = (CH_j, \psi_j) \in \Omega_i \times \Psi_i, i \neq j \mid \psi_j = max(\Psi) \tag{15}$$

If a CH receives data from another CH, it will forward the data directly to its next hop without aggregation, and go to sleep mode to save energy.

Principle of the protocol based on the algorithm BGF. Given that, once the CH selection phase is completed, the CH broadcast an announcement message "AdvMessage" to all nodes in the network whereby those nodes can join the closest CH to them. The CH is often in idle listening in order to access transmission medium (CSMA), which generates significant energy consumption. In fact, the advertisement message "AdvMsg" should be contained the ID and the weight ψ_i of the node. The goal is not only in order to avoid wasting this energy but also to prevent another route discovery after cluster formation, like we seen in [5, 10, 17]. Moreover, the others CHs can determine the number of its predecessor (NP) and the ID of its immediate successor (SID). By considering the memory capacity and computation constraint, the CHs just only memorize its SID and its NP. Then, a predecessor counting principle is done locally at each time the node hears the "AdvMsg" message. The following algorithm carries a comprehensive view of determination of the SID and NP of each CH.

Algorithm 2: Determination of the SID and NP

```
Input: ψi(r) of CHi
Result: SID[i], NP[i] of CHi
max ← 0
NP[i] ← 0
Broadcast AdvMSG(ID, ψi)
if ψi ≤ 0 then
    SID[i] ← BS.Id
else
    while ("Advertising stage" has not finished) do
        Receive AdvMSG(ID, ψi) message
        if ψi > ψj then
            if max ≤ ψj then
                max ← ψj
                SID[i] ← CHj.Id
            end
        else
            NP[i] = NP[i] + 1
        end
    end
end
```

This algorithm runs during the "Advertising Stage" of the cluster formation in the intra-cluster communication to reduce the overhead complexity of control message.

4. Analysis and Simulation

The performance evaluation of our algorithm has been simulated in Matlab, a series of simulation experiments were conducted. In order to investigate the feasibility and effectiveness of WeRoS, we compared it with LEACH, LEACH-C, e-LEACH and HEED. e-LEACH is one-hop version that considers only the optimized threshold (equation.8).

4.1. Complexity Analysis

Lemma 4.1. During clustering process, the control messages complexity in WeRoS protocol across the network is $O(N)$.

Proof. The cluster formation process of WeRoS is the same of LEACH protocol [7]. Let assume, there are N nodes in the network and suppose k CHs are generated. In each round, each CH broadcast the advertisement message "AdvMsg" and schedule message, while each cluster member node sends "Join Request" to the respective CH, where the total number is $N - k$. Thus, the total overhead control message in the entire network is: $2k + N - k = k + N$ Therefore, the control messages complexity is $O(N)$ □

HEED is a distributed protocol [19], the cluster construction phase is therefore localized. In addition, the CHs are selected with iterations. Thus, it provides an upper-bound control messages complexity of $N_{iter} \times N$. Where N_{iter} is the iterations numbers. Therefore, WeRoS algorithm is much better than HEED in term of control messages.

Lemma 4.2. Node information stored complexity in each node during data transmission is $O(1)$.

Proof. This analysis is proved according to the node's role respectively.
Case 1, For the CH node: After data collection from member nodes, each furthest CH to BS start to transmit its data. In this case, it should only consider successor ID information. Thus, the complexity of node information stored in a CH is $O(1)$.
Case 2, For member node: After receiving scheduling message, each node only keeps about its CH's information. Thus, the complexity of node information stored in each member node is $O(1)$.
Therefore, the node information complexity stored during data transmission phase in each node is $O(1)$. □

By considering lemma(4.1) and lemma(4.2), WeRoS proved that it presents a lower complexity in the different phases of its algorithm. Therefore, WeRoS is an energy-efficiency protocol.

4.2. Simulation parameters

Hundred sensor nodes have been deployed randomly in a $100x100$ square meter area. For our simulation, we have taken standard values from the Chipcon RFIC datasheet [1]. Transmit energy consumed for a one packet transmission, E_{tx}, was $0.121mJ$ and receiving energy consumed, E_{rx} was $0.208mJ$. The initial energy of each node was set to $100mJ$. The data payload have the same length $L = 320$ bits. The other radio parameters are the same as in [7]: the consumed energy for data aggregation $E_{da} = 5nJ$, the transmit amplifier free-space parameter $\epsilon_{fs} = 10pJ/bit/m^2$ and the transmit amplifier for multi-path fading parameter $\epsilon_{mp} = 0.0013pJ/bit/m^4$. First, we provide a study of the energy efficiency of the five algorithms by examining the network lifetime.

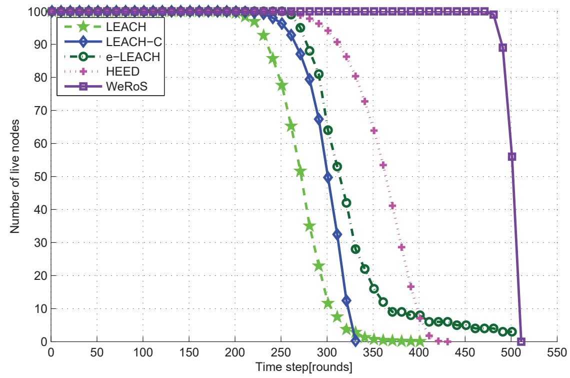

Figure 4. Comparison of network lifetime.

4.3. Network lifetime

Fig.4 illustrates the comparison between WeRoS, HEED, LEACH, LEACH-C and e-LEACH protocol in term of the lifetime per rounds. This results show that WeRoS algorithm outperforms the others algorithms regarding the number of alive nodes. This is the fact that in LEACH, LEACH-C as well as e-LEACH protocol, each CH directly sends the aggregated data to the BS. Therefore, the CHs that are further from BS used a higher transmission power. This implies the rapid depletion of batteries of the CH nodes. Nevertheless, the performance of HEED and WeRoS comparing with LEACH, LEACH-C and e-LEACH can be explained by the fact that these HEED and WeRoS use multi-hop communication between CHs and BS in the data transmission phase. But, HEED is an iterative algorithm; it requires a lot of information exchanges to form clusters, which are energy consuming. On the other hand, the WeRoS protocol considers the remaining energy of the node in routing process between CH to the BS. And, it considers the energy dispersion in the CHs selection stage. Moreover, it presents a low complexity algorithm to form a cluster, which allows it to outperform the HEED algorithm.

By means of Fig.4, if the lifetime metric is defined as number of rounds for which the first node died (FND), we can see that WeRoS can reach 471 rounds, while HEED reaches 270 rounds, and LEACH, LEACH-C and e-LEACH only reaches 200, 215 and 254 rounds, respectively. For half of the nodes being alive (HNA), WeRoS can reach 502 rounds, but LEACH, LEACH-C, e-LEACH and HEED reaches 250, 300, 320 and 366 rounds, respectively. For the Last Nodes Died (LND), WeRoS can reaches 511 rounds and HEED only reaches 421. By considering HNA metric, WeRoS carries out a HNA increase of approximately 85%, 67%, 56% and 37% on average compared to LEACH, LEACH-C, e-LEACH, and HEED. In brief, WeRoS can significantly increase a network lifetime compared to the other

	FND	HND	LND	CSD
LEACH	200	250	352	24,49
LEACH-C	215	300	332	18,99
e-LEACH	254	320	500	28,31
HEED	270	366	421	19,65
WeRoS	471	502	511	3,97

Figure 5. Coefficient of synchronous death

presented algorithms. Table(5) resumes those results, and also presents the coefficient of synchronous death (CSD).

4.4. Coefficient of synchronous death

The coefficient of synchronous death (CSD) of nodes can be defined as follows:

$$CSD = \frac{\dfrac{HND - FND}{HND} + \dfrac{LND - HND}{LND}}{2} \qquad (16)$$

This expression is used to evaluate the death rate dispersion of nodes around HNA metric. If CSD is equal to zero that means nodes are died simultaneously. Therefore, energy balance scheme is achieved, and by the way the energy hole problem is completely solved.

Table(5) shows that the CSD value of WeRoS is smaller than other protocols. This means that WeRoS effectively presents stable energy consumption across all the nodes. Resulting in the fact that the nodes die almost simultaneously which proves the efficiency of load balancing.

4.5. Average residual energy

The energy-efficiency provided by WeRoS is clearly confirmed by the results in Figure(6). This figure illustrates the global evolution of the average remaining energy of the network over time step (rounds) compared to LEACH, LEACH-C, e-LEACH and HEED. We clearly see that WeRoS and HEED improve the energy expenditure than the others one-hop considered protocols. This comes from the fact that HEED and WeRoS address the communication problem between CH and the BS. In WeRoS, just after data collection of all CH from member nodes, CHs that are further from the BS dynamically chooses the CH successors. To which, they deliver the aggregated data, by the construction of an elementary path. Thus, the CHs progressively route the data to the BS by adjusting the transmission power depending on the distance to CH successor. It helps balancing the energy consumption of CH. Speaking about the HEED protocol; the CH selection process depends on residual energy and distance by

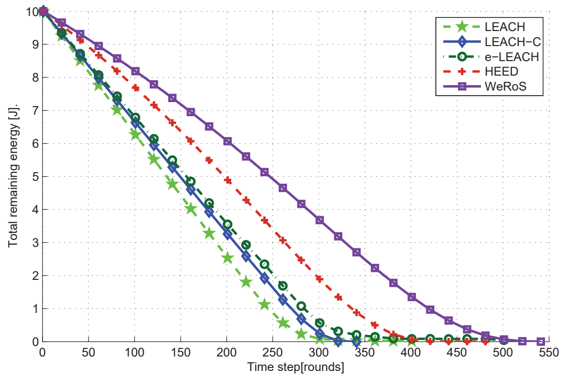

Figure 6. Average remaining energy of the network.

using iteration algorithm, which requires significant time when the nodes are far from the BS. And, the "Tentative Status" mechanism used for CHs election in HEED, favors the existence of isolated nodes. The isolated nodes can be self-elected to become CH, and can also communicate directly to the BS. Therefore, this direct communication between distant CHs and the BS is an energy-intensive. This is one of the main causes why WeRoS outperforms the HEED.

In Fig.6, as the network runs to 200 rounds, the average remaining energy of all the nodes in the network for LEACH, LEACH-C, e-LEACH, HEED and WeRoS, are respectively $2.4J$, $3.26J$, $4.2J$, $4.9J$ and $6.1J$. That means WeRoS exhibits 95%, 87%, 45% and 24% reduction in energy consumption over LEACH, LEACH-C, e-LEACH and HEED.

4.6. Amount of data collected at the BS

Figure 7 shows another comparison between LEACH, LEACH-C, e-LEACH, HEED and WeRoS protocol in term of amount of cumulative data collected at the BS according to the number of alive nodes in the network. The results presented in this figure demonstrate the effectiveness of WeRoS. It can deliver more amounts of data than others. According to HNA metric, WeRoS can deliver $7250kBit$ while LEACH, LEACH-C, e-LEACH, HEED cannot deliver more than $4121kBit$, $4624kBit$, $4782kBit$ and $5447kBit$. Hence, WeRoS increases the amount of delivered data by 76%, 57%, 51% and 33% comparing to LEACH, LEACH-C, e-LEACH and HEED, respectively.

5. Conclusion

Weighted Route Selection In Cluster-Based Protocol for WSNs has been proposed in this paper. It selects the CHs among nodes having the greatest residual energy by considering its coefficient of variation. It also rotates CH role periodically to

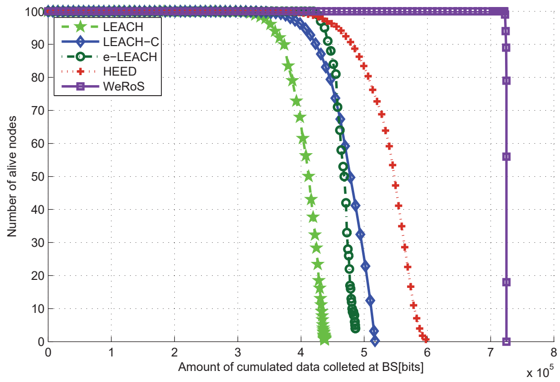

Figure 7. Amount of cumulative data collected at the BS.

balance the energy consumption between nodes. Furthermore, multi-hop routing process during data transmission was considered to reduce the global energy consumption. Based on the considered network assumptions, simulation results show that this protocol exhibits satisfactory comparative performances on energy consumption reduction, nodes synchronous death, and increases the network lifetime.

Inter-cluster and intra-cluster communication are always considered as no retransmission operations. However, a retransmission is required when a transmitted packet contains at least one error. In this context, an improvement of convex function and the threshold expression should be tuned: the energy consumed per bit transmitted to obtain a minimum variation of energy per bit transmitted successfully will be introduced.

Measurements on energy characterization of the cluster formation phase and data transmission phase currently are under study.

Acknowledgement. We acknowledge the financial support of structural funds of the European Community, the French Government and the Regional Council of Reunion Island (Region Reunion) for providing research grants (PO FEDER Convergence 2007/2013, 2013-32 933 Carerc Project).

References

[1] (2006) Cc2420 data sheet. URL http://www.chipcon.com/files/CC2420_Data_Sheet_14.pdf.

[2] Ali, M., Dey, T. and Biswas, R. (2008) Aleach: Advanced leach routing protocol for wireless microsensor networks. In *ICECE 2008. International Conf. on*: 909–914.

[3] Arampatzis, T., Lygeros, J., Member, S. and Manesis, S. (2005) A survey of applications of wireless sensors and wireless sensor networks. In *Proc. 13 th Mediterranean Conference on Control and Automation, Limassol*: 719–724.

[4] Baronti, P., Pillai, P., Chook, V.W., Chessa, S., Gotta, A. and Hu, Y.F. (2007) Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards. *Computer Communication* (30): 1655–1695.

[5] Chen, G., Li, C., Ye, M. and Wu, J. (2009) An unequal cluster-based routing protocol in wireless sensor networks. *Wireless Netw.* **15**(2): 193–207.

[6] Handy, M.J., Haase, M. and Timmermann, D. (2002) Low energy adaptive clustering hierarchy with deterministic cluster-head selection. In *Fourth IEEE Conference on Mobile and Wireless Communications Networks, Stockholm*.

[7] Heinzelman, W.B., Chandrakasan, A.P. and Balakrishnan, H. (2002) An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on wireless communication* **1**(4): 660 – 670.

[8] Kansal, A., Hsu, J., Zahedi, S. and Srivastava, M.B. (2007) Power management in energy harvesting sensor networks. *ACM Trans. Embed. Comput. Syst.* **6**(4).

[9] Li, L., Shu-song, D. and Xiang-ming, W. (2006) An energy efficient clustering routing algorithm for wireless sensor networks. *The journal of china universities of posts and telecommunications* **13**: 1005-8885(2006)03-0071-05.

[10] Liu, M., Cao, J., Chen, G. and Wang, X. (2009) An energy-aware routing protocol in wireless sensor networks. *Sensors* **9**(1): 445–462.

[11] Liu, X. (2012) A survey on clustering routing protocols in wireless sensor networks. *Sensors* **12**(8): 11113–11153.

[12] Mahmoudvand, R., Hassani, H. and Wilson, R. (2007) *Is The Sample Coefficient Of Variation A Good Estimator For The Population Coefficient Of Variation?* MPRA Paper 6106, University Library of Munich, Germany.

[13] Panichkitkosolkul, W. (2013) Confidence intervals for the coefficient of variation in a normal distribution with a known population mean. *Journal of Probability and Statistics* **2013**: 11.

[14] Raghunathan, V., Schurgers, C., Park, S. and Srivastava, M. (2002) Energy-aware wireless microsensor networks. *Signal Processing Magazine, IEEE* **19**(2): 40–50.

[15] Randriatsiferana, R., Lorion, R., Alicalapa, F. and Harivelo, F. (2013) Energy-efficient clustering algorithm based on energy variance for wireless sensor networks. In *Inter. Conf. on SaCoNet*, **01**: 1–5.

[16] Rashed, M.G., Kabir, M.H. and Ullah, S.E. (2011) An energy efficient protocol for cluster based heterogeneous wireless sensor network. *International Journal of Distributed and Parallel Systems* **2**(2): 54–60.

[17] Wang, W., Wang, B., Liu, Z., Guo, L. and Xiong, W. (2011) A cluster-based and tree-based power efficient data collection and aggregation protocol for wireless sensor networks. *Information technology journal* **10**(3): 557–564.

[18] Wu, X., Chen, G. and Das, S.K. (2008) Avoiding energy holes in wireless sensor networks with nonuniform node distribution. *IEEE Transactions on parallel and distributed systems* **19**(5): 710–720.

[19] Younis, O. and Fahmy, S. (2004) Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transaction on Mobile Computing* **3**(4): 366–379.