

ENERGY-EFFICIENT MULTI-HOP ROUTING WITH UNEQUAL CLUSTERING APPROACH FOR WIRELESS SENSOR NETWORKS

Younes El Assari , Samia Al Fallah, Jihane El Aasri, Mounir Arioua and
Ahmed El Oualkadi

Laboratory of Information and Communication Technologies, National School of
Applied Sciences, Abdelmalek Essaadi University, Tangier, Morocco

ABSTRACT

The development of an energy-efficient routing protocol, capable of extending the life of the network, is one of the main constraints of wireless sensor networks (WSN). Research studies on WSN routing prove that clustering offers an effective approach to prolong the lifetime of a WSN, particularly when it is combined with multi-hop communication that can reduce energy costs by minimizing the distance between transmitter and receiver. Most clustering algorithms using multi-hop in data transmission encounter the hotspot problem. In this work, an Energy-efficient Multi-hop routing with Unequal Clustering approach (EMUC) is proposed, to create clusters of different sizes, which depend on the distance between the sensor node and the base station. Equilibrate the energy dissipation between the cluster heads is the purpose of this approach by adopting multi-hop communication to relay data to the base station. The implementation of multi-hop mode to transmit data to the base station reduces the energy cost of transmission over long distances. The effectiveness of this approach is validated through performed simulations, which prove that EMUC balances energy consumption between sensor nodes, mitigates the hotspots problem, saves more energy and significantly extends the network lifetime.

KEYWORDS

Energy-efficient, Hotspot issue, Routing protocols, Unequal clustering, Wireless sensor networks, Multi-hop communication

1. INTRODUCTION

Due to their great usefulness in many fields such as military, medical and environmental sciences, traffic management, disaster prediction, manufacturing industry, the Wireless Sensor Network (WSN) has recently become a point of interest for many researchers. [1-4]. WSN is a set of tiny sensor nodes with low-cost and low power, connected by wireless links and distributed over a physical area. Those sensor nodes can assemble data from the environment, process data, and communicate with each other to route the collected data to the base station (BS). The sensors are typically equipped with small batteries, and it cannot be recharged due to unattended environments and dense deployment [5]. Therefore, the most important challenge in an energy-constrained sensor network is the reduction of energy consumption. In this context, several studies on routing protocols for WSNs have been carried out.

To increase energy efficiency and network lifetime, the network can be divided into several groups (clusters), with two types of nodes in the cluster: cluster head (CH) and cluster member. The CH is responsible for gathering, aggregating data from its cluster members, and sending the collected data to the BS [6,7].

The dissipation of energy in CH is higher than the cluster member when data is transmitted to longer distances [8,9]. The energy consumption must be balanced along with nodes to achieve a maximum network lifetime. Accordingly, all nodes inhomogeneous networks are supposed, to begin with the same initial energy, it is possible to regularly rotate the role of the CH between the sensor nodes to distribute the energy dissipation. Also, clustering can reduce the consumption of energy by eliminating redundant transmissions and merging the collected data into a single packet and then transmitting it to the sink [10-12].

The clustering techniques used in WSN help to increase the lifetime of the network. However, some clustering algorithms encounter some issues, for instance, additional overheads when selecting the CHs, forming clusters, and hotspot problems. Generally, the hotspot issue is occurring while using multi-hop communication; the sensor nodes do not send the sensed data directly to the BS. However, they send it with a multi-hop approach. The nodes closer to the BS have to communicate their own sensed data and should also forward data from other nodes. This implies the depletion of their energy quicker than other nodes [13]. Consequently, the nodes die prematurely and may cause network partitions [14]. Several hierarchical clustering protocols have been developed with a structure that reduces the path cost and the hotspot issue when communicating with the BS.

In this paper, an Energy-efficient Multi-hop Unequal Clustering (EMUC) based routing approach is presented. It is relevant for periodical data collection applications in WSNs. This approach is intended to distribute energy consumption within the network and extend the life of the WSN. The EMUC is a distributed energy-efficient algorithm with an unequal clustering routing mechanism. During the CH election phase, several tentative CHs are chosen according to a probabilistic model and compete to be CHs based on residual energy of nodes. The main feature of the routing protocol we propose is the association of the clustering approach with a multi-hop algorithm. In intra-cluster, the transmission of the detected data from the nodes to the CH is performed by multi-hop propagation. We employed the minimum transmission energy algorithm to create the shortest path between the node and its CH, this path is established using the Dijkstra routing algorithm. Furthermore, to reduce the energy consumption of CHs, inter-cluster communication is performed in multi-hop mode. The choice of routing path takes into account the compromise between the energy of relay nodes and the energy cost of relay links.

This paper work is structured as follows. Section 2 summarizes the related works based on the clustering and unequal clustering algorithms. The channel propagation and energy consumption models for WSN are described in section 3. Subsequently, the proposed EMUC approach is described in detail in section 4, which is divided into sub-sections, including the clustering module and the data transmission strategy for both intra-cluster and inter-cluster communications. In section 6, the results of the simulation are analyzed and discussed. Finally, section 7 concludes this paper.

2. RELATED WORKS

The clustering method is used in WSNs to design various energy-efficient protocols [15], and it has been proven to be a resource-aware and highly adaptive method [16]. The principle of clustering is to divide the area network into clusters. The cluster-based protocols have three major phases during the clustering process: CH election, cluster formation, and data transmission phase. Several techniques are used to implement each one of these phases. For example, in CH election it can be probabilistic or based on the choice of nodes with more residual energy. In clustering methods, communication can be partitioned into intra-cluster or inter-cluster. Intra-cluster communication is based on communication between the CH and its cluster member nodes. However, inter-cluster communication is based on data exchanges between CHs or between CHs and the BS.

Several routing protocols use clustering method. The difference between these various protocols resides in the criteria for CHs election of each protocol, in the communication between the nodes, and in data transmission to the BS. In order to minimize the power consumption of WSNs, several clustering protocols [17-21] and uneven clustering protocols [22-25] have been proposed.

LEACH (Low Energy Adaptive Clustering Hierarchy) is an adaptive and self-organizing clustering routing protocol for WSN [17]. This distributed clustering approach does not require centralized control. The CH role is rotated between all nodes to evenly distribute the energy load. Clusters are created after the selection of CHs, and this sequence of steps is similar for most cluster approaches, they select the CHs and then form the clusters [18]. Each CH gathers information from the members of its cluster and transmits it to the BS to reduce overall communication. Therefore, it is not suitable for networks deployed in larger regions, and has low energy balance and energy consumption [19].

PEGASIS (Power-Efficient GATHERing in Sensor Information Systems) is an extension to LEACH [20]. Unlike LEACH, PEGASIS employs multi-hop routing by creating chains and choosing only one node to broadcast to the BS instead of requiring several nodes. Every node can only establish communication with two adjacent neighbors in the chain until the aggregated data reach the BS. All nodes in the chain will send in turn to the BS. Thus, the network lifetime is extended through this protocol, which reduces the amount of energy expended per round. Accordingly, due to multi-hop propagation throughout the entire chain, network delay is very significant and is not relevant for applications that do not tolerate delays.

HEED (Hybrid Energy-Efficient Distributed clustering approach) is a multi-hop clustering protocol [21]. The CHs are periodically selected according to the residual energy of the node and the proximity of the node to its neighbors. It constructs clusters in a distributed manner, with each node decides according to local information. It prolongs the network lifetime by decreasing the amount of energy consumed in intra-cluster communications and maintains a good distribution of CH nodes. However, in the cluster formation phase, there are many more messages to be broadcasted which increases energy consumption. Also, the premature death of some nodes due to the hot spot problem significantly reduces the network lifetime [5].

EEUC (An Energy-Efficient Unequal Clustering protocol) creates clusters of uneven sizes to distribute energy consumption across the CHs [22]. The CHs are selected through a partial competition. The closest clusters to the BS are smaller in size than those further from the BS. For instance, the closer the CHs are to the base station, the more energy can be saved for data transmission between clusters. The CH located far from the BS selects relay nodes having more residual energy to forward its data.

UHEED (Unequal Clustering Algorithm for Wireless Sensor Networks) [23], is an unequal clustering algorithm based on the HEED algorithm, which prevents the death of the CHs closest to the BS. It allows creating clusters of unequal size by editing the competition range of the HEED protocol. Several competition ranges are defined based on the EEUC competition radius formula [22].

EADUC (Energy-Aware Distributed Unequal Clustering) [24] is a distributed unequal clustering method that supports both heterogeneous and homogeneous networks. The protocol consists of three sub-phases for cluster formation: gathering information on neighboring nodes, CH competition, and cluster formation. CHs that are further away from BS and have higher residual energy have larger clusters.

LUCA (Location-based unequal clustering algorithm) [25] is another unequal clustering algorithm, in which each cluster has a different size depending on its distance to the BS. LUCA creates clusters that become larger as the distance to the BS increases. Clusters are organized unevenly according to the location from BS, and the location is determined by GPS [26]. The LUCA nodes are location-aware, making it unsuitable for many real-time applications and increasing energy overload.

Finally, several routing protocols organize the network into clusters and address the issue of load imbalance in WSNs. Moreover, routing protocols use specific methods to save energy and extend the network lifetime. Each routing protocol has strengths and weaknesses, however, it is very important to design an energy-efficient routing algorithm that can reduce energy consumption, balance the load in sensor networks, and prevent premature death of nodes. To address these issues, we propose a routing scheme combining multi-hop with clusters of unequal sizes.

3. SYSTEM MODEL

3.1. Network Model

In this contribution, a WSN of N sensor node scattered over the studied area is considered. All nodes permanently sense their region and periodically transmit the data to the BS via their CH. Each sensor may have various functionalities depending on its role. If it is a cluster member, its role is to monitor its region and transmit the data to its CH. If it is a CH, it is necessary to collect data from its members as well as monitoring its region, aggregate the collected information into a packet and transmit it to the BS. Moreover, a CH can be used as a relay node to transmit other CH packets to the BS. The assumptions that were made for this network model are as follows:

- All sensors and the BS are immobile after deployment.
- The BS is placed away from the monitoring area.
- Sensors are homogeneous, they start with a similar initial energy level and they have the same capabilities.
- Sensors are kept unsupervised. Consequently, charging the battery is not possible.
- Every sensor is individually labeled by a unique identifier (ID).
- Sensors can diverge the emission power depending on the distance to the receiving node.
- The length between each sensor pair can be computed based on the received signal strength if the transmit power is known.
- The radio connections are symmetrical. Consequently, communication between two arbitrary nodes will require the same transmission power.

3.2. Channel Propagation Model

The electromagnetic waves propagate in the wireless channel can be modeled as a function of the power-law of the transmitter/receiver distance. The free space propagation model accepts that the ideal condition for propagation is that between the transmitter and the receiver there is only one free path in the line-of-sight. Both the direct path and a ground reflection path are taken into account in the two-ray ground reflection model. This model provides more accurate predictions than the free space model for long distances. The two-ray ground propagation model is used if the distance between transmitter and receiver is greater than a crossover distance (d_{co}) [27]. The crossover distance is defined as follows:

$$d_{co} = \frac{4\pi\sqrt{L}h_t h_r}{\lambda} \quad (1)$$

Where, $L \geq 1$ designates the loss factor of the system and it is not associated with propagation. h_t is the height of the transmitting antenna, h_r is the height of the receiving antenna, and λ is the transmitting frequency wavelength.

The transmission power is reduced when the distance is less than d_{co} . Otherwise, if the distance is higher than d_{co} , the two-ray ground propagation is used. The received signal power ($P_r(d)$) at distance d from the transmitter in free space ($d < d_{co}$), and two-ray ground propagation ($d \geq d_{co}$) is predicted by:

$$P_r(d) = \begin{cases} \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2} & \text{if } d < d_{co} \\ \frac{P_t G_t G_r h_t^2 h_r^2}{d^4} & \text{if } d \geq d_{co} \end{cases} \quad (2)$$

G_r and G_t are the antenna gains of the receiver and transmitter respectively. P_t is the transmitted signal power.

We used omnidirectional antennas in this paper, as described in [28], with the following parameters: The antenna gains are $G_t=G_r=1$, and the height of the antenna is $h_t=h_r=1.5$ m. We assumed that the system is lossless ($L=1$), a frequency of transmission of 914 MHz, and $\lambda=0.328$ m. If these values are replaced in (1), a crossover distance of 87 m is obtained.

3.3. Energy Consumption Model

The radio model we have adopted is the first-order radio model, as illustrated in Figure 1. In the literature, this model is generally used to deduct the reception and transmission power dissipated in communication in WSNs [20].

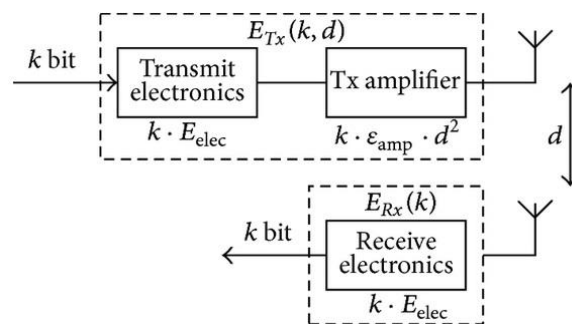


Figure 1. First order radio energy dissipation model

We assume that the sensor consumes energy during data transmission and reception. In transmission, the node consumes energy to operate the radio electronics and the power amplifier. However, in reception, the node consumes energy to operate the radio electronics. Thereby, the energy consumption for transmitting a message of size k -bits over a distance d , is defined by:

$$E_{Tx}(k, d) = E_{Tx-elec}(k) + E_{Tx-amp}(k, d) \quad (3)$$

$$E_{Tx}(k, d) = \begin{cases} E_{elec} k + \varepsilon_{fs} k d^2 & d < d_{co} \\ E_{elec} k + \varepsilon_{mp} k d^4 & d \geq d_{co} \end{cases} \quad (4)$$

The energy consumed in receiving a k -bits message in the receiver circuitry is given by:

$$E_{Rx}(k) = E_{Rx-elec}(k) = E_{elec} k \quad (5)$$

Where, E_{Tx-amp} is the electrical energy used to transmit k -bits message over a distance d , E_{elec} represents the energy necessary per bit to transmit and receive the electronics required to process the information, and the ε_{fs} and ε_{mp} are the energies required in the transmit amplifier to communicate at short range and long-range, respectively.

The ε_{fs} and ε_{mp} parameters vary depending on the sensitivity required of the receiver and the receiver noise factor, as the transmit power should be adjusted so that the receiver power is above a certain threshold $P_{r-thresh}$. The transmit power P_t corresponds to the transmit energy per bit $E_{Tx-amp}(1, d)$ multiplied by the radio bit rate R_b :

$$P_t = E_{Tx-amp}(1, d) R_b \quad (6)$$

Replacing in the value of $E_{Tx-amp}(1, d)$, we obtain the following:

$$P_t = \begin{cases} \varepsilon_{fs} R_b d^2 & d < d_{co} \\ \varepsilon_{mp} R_b d^4 & d \geq d_{co} \end{cases} \quad (7)$$

Based on the propagation model outlined above in (2), the received power is formulated as follows:

$$P_r = \begin{cases} \frac{\varepsilon_{fs} R_b G_t G_r \lambda^2}{(4\pi)^2} & d < d_{co} \\ \varepsilon_{mp} R_b G_t G_r h_t^2 h_r^2 & d \geq d_{co} \end{cases} \quad (8)$$

By setting (8) equals to $P_{r-thresh}$ we can determine the parameters ε_{fs} and ε_{mp} :

$$\varepsilon_{fs} = \frac{P_{r-thresh} (4\pi)^2}{R_b G_t G_r \lambda^2} \quad (9)$$

$$\varepsilon_{mp} = \frac{P_{r-thresh}}{R_b G_t G_r h_t^2 h_r^2}$$

Therefore, depending on the receiver threshold and the distance between transmitter and receiver, the required transmit power P_t is given by:

$$P_t = \begin{cases} \frac{(4\pi)^2}{G_t G_r \lambda^2} P_{r\text{-thresh}} d^2 & d < d_{co} \\ \frac{1}{G_t G_r h_t^2 h_r^2} P_{r\text{-thresh}} d^4 & d \geq d_{co} \end{cases} \quad (10)$$

The received power threshold $P_{r\text{-thresh}}$ is defined by assessing the noise at the receiver. While the thermal noise threshold equals to 99 dBm, the receiver noise factor is 17 dB and the SNR required is at least 30 dB in order to receive the signal without errors [28], for efficient reception, the minimum received power $P_{r\text{-thresh}}$ required for successful reception is:

$$P_{r\text{-thresh}} \geq -52 \text{ dBm} \quad (11)$$

Hence, the received power must have a value not less than -52 dBm (or 6.3 nW) to ensure successful packet reception. By substituting the real parameter's values of (9) ($G_t = G_r = 1$, $h_t = h_r = 1.5$ m, $\lambda = 0.328$ m, and $R_b = 1$ Mbps), the values of ε_{fs} and ε_{mp} can be determined:

$$\begin{aligned} \varepsilon_{fs} &= 10 \text{ pJ} / \text{bit} / \text{m}^2 \\ \varepsilon_{mp} &= 0.0013 \text{ pJ} / \text{bit} / \text{m}^4 \end{aligned} \quad (12)$$

4. PROPOSED ALGORITHM

The proposed EMUC protocol is a distributed energy-efficient algorithm with unequal clustering routing mechanism that balances the energy consumption. To address the hotspot problem, the algorithm organizes the network into clusters of unequal size, with the clusters closest to the BS being composed of a reduced number of nodes. The farthest nodes from the BS are unable to communicate directly with the BS as the transmission range is limited. Therefore, it is important to use the closest CHs as a relay to forward their data. During intra-cluster data processing, the CHs closer to the BS will consume less energy to conserve more energy for inter-cluster relay traffic. Figure 2 presents an overview of the EMUC mechanism, where the arrows between nodes represent the communication links, and each CH represents a cluster. The number of member nodes that join a CH represents the size of the cluster.

While the network is under deployment, an advertisement message is sent by the BS, with a certain power level, to all sensor nodes. This allows every sensor node to determine the estimated distance to the BS using the received signal intensity, which assists the nodes in using the appropriate power setting when communicating with the BS. The advertisement message provides some information necessary for the clustering process to create clusters of unequal sizes, such as the farthest and closest distance between sensor nodes and the BS.

The EMUC process is divided into rounds and each round is composed of steps:

1. A CH selection and cluster definition phase.
2. Data transmission phase where the data is routed from the nodes to the BS via the CHs.

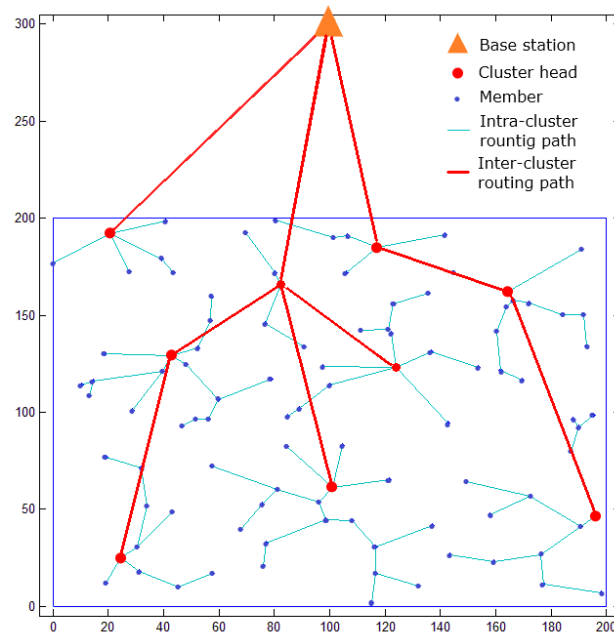


Figure 2. The EMUC network topology; random 100-node topology for a 200m×200m.

4.1. Cluster Head Election

The proposed approach is a probabilistic clustering algorithm, during each round, sensor nodes decide in a probabilistic manner whether or not will participate in the competition to be elected as a CH. This decision is based on a predefined threshold T , the sensor node that has decided to compete to become a CH first becomes a tentative CH. In local areas, tentative CHs are competing with each other to become a CH based on residual energy.

Each sensor node calculates its competition range (R_{comp}) which is determined according to the relative value of the distance between the sensors and the BS. Based on this distance, the network will produce clusters of different sizes. The R_{comp} for an arbitrary node n_i is given by:

$$R_{comp}(n_i) = \left(1 - c \frac{d_{max} - d(n_i, BS)}{d_{max} - d_{min}} \right) R_0 \quad (13)$$

The maximum and minimum distance between the base station and the sensor nodes is represented respectively by d_{max} and d_{min} . c is a constant factor that takes values between 0 and 1, $d(n_i, BS)$ denotes the distance between n_i and the BS, and R_0 represents the maximum competition radius that is predefined.

Each node generates a random number that is greater than 0 and less than 1, if the number generated is below the threshold T (approximate percentage of the desired tentative CH), the node becomes a tentative CH. Afterward, the CH competition begins. Every tentative CH sends a message “COMPETITION_CH_MSG” to its surrounding tentative CHs within its maximum competition radius, the message includes the node ID, residual energy and competition radius. If a tentative CH th_i receives a “COMPETITION_CH_MSG” from the node th_j , which is within its competition radius, and the residual energy of th_i is greater than that of th_j , th_j gives up the CH competition and broadcast a message “QUIT_CH_MSG”. Then if th_i has the highest residual energy among the tentative CHs within its competition radius, it becomes a CH.

After the election of the CH is completed, the CHs broadcast a message “CH_ADV_MSG” over the network. Every non-CH node selects the nearest CH that has the highest received signal power and then sends a “JOIN_CLUSTER_MSG” message to the CH. Algorithm 1 depicts the CH election phase. R_{comp} and $resEnergy$ are the competition range and the residual energy of a particular sensor node, respectively.

When the cluster creation process is achieved, all CHs broadcast TDMA messages to allocate a time slot for their cluster members. Once the cluster nodes are notified of the TDMA schedule, the setup phase is completed, and data transmission can be started.

Algorithm 1 : Cluster head election

1. Calculate R_{comp}
 2. $T \leftarrow$ probability to become a tentative cluster-head
 3. $\mu \leftarrow rand(0, 1)$
 4. **if** $\mu < T$ **then**
 5. $beTentativeHead \leftarrow TRUE$
 6. **end if**
 7. **if** $beTentativeHead = TRUE$ **then**
 8. broadcast *COMPETITION_CH_MSG*($ni.ID, ni.R_{comp}, ni.resEnergy$)
 9. **else**
 10. EXIT
 11. **end if**
 12. on receiving a *COMPETITION_CH_MSG* from node n_j
 13. **if** $d(ni, n_j) < \max(ni.R_{comp}, n_j.R_{comp})$ **then**
 14. **if** $ni.resEnergy > n_j.resEnergy$ **then**
 15. broadcast a *CH_ADV_MSG*($ni.ID$)
 16. EXIT
 17. **end if**
 18. **end if**
 19. on receiving a *CH_ADV_MSG* from node n_j
 20. **if** $d(ni, n_j) < \max(ni.R_{comp}, n_j.R_{comp})$ **then**
 21. broadcast a *QUIT ELECTION MSG*($ni.ID$)
 22. EXIT
 23. **end if**
-

In order to find the optimum values of R_0 and c parameters, we have studied the longevity of the network for various settings of c and R_0 . Figure 3 illustrates how these parameters impact the network lifetime. There is a correlation between c , R_0 and network longevity. R_0 is the main factor influencing the network lifetime, this is due to the fact that R_0 determines the number of clusters in each network scale. The longest network lifetime is reached when R_0 is set to 90 m. On the other hand, c determines the variation of cluster sizes. We can study the evolution of the lifetime of the network versus c values under each setting of R_0 . When the value of c is equal to 0, the algorithm behaves as an equal clustering approach, since all nodes have the same competition range ($R_{comp} = R_0$). As the value of c rises from 0, the energy consumption among the CHs is progressively balanced, which increases the lifetime of the network. However, if c is close to 1, this results in the formation of several clusters closer to the BS, each cluster provides an aggregated packet to the BS, leading to a loss of energy and a decrease in the lifetime of the network. Therefore, according to Figure 3, for a given R_0 , there is an optimal value of c that could optimally extend the network lifetime. We observed that the network performs perfectly for three values of R_0 equal to 70, 80 and 100 m when the values of c are equal to 0.2, 0.5 and 0.7, respectively. For this simulation, the most efficient value of c is about 0.4 when R_0 is equal to 90 m, which is considered as the best performed combination.

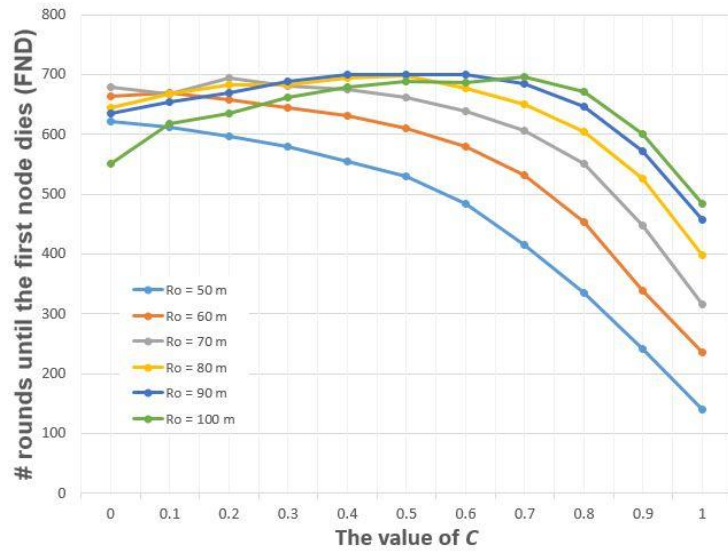


Figure 3. The network lifetime with various values of R_0 versus c

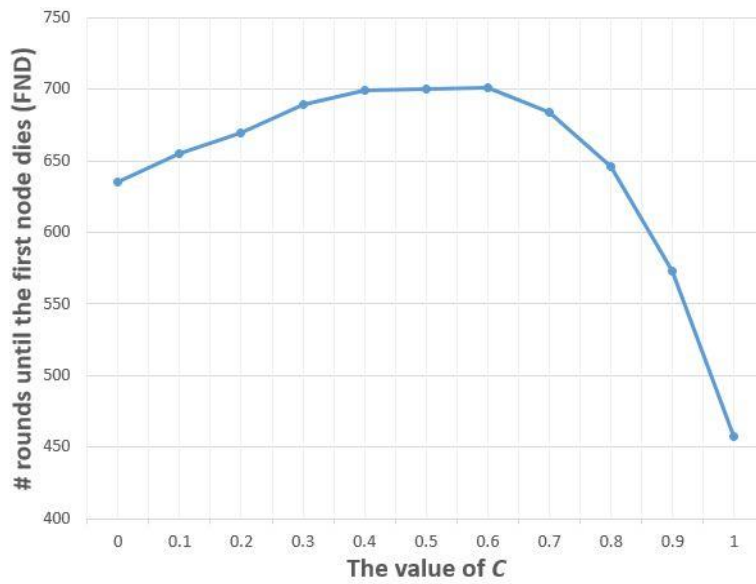


Figure 4. The network lifetime versus c for $R_0=90$ m

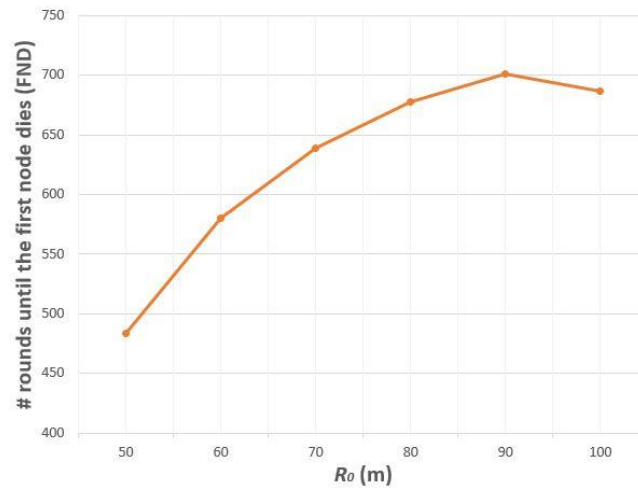


Figure 5. The network lifetime versus R_0 for $c=0.6$

To verify the relationship between c and the longevity of the network, we vary the value of c from 0 to 1 (Figure 4). In the case where R_0 is set at 90 m and c is increasing from 0. The effect of varying c value becomes distinct. However, the network lifetime increases with the rise of c until a value equals to 0.6, after that the network lifetime decrease when the value of c is too high. In fact, when c equals 0, the algorithm generates clusters of equal size, and when c is greater than 0 the size of clusters varies which generates more clusters. Each cluster must provide a data packet to the BS, with a high number of clusters, the network consumes more energy in inter-cluster communication, causing energy dissipation. Consequently, there is an appropriate value of c if R_0 is specified to obtain an adequate number of clusters that is approximately 0.4 in this case as shown in Figure 4. In addition, the effect of varying R_0 on network longevity is investigated in Figure 5. The result shows that R_0 significantly affects the durability of the network.

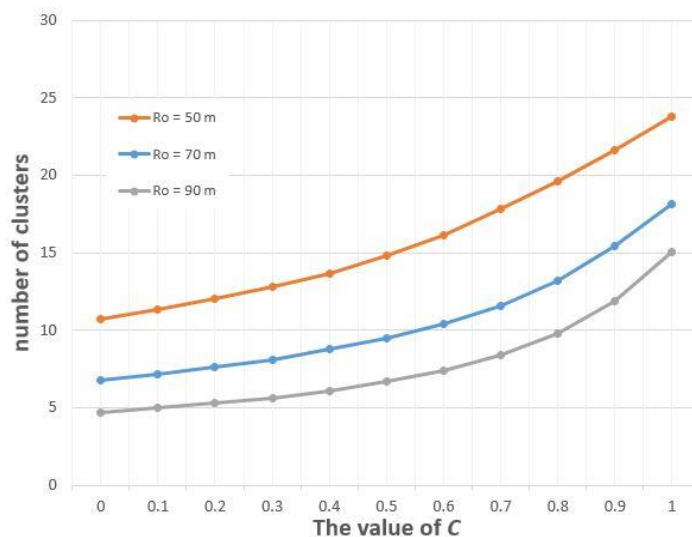


Figure 6. The number of generated clusters versus c for three values of R_0 : 50 m, 70 m, and 90 m

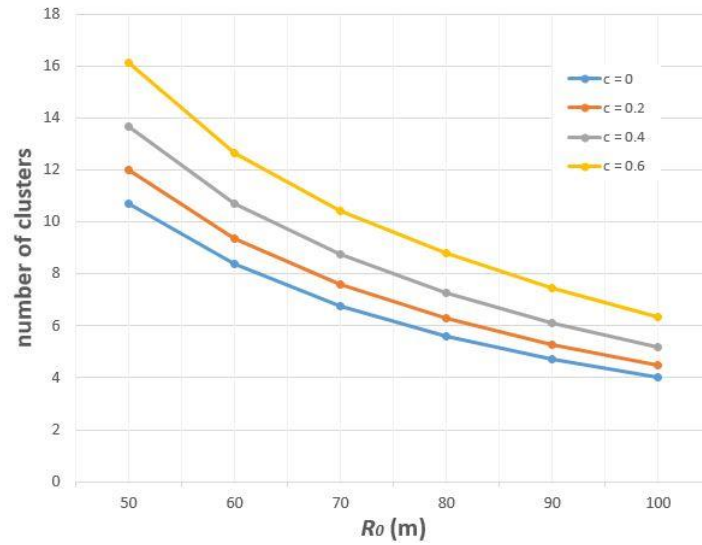


Figure 7. The number of generated clusters versus R_0 for values of C : 0, 0.2, 0.4, and 0.6

All CHs are in charge of aggregating the data of their members into a single packet, and each cluster sends its packet to the BS. Therefore, if there are more clusters, more messages must be transmitted to the BS, the large number of clusters leads to an increase in overall energy consumption. Based on the defined c and R_0 , there is a variation in the selected number of CHs. The mean number of CHs produced is shown in Figures. 6 and 7. When R_0 is given and the value of c goes from 0 to 1, the radius of competition reduced consequently, and further clusters are generated as shown in Figure 6. When the value of c is set and the value of R_0 increases, the number of clusters decreases, as shown in Figure 7. With a small competition radius, the network should produce more clusters to cover the network.

4.2. Intra-Cluster Communication

In Intra-Cluster communication, several clustering protocols require all source nodes to transmit their data directly to the corresponding CH, which influences the amount of energy consumed by source nodes through high cost of long-distance transmissions [16,29]. Hence, those source nodes located away from the CH hastily exhaust their energy compared to other nodes. In order to overcome this energy limitation in intra-cluster communication, the proposed EMUC protocol uses multi-hop communication using Minimum Transmission Energy (MTE) algorithm to create the shortest path between the node and its CH. This path is determined using the Dijkstra routing algorithm of the shortest path. Within each cluster, the nodes transmit their own sensed data to the CH via relay nodes. To minimize transmission energy, each node selects the nearest node on the path to the CH to forward its message.

Considering the network density specifications and the distance of communication, the direct transmission from any member node to CHs might not be convenient for large-scale network structures [16,30]. Consequently, in the case of a large-scale WSN, it is important to have a structure of multi-hop communication that does not restrict the size and area coverage of the cluster. MTE routing is an efficient multi-hop algorithm that is commonly applied in WSN [16,31]. This algorithm consists of using certain nodes that act as routers to forward data. These intermediate nodes not only route data from other sensors to the CHs, but also monitor the environment. The router nodes are selected to minimize the power of the transmission amplifier.

If the minimum distance between a source node and its CH is by one hop, the source node transmits its data directly to the CH. Otherwise; the source node adopts the MTE algorithm to transmit the data to the CH with several minimum hops. The nodes have then to perform m transmissions within a distance of d and $m-1$ receptions. In each cluster, the amount of energy consumed for reception and transmission in MTE mode is expressed by:

$$E_{MTE} = E_{Tx-MTE} + E_{Rx-MTE} = mE_{Tx} + (m-1)E_{Rx} \quad (14)$$

For free space propagation:

$$\begin{aligned} E_{MTE} &= m(kE_{elec} + k\varepsilon_{fs}d^2) + (m-1)kE_{elec} \\ &= k((2m-1)E_{elec} + \varepsilon_{fs}md^2) \end{aligned} \quad (15)$$

For multi-path propagation:

$$\begin{aligned} E_{MTE} &= m(kE_{elec} + k\varepsilon_{mp}d^4) + (m-1)kE_{elec} \\ &= k((2m-1)E_{elec} + \varepsilon_{mp}md^4) \end{aligned} \quad (16)$$

4.3. Inter-Cluster Communication

After all nodes in the cluster have sent their data, the CH aggregates the data collected from its cluster members and sends its packet to the BS. When the CH is situated far from the BS, the CH should use a relay node to forward its packet to the BS, the CH reduces the cost of transmission energy by reducing the transmission distance, this relay node is another CH, which is on the way to the BS. The energy expended by the CH for receiving packets and transmitting k -bit message over distance d is given by:

$$E_{CH} = \begin{cases} mkE_{elec} + (m+1)kE_{DA} + \\ kE_{elec} + k\varepsilon_{fs}d^2 & \text{if } d < d_{co} \\ mkE_{elec} + (m+1)kE_{DA} + \\ kE_{elec} + k\varepsilon_{mp}d^4 & \text{if } d \geq d_{co} \end{cases} \quad (17)$$

Where m is the number of cluster members, E_{DA} is the energy cost in aggregating the received data, and d is the distance between the CH and the BS or the relay.

Several authors assume that CHs could aggregate several incoming packets into a single fixed-length outgoing packet [24], which is useful when the data is highly correlated. This is more acceptable for intra-cluster communication when the network can have multiple adjacent sensors in the same cluster, producing duplicate monitoring data. Data aggregation can, therefore, be applied to reduce the communication load and avoid redundancy [22]. However, in cases where higher reliability is desired or where the nodes are not close enough and have different detection data, the aggregation cannot be considered as a consistent solution. Like in inter-cluster communication, where the CHs are far away and have to deal with various data coming from other CHs. Therefore, we consider this scenario of inter-cluster communication: The CH (relay)

combines the received packet into one packet, with a packet length equals to $r*k$, where r is the number of packets received from its predecessors (CHs). Then the relay sends this packet to the next relay or directly to the BS. According to this scenario, the energy consumed by the CH relay in transmission and reception during one round is expressed as:

$$E_{relay} = \begin{cases} (m+r)k E_{elec} + (m+1)k E_{DA} + \\ (r+1)(k E_{elec} + k \varepsilon_{fs} d^2) & \text{if } d < d_{co} \\ (m+r)k E_{elec} + (m+1)k E_{DA} + \\ (r+1)(k E_{elec} + k \varepsilon_{mp} d^4) & \text{if } d \geq d_{co} \end{cases} \quad (18)$$

Where m is the number of cluster members and E_{DA} is the energy cost to aggregate the collected data.

5. SIMULATIONS AND RESULTS

In this section, to examine and strengthen the proposed EMUC protocol, we have implemented it with MATLAB. We have considered the first order radio model, a perfect MAC layer, and a faultless communication link. The energy was expended whenever a sensor sent, received, and aggregated data. The performance of EMUC was evaluated against LEACH and EEUC protocols.

The sensors are randomly distributed over a 200 m x 200 m area; there is only one BS, located at (100 m, 300 m). We set the threshold T to 0.4, R_0 to 90 m, and c to 0.4 in equation (13). These values are estimated through the performed analysis and simulations (Figure3). Table I presents the parameters used for the simulations of the proposed model.

Table 1. Simulation Parameters.

Parameters	Values
Number of nodes	100
Network size	200 m,200 m
Base station position	100 m,300m
D_{CO}	87 m
Initial energy	0.5 J
K	4000 bits
E_{elec}	50 nJ/bit
ε_{mp}	0.0013 pJ/bit/m ⁴
ε_{fs}	10 pJ/bit/m ²
E_{DA}	5 nJ/bit/message

Table II summarizes the results of nodes elimination as a function of rounds, represented by important key performance metrics such as: FND (first node died), HND (half of nodes died), LND (last node died), SP (rate of stability period), and EDR (energy depletion rate) of LEACH, EEUC, and EMUC.

Table 2. Nodes Die-Off Statistics

Protocol	FND	HND	LND	% SP	EDR
LEACH	150	450	743	20	0.067
EEUC	425	688	780	54	0.064
EMUC	478	736	859	56	0.058

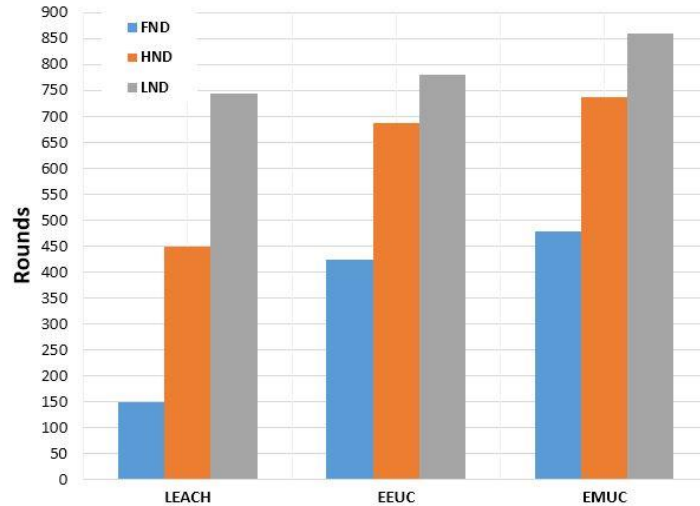


Figure 8. Lifetime metrics (FND, HND and LND) for LEACH, EEUC, and EMUC

The first EMUC node died after 475 rounds; however, in the case of the first node of LEACH and EEUC, it died after 150 and 425 rounds, respectively. The duration between the network start and the FND represents the stability period that is prolonged in EMUC. EMUC increases the stability period by 68.6% compared to LEACH and 11.1% compared to EEUC. Similarly, with respect to the HND metric, EMUC extends the network lifetime by 38.9% until half of the nodes are dead compared to LEACH and 6.5% compared to EEUC. However, with the LEACH protocol, all sensor nodes are dead after 743 rounds, while at the same round, about 50% of the nodes are still alive for EMUC. As a result, the approach has extended the network lifetime by 13.5% compared to LEACH and 9.2% compared to EEUC. The performance of LEACH is the lowest correlated to the EEUC and EMUC since it uses an entirely probabilistic approach for clustering then data are disseminated straightly from the CHs to the BS. Consequently, the performed sensor network, which uses multi-hop routing to route data packets and which adopts the unequal clustering approach, demonstrates its reliability and extends the life of the WSN. The bar chart in Figure 8 shows a detailed comparison of the three algorithms considering FND, HND, and LND metrics.

By taking into account the maximum rounds of each protocol, the percentage of the stability period in relation to the total lifetime of the network, is 56% for EMUC, 54% for EEUC and 20% for LEACH. These obtained results show a significant extension for the stability period of EMUC.

To evaluate the efficiency of EMUC in comparison with other protocols, we have considered the energy depletion rate. The obtained results in Table II indicate that EMUC reduces the EDR by 16% and 10% in comparison to LEACH and EEUC respectively.

Figure 9 presents the number of active nodes as a function of rounds for LEACH, EEUC, and EMUC. As observed, our approach achieves better energy efficiency than LEACH and EEUC. Indeed, EMUC uses multi-hop for both intra-cluster and inter-cluster communications to minimize transmission costs. Instead of sending data directly from the nodes to their appropriate CH, the nodes use intermediate relay nodes, allowing the nodes to communicate with each other over short distances. In fact, the distances upon which most nodes transmit data to their CHs are much smaller than those of LEACH and EEUC. In large-size clusters, the distance between some nodes and their CH is greater, which leads these nodes to use more energy to communicate with the CH.

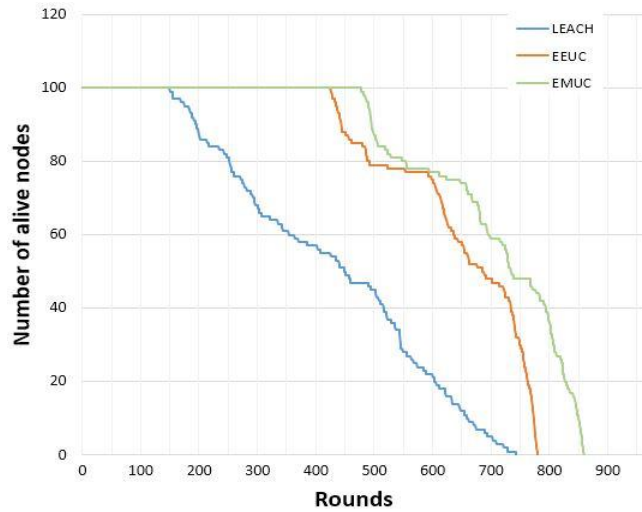


Figure 9. Number of alive nodes per round

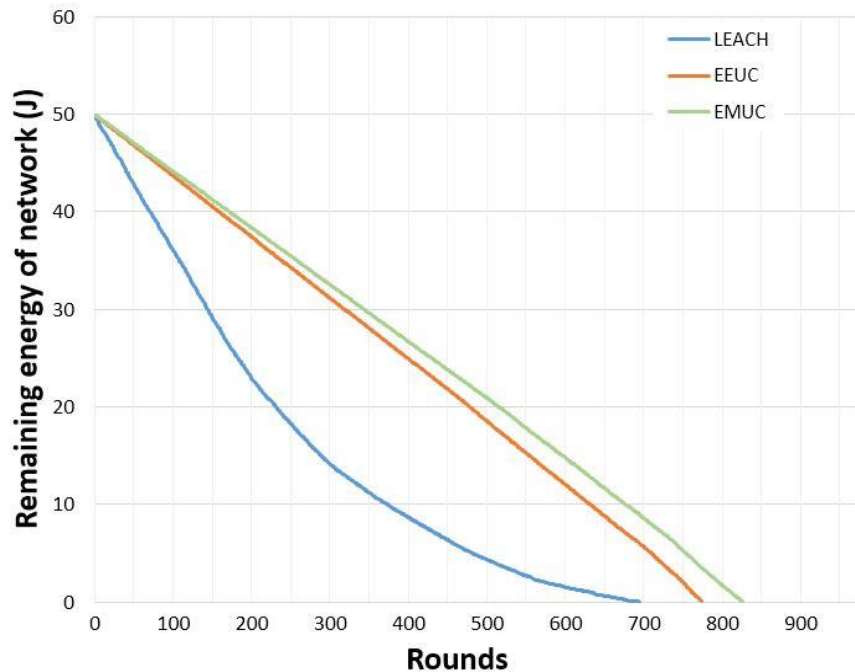


Figure 10. The total remaining energy of the network

Figure 10 shows the total system energy as a function of rounds for LEACH, EEUC, and EMUC. The balance of energy consumption in the network is influenced by the energy consumed at each round. Therefore, the energy consumed per round should be balanced in order to maximize the network lifetime. Due to the fact that EMUC balances the energy consumption between CHs and nodes, the durability of the EMUC network is extended and lasts longer than the LEACH and EEUC protocols. In addition, EMUC improves the reliability of transmission in the clustering process. In this case, the minimization of inter-nodes distance ensures that the transmission inside the clusters is mostly free space propagation. Using this propagation model, the energy efficiency of the network is effectively increased.

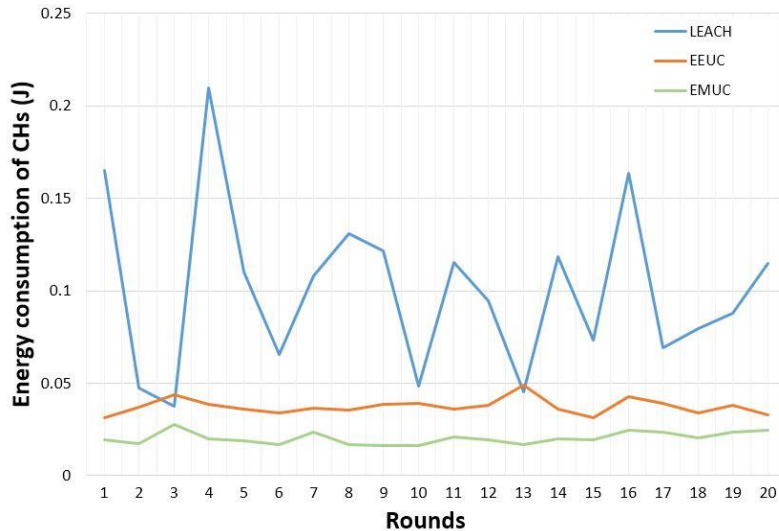


Figure 11. The energy consumed by CHs

Figure 11 shows the energy consumption of all CHs for the three protocols over 20 rounds. The energy consumed by CHs using the EMUC algorithm is lower than the energy consumed by EEUC's CHs, and both protocols consume less energy than LEACH. The CHs in LEACH send their data directly to the BS via one-hop communication, which implies a very high energy consumption. Both EMUC and EEUC use multi-hop in inter-cluster communication to send data to the BS, allowing CHs to save a significant amount of energy. In EMUC, the amount of energy consumed by CHs fluctuated slightly, reflecting that CHs consume nearly the same amount of energy in each round. For EEUC, energy consumption is higher than EMUC but shows a similar trend throughout the rounds. Unlike LEACH, the energy consumption fluctuates dramatically since, in LEACH, the number of elected CHs may vary from one round to the next.

5. CONCLUSIONS

In this paper, an unequal and efficient clustering protocol for WSNs has been proposed, to solve the hotspot issue, to save and to balance the energy usage between all network nodes. The EMUC protocol is a distributed energy-efficient algorithm with an unequal clustering routing scheme; it is relevant for periodic data collection applications in WSN. The proposed algorithm divides the network into several clusters with different cluster sizes. The clusters that are closer to the BS have smaller cluster sizes and the furthest has larger cluster sizes. The nodes compete to be CHs based on a probabilistic model and the residual energy of nodes. The minimum transmission energy algorithm is adopted to calculate the shortest path between the CH and its member nodes. The simulation results prove that our proposed approach performs better than

LEACH and EEUC, offers better energy efficiency and longer network life. These results confirm the benefits of the proposed solution and reveal that the EMUC has provided a larger stability period than the EEUC and LEACH. Our future work will consist of configuring automatically the parameters of c and R_0 in order to adapt the EMUC to each sensor network application.

REFERENCES

- [1] Q. Jiang, S. Zeadally, J. Ma, and D. He, "Lightweight threefactor authentication and key agreement protocol for internetintegrated wireless sensor networks," *IEEE Access*, vol. 5, pp. 3376–3392, 2017.
- [2] Y. El Assari, M. Arioua, A. El Oualkadi, and I. Ez-zazi, "Zone Divisional Approach for Energy Balanced Clustering Protocol in Wireless Sensor Network," in *Proceedings of the International Conference on Future Networks and Distributed Systems*. ACM, London, UK, 2017, p. 19.
- [3] Z. Chen, M. Ma, X. Liu, A. Liu, and M. Zhao, "Reliability Improved Cooperative Communication over Wireless Sensor Networks," *Symmetry*, vol. 9, no. 10, p. 209, 2017.
- [4] G. Han, J. Jiang, C. Zhang, T. Q. Duong, M. Guizani, and G. K. Karagiannidis, "A Survey on Mobile Anchor Node Assisted Localization in Wireless Sensor Networks," *IEEE Communications Surveys and Tutorials*, vol. 18, no 3, p. 2220-2243, 2016.
- [5] P. Neamatollahi, and M. Naghibzadeh, "Distributed unequal clustering algorithm in large-scale wireless sensor networks using fuzzy logic," *The Journal of Supercomputing*, Vol. 74, no 6, pp. 2329-2352, 2018.
- [6] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no 12, p. 2292-2330, 2008.
- [7] Z. Xinlian, W. Min, and X. Jianbo, "BPEC: an energy-aware distributed clustering algorithm in WSNs," *Journal of Computer Research and Development*, vol. 46, no 5, p. 723-730, 2009.
- [8] Z. Han, J. Wu, J. Zhang, L. Liu, and K. Tian, "A general self-organized tree-based energy-balance routing protocol for wireless sensor network," *IEEE Transactions on Nuclear Science*, vol. 61, no 2, p. 732-740, 2014.
- [9] A. Al-Baz and A. El-Sayed, "A new algorithm for cluster head selection in LEACH protocol for wireless sensor networks," *International journal of communication systems*, vol. 31, no 1, p. e3407, 2018.
- [10] L. Shen, J. Ma, X. Liu, F. Wei, and M. Miao, "A Secure and Efficient ID-Based Aggregate Signature Scheme for Wireless Sensor Networks," *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 546–554, 2017.
- [11] X. Li, W. Liu, M. Xie, et al., "Differentiated data aggregation routing scheme for energy conserving and delay sensitive wireless sensor networks," *Sensors*, vol. 18, no 7, p. 2349, 2018.
- [12] M. Ambigavathi and D. Sridharan, "Energy-aware data aggregation techniques in wireless sensor network," In *Advances in Power Systems and Energy Management*, Springer, Singapore, pp. 165-173, 2018.
- [13] J. Yu, Y. Qi, G. Wang, Q. Guo, and X. Gu, "An energy-aware distributed unequal clustering protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 7, no 1, 2011.
- [14] D. Agrawal and S. Pandey, "FUCA: Fuzzy-based unequal clustering algorithm to prolong the lifetime of wireless sensor networks," *International Journal of Communication Systems*, vol. 31, no 2, p. e3448, 2018.
- [15] S. Tyagi, and N. Kumar, "A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 36, no 2, p. 623-645, 2013.
- [16] M. Arioua, Y. El Assari, I. Ez-zazi, and A. El Oualkadi, "Multi-hop Cluster Based Routing Approach for Wireless Sensor Networks," *Procedia Computer Science*, 2016, vol. 83, p. 584-591.
- [17] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, (2000, January). "Energy-efficient communication protocol for wireless microsensor networks," In *System sciences, Proceedings of the 33rd annual Hawaii international conference on* (pp. 10-pp). IEEE.

- [18] Y. Chen, C. Shen, K. Zhang, H. Wang, Q. Gao “LEACH Algorithm Based on Energy Consumption Equilibrium,” Proceedings of the 2018 International Conference on Intelligent Transportation, Big Data Smart City (ICITBS); Xiamen, China. 25–26 January 2018; pp. 677–680.
- [19] S. Liu, “Novel unequal clustering routing protocol considering energy balancing based on network partition & distance for mobile education,” Journal of Network and Computer Applications. Vol 88, no 15, pp. 1–9, 2017.
- [20] S. Lindsey and C. S. Raghavendra, “PEGASIS: power-efficient gathering in sensor information systems,” in Proceedings of the IEEE Aerospace Conference, vol. 3, Big Sky, Mont, USA, March 2002, pp. 1125–1130.
- [21] O. Younis and S. Fahmy, “HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,” IEEE Transactions on Mobile Computing, vol. 3, no. 4, pp. 366–379, 2004.
- [22] C. F. Li, M. Ye, and G. Chen, “An energy-efficient unequal clustering mechanism for wireless sensor networks,” in Proceedings of the 2nd IEEE International Conference on Mobile Ad hoc and Sensor Systems (MASS '05), Washington, DC, USA, 2005.
- [23] E. Ever, R. Luchmun, L. Mostarda, A. Navarra, and P. Shah, “UHEED - An Unequal Clustering Algorithm for Wireless Sensor Networks,” in Sensornets, Rome, Italy, February 2012.
- [24] J. Yu, Y. Qi, G. Wang, Q. Guo, and X. Gu, “An energy-aware distributed unequal clustering protocol for wireless sensor networks,” International Journal of Distributed Sensor Networks, vol. 7, no. 1, Article ID 202145, 2011.
- [25] S. Lee, H. Choe, B. Park, Y. Song, and C.-K. Kim, “LUCA: an energy-efficient unequal clustering algorithm using location information for wireless sensor networks,” Wireless Personal Communications, vol. 56, no. 4, pp. 715–731, 2011.
- [26] S. Arjunan, P. Sujatha, “A Survey on Unequal Clustering Protocols in Wireless Sensor Networks,” Journal of King Saud University - Computer and Information Sciences, 2017.
- [27] R. Patel, S. Pariyani, and V. Ukani, “Energy and throughput analysis of hierarchical routing protocol (LEACH) for wireless sensor network,” International Journal of Computer Applications, vol. 20, no 4, p.32-36, 2011.
- [28] W. B. Heinzelman, A. P. Chandrakasan, H. Balakrishnan, “An application-specific protocol architecture for wireless micro-sensor networks,” IEEE Transactions on Wireless Communications, vol. 1, no 4, p. 660-670, 2002.
- [29] A. Mahboub, M. Arioua, E. M. En-naimi, I.Ez-Zazi. “Performance evaluation of energy-efficient clustering algorithms in wireless sensor networks,” Journal of Theoretical and Applied Information Technology, Vol. 83, no 2, 2016.
- [30] D. Wei, Y. Jin, S. Vural, K. Moessner, and R. Tafazolli, “An energy-efficient clustering solution for wireless sensor networks,” IEEE Transaction on Wireless Communication, Vol. 10, no. 11, pp. 3973–3983, 2011.
- [31] M. Zimmerling, D. Walteneagus, and J.M. Reason. “Energy-Efficient Routing in Linear Wireless Sensor Networks,” IEEE International Conference on Mobile Adhoc and Sensor Systems, Oct. 8-11, 2007, pp.1-3.