DCOPA: a distributed clustering based on objects performances aggregation for hierarchical communications in IoT applications

Foudil Mir[1] · Farid Meziane[2]

Abstract

Developing clustering algorithms for energy optimization, for the Internet of Things (IoT) applications based mainly on wireless sensors networks services, is a major research challenge. The large amount of data generated by this type of networks, requires a huge quantity of energy and a large number of Internet connections to support communication between different devices via the Internet. Therefore, clustering protocols for data routing in IoT, also referred to as hierarchical routing protocols, need to take into account energy efficiency, the number of individual connections and scalability. In this paper, we propose a novel distributed clustering protocol based on the objects' performances aggregation for hierarchical communications in IoT applications focused on Wireless Sensors Networks (WSNs) for the self election of Cluster Heads (CHs) and the clusters construction. Each node calculates a threshold $T(i)$ considered as a competition timer to become a CH. A node can be a CH only once among a fixed number of rounds. An optimum number of clusters is computed based on the network parameters. Simulation results show that our proposed solution is better with energy management performance parameters compared to the Low Energy Adaptive Clustering Hierarchy (LEACH) protocol and two other developed energy-efficient protocols namely LEACH-MAC and TB-LEACH.

Keywords Internet of Things · Wireless Sensors Networks · Clustering · Data communication · Energy efficiency · Optimum number of clusters

1 Introduction and motivation

1.1 Introduction

Cities, homes and industries are becoming increasingly intelligent thanks to the Internet of Things (IoT). IoT is part of the Internet of the future that contains billions of intelligent interconnected "things" [1]. In 2010, the number of objects connected to the Internet surpassed the human population on the planet [2]. The IoT has grown with the emergence of Radio-Frequency Identification (RFID), smart sensors, communication technologies and Internet protocols. These developments allowed direct collaborations between intelligent sensors without human intervention by exploiting technologies such as ubiquitous and pervasive computing, embedded devices, communication technologies, sensor networks and Internet protocols [3]. Wireless Sensors Networks (WSNs) are ubiquitous and a taxonomy of their applications was constructed in [4] where the authors have summarized the specific requirements of each described application domain. They have classified these applications into several categories namely, healthcare, public safety and military, environment and agriculture, industry and transportation systems. Some of these applications can operate underground or underwater. In this technological age, one cannot overlook the extraordinary development of the Internet of Flying Things (IoFT), in particular Unmanned Aerial Vehicles (UAVs), commonly known as drones, supported by wireless communications and networking [5]. Sensor nodes and multiple

✉ Foudil Mir
foudil.mir@univ-bejaia.dz

Farid Meziane
F.Meziane@derby.ac.uk

1 Laboratoire d'Informatique Médicale (LIMED), Faculté des Sciences Exactes, Université de Bejaia, 06000 Bejaia, Algeria

2 Data Science Research Centre, University of Derby, Derby, UK

devices dedicated for some of these applications cannot be equipped with a large power source but use batteries with limited capacity. This exposes the devices to failure when these batteries are depleted, which directly affects the lifetime of the network resulting in the end of the services provided or requested. In this paper, we focus on WSN-based IoT applications, where the energy resource plays a crucial role in the lifetime of the network and the continuity and availability of the services provided or requested via the Internet. In this case, the network is made up of sensor nodes connected to the Internet and are able to detect, process and communicate or receive data. This architecture requires excessive energy consumption and individual Internet connections. Scalability is not guaranteed due to the large number of individual Internet connections of each sensor node. To overcome these limitations, it is possible to aggregate data at specific sensor nodes to eliminate the redundancies caused by collecting large amounts of data, thus reducing the number of direct transmissions to the Internet which is considered to be very energy-intensive. Data aggregation is an energy-efficient data reduction mechanism [4] that mitigates the scalability problem in IoT networks by ensuring that a representative node in each cluster connects to the Internet on behalf of the other nodes [6]. To support data aggregation in a WSN-based IoT network, sensor nodes can be divided into a number of small groups called clusters. Each cluster has a node coordinator, named the Cluster Head (CH), and a number of normal nodes (N) or members. This organization is called clustering. Each CH sends its data to a particular node called the Base Station (BS), which acts as a router to and from the Internet. The data is collected, compressed and transmitted directly to the BS in a single hop or if necessary using other sensors in multi-hops.

Hence, much research is focused on proposing clustering protocols for data routing in IoT applications based on WSNs [7–9], also called hierarchical routing protocols [10].

1.2 Motivation

The LEACH protocol (see Sect. 3.1) [11, 12] is one of the most popular protocols. Despite its success, it suffers from many disadvantages and these include (i) no criteria such as residual energy or distance to the Base Station BS are defined for the election of a node as a CH; (ii) nodes with very low energy can be CHs, which causes them to fail sometimes without completing the current round; (iii) several nodes that are very far from the BS can also be CHs, which can lead to their failures in the phase of direct transmission to the BS; (iv) the election of CHs is strictly probabilistic; (v) the number of CHs is unpredictable hence, the desired optimum number of CHs is not

achieved; (vi) a very significant number of rounds are without CHs; (vii) rounds with a very large number of CHs which makes a large number of nodes will be in direct communication with the BS; (viii) two CHs can be very close and there is no minimum distance between them; (ix) one CH announces itself by sending a message for the whole network, which further increases its power consumption.

As the issue of energy consumption is very important for IoT, many researchers have attempted to develop protocols that can minimise energy consumption and provide better approaches for energy management. Our work is inline with current research approaches for improving such protocols particularly in the use of clusters (see for example [13] for a state of the art approaches using clusters). Hence, the research question is: Can we improve on the current protocols on energy consumption and management using new approaches? More precisely, we are attempting to answer the following research sub-questions: (i) How to elect a set of nodes as CHs? (ii) which approach is better, a distributed or centralized clustering approach? (iii) What are the criteria of a node that decide its election as a CH? (iv) Are these criteria local or global? (v) How do you calculate the radius over which a CH sends a message to solicit membership from normal nodes? (vi) How does a normal node choose its CH in case it receives several membership request messages? (vii) For how many consecutive rounds can a node be elected as CH? (viii) How is intra-cluster and inter-cluster data communication ensured? The aim of our work is to design an energy-efficient, distributed and balanced clustering protocol for routing in IoT applications. We propose the DCOPA protocol which is a Distributed Clustering Based on the Objects' Performances Aggregation for Hierarchical Communications in IoT Applications.

1.3 Research contributions

Our main contribution is the development of the DCOPA protocol that outperforms the state of the art energy management protocols. The superiority of our approach is due to the use of Multiple-Criteria Decision-Making (MCDM) [14] for the election of CHs. It focuses on the modeling and formalization of the CHs election process as a MCDM problem based on the weighted sum of the considered criteria, which are the residual energy and the distance to the BS by associating them to a weight that depends on the importance of the criterion in the clustering process. The aggregation value is then considered as a timer enabling a node to engage in a competition with other nodes of the network for the role of CH. To guarantee a better distribution of CHs through the network, a minimum distance

that must separate two *CHs* is calculated according to the required optimum number of *CHs*.

2 Related works

The most important and crucial step in a clustering algorithm for routing is the designation of the optimum number of *CHs* and the *CHs* nodes needed for each round. These are needed to balance the number of nodes per cluster as well as optimizing and balancing power consumption for a better network lifetime, controlling overhead, optimizing the number of Internet connections and ensuring scalability. The authors in [15] have made an exhaustive study of existing clustering techniques in WSNs. They presented the properties of a network based on these techniques and their objectives such as energy consumption, load balancing, fault tolerance, quality of service and scalability. The clustering techniques proposed in the literature are classified in [16] according to their required objective and the overall architectural and operational model of the network, including the desired number and characteristics of the resulting clusters. Data transmission based on clustering is considered in [17] as a typical hierarchical routing. Clustering makes it possible to considerably and significantly reduce energy consumption in sensor network applications, mainly in data communication, which remains to the present day one of the most important areas of research that attracts and excites the research community working on WSNs issues, particularly in data routing. The challenges of this type of architectures are numerous and include (i) the approach to choose the *CHs* that can be distributed [11] or centralized [12]; (ii) defining the algorithm to rotate the role of the *CH* between the nodes of the network respecting a set of calculated thresholds and considered criteria [18]; (iii) the choice of *CHs* that can be dynamic [19] or static for a certain number of rounds; (iv) the optimal number of *CHs* [20] and (v) the radius and size of the clusters [21]. All these challenges and steps take into consideration the efficient and balanced management of the energy resources to increase the network lifetime. Several other challenges in WSNs such as security, quality of service and coverage are also attracting a lot of interest from the research community. Research on the use of clustering for energy efficiency in sensor networks is continuously growing. In this section, we will discuss some energy efficient clustering protocols proposed for WSNs that can also be applied to IoT environments using a variety of devices.

The field of clustering in WSNs was investigated and explored in depth in the last two decades. For this reason, the focus of this state of the art review is on the LEACH protocol[11, 12] and several of its variants. After two

decades from its development, LEACH is still of interest to researchers in the field. It is one of the first and most popular distributed hierarchical routing protocols which runs in two phases. The first one is dedicated to the probabilistic election of a certain percentage *p* of nodes for the role of *CHs*. In the second phase, the nodes choose a nearest *CH* to build the clusters. A node can play the role of *CH* only once in *(1/p)* rounds. The data sent by the nodes to their *CHs* will be aggregated and communicated to the *BS* directly. Furthermore, the same authors developed a method to compute the optimum number of clusters in the network, called K_{opt}, based on the size of the network, the position of the *BS*, the constants of the radio model and the number of the deployed nodes [12]. Several surveys on the improvements of LEACH have been published [22–26]. Interest in LEACH protocol is due to its advantages such as the probabilistic and distributed aspects, the simplicity of the algorithm and the proposed formula, a very interesting condition imposed for the rotation of the role of *CH* between the nodes of the network and its demonstrated energy efficiency.

There are many protocols that have been developed with the attempt to modify the original LEACH protocol to improve its performance. Heinzelman et al. [12] proposed a centralized control algorithm called Low Energy Adaptive Clustering Hierarchy Centralized (LEACH-C). The election of *CHs* and cluster formation is carried out by the *BS* after each node sends its position and energy level at the beginning of each round. Nodes with an energy level higher than the average energy level of the other nodes of the network can be a *CH* in the current round. To identify clusters, the *BS* uses the simulated annealing algorithm [27] to solve a NP-Complete problem to find the optimum number of clusters [28]. After selecting the *CHs* and associated nodes, the *BS* broadcasts a message to the remaining nodes over the network. The last step is similar to the LEACH protocol.

MELEACH-L [29] is proposed as an improvement of the ME-LEACH protocol for More Energy Efficient-LEACH [30] which is an extension of LEACH. Each round is composed of four phases. The first phase is the selection of *CHs* using a *T(i)* timer. In the second phase, a backbone tree is built with Energy-Aware Virtual backbone Tree (EAVT) [31] by selecting some nodes that are not *CHs*. During the third phase, each *CH* chooses the closest EAVT node to be used as a relay to reach the *BS*. Non *CH* nodes choose their *CHs*, called leaders, and a tree is built in intracluster for data communication. During the last phase, each non *CH* node sends its data to the *CH* which aggregates them to be sent by the tree backbone through its previously chosen relay Node.

Junping et al. [32] presented the Time Based LEACH (TB-LEACH) distributed protocol to modify the *CHs* set selection process. A *CH* is selected based on a random time interval. The node with the shortest time interval becomes *CH*. The authors have defined the desired number of *CHs*. A node generates a random number considered as a time at the beginning of each round. The node decreases its time, if it expires before the number of *CHs* in the network is reached and declares itself *CH*. Otherwise, the node leaves the competition. Once the *CHs* have been elected, the remaining of the process is the same as in LEACH.

In Advanced-LEACH (ALEACH) [33], the authors proposed a new threshold $T(n)$ calculated in the set up phase based on the sum of current state probability (CSp) and general probability (Gp). $T(n)$ is a function of the residual energy of the node, the initial energy of the network, the number of *CH* that is planned in a round, the number of the current round and the total number of nodes in the network. The phase steady state is identical to LEACH. Differently from the dynamic clustering technique, Hong et al. in [34] proposed Threshold-LEACH (T-LEACH) where the election of *CHs* becomes hybrid. The elected *CHs* will remain *CHs* in the following rounds as long as their energy are not below a given threshold. The election process is activated again otherwise. The focus of LEACH-Balanced (LEACH-B) [35] is the balancing of clusters. An additional stage of *CHs* competition is added in the set up phase so as to keep the number of *CHs* constant. The *CHs* randomly elected in the set up step of LEACH, broadcast a message for the whole network containing their residual energies. If the number of *CHs* is greater than $(N * P)$, where N is the number of nodes and P is the desired percentage of *CHs*, some *CHs* according to their residual energies, will abort the role of *CH*. However, if the number of *CHs* is not greater than $(N * P)$, some normal nodes will become *CHs* according to a time calculated to maintain the number $(N * P)$ of *CHs* for each round.

The protocol Improved-LEACH (I-LEACH) [36], improves the formula $T(n)$ for choosing *CHs* in LEACH. In addition to the desired percentage of clusters and round number, new parameters such as the distance to the *BS*, the number of neighbors and the residual energy of the node are also taken into account. Chen et al. [37] proposed LEACH-G to overcome the disadvantages of LEACH, namely the number of *CHs* which is not ensured, the optimum number K of *CHs* is not guaranteed and the unbalanced distribution of *CHs* in the network. LEACH-G combines the centralized and the distributed approaches in order to select the *CHs* and their nodes in each round. A new formula to calculate the optimum number of clusters has been defined, which is a modification of the one given in [12]. The goal of the contribution by Salim et al. [38],

called Intra-Balanced LEACH (IB-LEACH) is to minimize the costs of intra-cluster communication. Three phases constitute a round. After the set up phase which is similar to that of LEACH, a new phase called pre-steady state where nodes are divided into three classes, *CHs*, aggregators and normal nodes is introduced. The *CH* manages its cluster and chooses the aggregator nodes. A message containing the Time Division Multiple Access (TDMA) schedule and the list of aggregators will be sent to the cluster nodes. In the steady state phase, which is divided into frames, the nodes send their data to the aggregators in their reserved slots, which in turn aggregate the data and forward it to the *BS*.

The authors of Vice Cluster LEACH (V-LEACH) [39] reconsider the weakness of LEACH which is the absence of the energy parameter in the selection of *CHs*. V-LEACH introduces a new type of node called vice *CH*. The stage of clustering is the same as LEACH. The node with the highest energy in the cluster is designated as Vice *CH* to replace the *CH* in case of failure in a round. The steady state phase is similar to LEACH. Behera et al. proposed R-LEACH [9] which consists of two phases, namely set up and steady state phases. In the first round the selection of *CHs* is similar to LEACH. From the second round a new $T(n)$ is introduced. Precisely, $T(n)$ of the nodes that do not belong to the set of nodes that were *CH* during the $(1/p)$ rounds is based on the old LEACH formula, the residual energy of the node, the initial energy of the node and the optimum number of clusters (K_{opt}) introduced in [40]. The steady state phase is similar to that of LEACH.

Nihar et al. [20], have reviewed some assumptions of LEACH mainly on the optimum number of clusters. They have made an important observation namely that the parameter Energy dissipation by the Receiver circuitry (EelecRx) has a crucial role in the calculation of the optimum number of clusters compared to the parameter Energy dissipation by the Transmitter circuitry (EelecTx) which has no impact. After a successful proof of concept, the authors have modified the K_{opt} formula given in [12] and defined a new one which integrates the fraction (EelecRx /Emp) (Emp: Energy required for multi path propagation) neglected in [12] for the calculation of the optimum number of the desired clusters. Zhen et al. [41], focused their proposal, the Modified LEACH (MLEACH), on the optimum number of *CHs* in each round as well as the integration of energy as a weighting factor in the election of *CHs*. A new $T(n)$ is provided containing the residual energy of the node, the energy consumed in the previous round by a node, the average energy consumed in the previous round by the whole network and the total residual energy of the network. The authors in [42] proposed ALEACH-Plus, an improvement in the selection

process of *CHs* which considers the node energy, the distance from the *BS* and the number of neighbors in the formula $T(n)$, unlike LEACH which is probabilistic and does not take into account the node criteria. In the transmission phase, a node that is closer to the *BS* than to its *CH* and its energy is lower than its *CH*, communicates directly with the *BS*. Two elected *CHs* respect a certain distance between them. The authors integrate a new condition for ALEACH-Plus to obtain another contribution which is ALEACH-Plus2. A node becomes *CH* if its residual energy is higher than the average of the energies of the nodes of the network.

Multi-Energy Threshold LEACH (MET-LEACH) is presented by Anika et al. [43]. $T(n)$ is computed in the same way as LEACH. The MET-LEACH process is in two phases, set up and steady state. The authors introduced four energy thresholds, $E_{Th1} = 75\%$, $E_{Th2} = 50\%$, $E_{Th3} = 25\%$ and $E_{Th4} = 1\%$ of the initial energy. A node participates in the *CH* selection process if its residual energy is higher than the threshold considered in the current round. Initially E_{Th1} is used. Once most nodes have an energy lower than E_{Th1}, E_{Th2} is applied. Similarly E_{Th3} is considered once most nodes have an energy less than E_{Th2}. The E_{Th4} threshold is applied if the majority of the nodes have an energy less than E_{Th3}. The technique proposed in [44] extends the LEACH protocol to the LEACH-V protocol by including the Vector Quantization (VQ) technique for intra-cluster cooperative communication considered as a powerful quantization technique that is used for image compression. With VQ, the minimum distance is calculated between multiple *CHs* to create a shortest path that saves energy. The authors in [45], proposed the Low energy adaptive clustering hierarchy with VQ and Dijkstra's algorithm (V-LEACH-VD) which is a new improvement of the LEACH protocol with the integration of the VQ technique and Dijkstra's algorithm. In this improvement, LEACH clusters are replaced by voronoi regions. VQ is used for intra-cluster cooperative communication and the Dijkstra algorithm for designing the minimum energy cost path. Maximum LEACH (MaxLEACH) is proposed in [46] to increase the lifetime of the network as well as the integration of residual energy in the *CHs* selection process. The first round is similar to LEACH with the inclusion of the node energy in the comparison threshold. After the first round, the node with the highest energy will be elected as *CH* for the next round. Residual Energy based Cluster Head Selection in (RCH-LEACH), is proposed in [47], at the beginning of each round, each node sends its position and residual energy to the *BS* which will proceed to the selection of *CHs* using the process which is similar to the LEACH protocol with the addition of another condition which is the satisfaction of a calculated energy threshold. The set of *CHs* obtained is subjected to another check, if the number of *CHs* obtained is higher than the optimal number of clusters *K*, then the *BS* chooses *K CHs* that have the most energy. If the opposite is true, all *CHs* are retained. The *BS* broadcasts the list of *CHs*, the nodes choose their *CHs* in a similar way to LEACH.

Within the context of the IoT and Industry 4.0, sensors play an important role as they manage large scale infrastructure and require communication protocols with high resilience. Furthermore, they require strong message exchange with very few misses and capable of achieving scalability and elasticity. Esposito et al. [48] developed a system that exhibits these properties. They have developed a system for resilient event based communications among the sensors, without requiring a pre-deployed brokering infrastructure supporting the adopted protocol. Their system required low energy consumption as they have introduced a proper strategy that was able to reduce the number of exchanged messages using a game theoretic formulation. Furthermore, a simulation of their system has successfully delivered notifications at a lower cost than the available state-of-the-art solutions.

More recent works have adopted clustering based approaches. For example, to address two main drawbacks of the LEACH protocol, namely, the random selection of *CHs* and the unequal clustering. The authors in [49] proposed an improvement based on node distribution density and allocating remaining nodes. They introduced the peak density clustering algorithm called MKN-DPC which is based on mutual *K* nearest neighbors. First the local characteristics of the network node distribution are considered to dynamically calculate the local density of each data point. Then, the automatic selection of the initial *CH* and the allocation of the remaining nodes is done by combining the two parameters of density and signal strength. Finally, to avoid excessive cluster partitioning, a cluster merging strategy was applied.

Mtopi et al. [50] proposed a multiHop constant-time complexity clustering algorithm (MultiHopFast) for IoT networks to address the load balancing and scalability challenges. The proposed MultiHopFast algorithm reduces the computing burden from IoT nodes with smart load balancing to ensure IoT network scalability. The algorithm addresses the network load, scalability, and time efficiency challenges. Using neighbourhood heuristics, the MultiHopFast algorithm builds appropriate size of clusters with participating IoT nodes. Each cluster is associated with a *CH* (or a coordinator). The MultiHopFast algorithm selects *CH* for each cluster using a probabilistic model.

For a comprehensive review of the approaches using clustering, Hosseinzadeh et al. [13] have carried out a high quality study of clustering algorithms in IoT especially in the context of smart cities to draw up a technical taxonomy

for the categorization of clustering in IoT. They have reviewed the papers published between 2017 and 2021.

Following on the conclusion drawn by [13] on the great interest of researchers on clustering for energy efficiency, mainly on scalability, robustness, and load balancing, we have proposed a novel approach that uses Multiple-Criteria Decision-Making for the election of *CHs*. Furthermore, to our knowledge this is the first use of a competition between nodes to select the *CHs*. The simulation results show that our protocol outperforms the state of the art protocols.

3 The LEACH energy model

3.1 LEACH protocol architecture

LEACH is a distributed dynamic and probabilistic clustering protocol for single-hop routing. Nodes do not need global information to pronounce themselves as *CHs*. Nodes can decide to elect themselves as *CHs* by comparing a value called $T(i)$ given in formula (1), calculated from the percentage P of the desired *CHs* and the number of the current round with a generated random number between [0,1]. LEACH consists of two phases namely the set up and steady phases. In the set up phase, each node $N(i)$ generates a random number between [0,1], and calculates the value of $T(i)$ called the threshold. P is the percentage of *CHs* desired to be calculated based on the optimum number of *CHs*, r is the number of the current round. G is the set of nodes which were not *CHs* during the previous $(1/P)$ rounds.

$$T(i) = \begin{cases} \dfrac{P}{1 - P * (r \bmod \frac{1}{p})} & \text{if } i \in G \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

If the random number is less than $T(i)$ then the node $N(i)$ becomes *CH* and sends an ADV-CH message for the whole network. For the steady state phase, the nodes receiving ADV-CH will choose the nearest *CH* and will send a JOIN-CH message. The *CHs* broadcast the Time Division Multiple Access (TDMA) schedule for intra-cluster communication to avoid collisions. The *CHs* receive the data from all the nodes of the cluster and after aggregating them, they are sent directly to the *BS*.

3.2 Energy consumption model

The energy dissipation model used in the simulations is that of Heinzelman et al. [12]. It counts the transmission energy according to the size l of the transmitted message and the communication distance d between the sender and the receiver. The reception energy is counted according to the size of the received message l. The data aggregation energy at the *CHs* level is also taken into consideration in this model. During transmission, the energy consumed is defined by $E_{Tx}(l, d)$ in formulas (2–3). Depending on the distance d, two channel models and two power control settings are used as follows:

- If $(d \leq d_0)$, d_0 is given in equation (4), the channel of the free space model (d^2 power loss) and the free space power amplifier E_{fs} are used.
- Otherwise $(d > d_0)$, the channel of the multipath fading model (d^4 power loss) and the multipath power amplifier E_{mp} are used.

$$E_{Tx}(l, d) = E_{Tx-elec}(l) + E_{Tx-amp}(l, d) \quad (2)$$

$$E_{Tx}(l, d) = \begin{cases} E_{elec} * l + E_{fs} * l * d^2 & si\, d \leq d_0 \\ E_{elec} * l + E_{mp} * l * d^4 & si\ d \leq d_0. \end{cases} \quad (3)$$

$$d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}} \quad (4)$$

E_{elec}, E_{mp}, E_{fs} are defined in Table 1.

The reception energy $E_{Rx}(l, d)$ of l bits is defined in formula (5) and is not sensitive to distance.

$$E_{Rx}(l, d) = E_{Rx-elec}(l) \quad (5)$$

The *CH* nodes aggregate the data (signals) with an energy called E_{DA}(See Table 1) to send them to the *BS*.

3.3 Optimum number of cluster

The optimum number of clusters K_{opt} is calculated in [12], and given in formula (11), depends on the energy consumption of a *CH* node, a non-CH node and the energy consumed by a cluster. Network and radio parameters are also considered. In this section, we only give the essential steps to calculate it using formula (6–11), the complete formal demonstration is given in [12] and the variables are:

- M*M is the area of the monitoring zone.
- K is the number of clusters.
- d_{toCH} is the average of the distances from the *CH* to the *BS*.
- N is the number of nodes.
- E_{CH} is the energy consumed by a *CH* node.
- E_{Non-CH} is the energy consumed by a non-CH node (normal node).
- $E_{Cluster}$ is the energy consumed by a cluster.

$$E_{CH} = lE_{elec}\left(\frac{N}{K} - 1\right) + lE_{DA}\frac{N}{K} + lE_{elec}d_{toBS}^4 \quad (6)$$

$$E_{Non-CH} = lE_{elec} + lE_{fs}d_{toCH}^2 \quad (7)$$

Table 1 Energy model parameters

Parametres	Values	Description
E_{elec}	50 nJ /bit	Required energy to run electronic circuit
E_{fs}	0.0013 pJ /bit/ m4	Free space propagation
E_{mp}	10 pJ /bit/ m2	Multi path propagation
E_{DA}	5 nJ /bit/ signal	Required energy for data aggregation

$$E[d_{toCH}^2] = \frac{1}{2\Pi}\frac{M^2}{K^2} \tag{8}$$

$$E_{Non-CH} = lE_{elec} + lE_{fs}\frac{1}{2\Pi}\frac{M^2}{K^2} \tag{9}$$

$$E_{Cluster} = l(NE_{elec} + NE_{DA} + KE_{mp}d_{toBS}^4$$
$$+ NE_{elec} + NE_{fs}\frac{1}{2\Pi}\frac{M^2}{K^2}) \tag{10}$$

$$K_{opt} = \frac{\sqrt{N}}{\sqrt{2\Pi}}\sqrt{\frac{E_{fs}}{E_{mp}}}\frac{M}{d_{toCH}^2} \tag{11}$$

Another formula (13) for the optimum number of clusters K_{New} is defined by Nihar et al. [20] and is calculated using the same parameters by demonstrating that the optimum number of clusters depends on C as given in (12).

$$C = \sqrt{d_{toCH}^2 - \frac{E_{elec}}{E_{mp}}} \tag{12}$$

For this reason a K_{New} is given in formula (13).

$$K_{New} = \frac{\sqrt{N}}{\sqrt{2\Pi}}\sqrt{\frac{E_{fs}}{E_{mp}}}\frac{M}{\sqrt{d_{toCH}^2 - \frac{E_{elec}}{E_{mp}}}} \tag{13}$$

3.4 The concept of the network lifetime

Lifetime is an essential aspect for evaluating the performance of WSNs. The number of alive nodes, coverage and connectivity are the parameters of the lifetime. The Quality of Service (QoS) metrics can also be brought back to the lifetime context [51]. Several definitions and an exhaustive study of the WSNs metrics lifetime are given in [51]. In our simulations we use the definitions related to the number (percentage) of alive (dead) nodes in the network as a function of the number of rounds.

4 The DCOPA algorithm

4.1 The mathematical and theoretical foundation of DCOPA

The DCOPA protocol is modelled using the mathematical method of Multiple-criteria decision-making (MCDM), more precisely with a weighted sum-based multiple criteria method which is a method of a priori aggregation of criteria into a single criterion which we then consider as a timer allowing a node to engage in a competition with other nodes in the network for the role of *CH*. We were able to bring the *CHs* election problem in an IoT network and especially in a WSN to a multiple criteria decision-making problem. The aggregated criteria in our case are the residual energy of the node and the distance to the *BS*. Each node aggregates these two criteria by associating them with a weight that depends on the interest in each criterion and its degree of implication in the expected objective which is the minimization of energy consumption in clustering for data communication.

4.2 Problem formulation

We make the following considerations:

- The considered criteria are : Er_i and d_{itoBS} (see Table 2). Each device has its residual Er_i energy and distance from the *BS* d_{itoBS}.

Table 2 Variables of T(i)

Variables	Description
d_{itoBS}	The distance between the node N(i) and the *BS*.
$d_{MaxtoBS}$	The maximum distance to the *BS*.
$d_{MintoBS}$	The minimum distance to *BS*.
E_{Max}	The initial energy of the node.
Er_i	The residual energy of node N(i).
α	The weight of the energy criterion.
β	The weight of the distance criterion.
τ	The time of the self-election period of *CHs*.
δ	A small positive real number.

- The Weight Vector (α, β).
- Maximise the Er_i criterion, so it is equivalent to minimising $(E_{Max} - Er_i)$.
- Minimise the d_{itoBS} criterion.

With the minimisation of these two criteria, we aim to minimise $T(i)$ in formula 14, which aggregates the two criteria, considered as a unique criterion which is a temporal value. Each node with a $T(i)$ value close to zero will have a higher chance of being declared as CH. This means that the devices closest to the BS with high energy have the highest chance of being declared as CHs.

4.3 Description

DCOPA is a distributed hierarchical routing protocol based on the auto election of CHs after a competition between network nodes. The competition focuses on the calculation of a timer $T(i)$ by each node according to its residual energy and its distance to the BS. The nodes will not need global data. $T(i) \in\,]0, \tau - \delta]$, is considered as a time strictly inferior to the duration of the dedicated period of election of CHs which is τ. $\delta \in\,]0, 1[$ is a very small time to avoid that a node declares itself CH at the end of the period τ. At the beginning of this period, each node decrements its $T(i)$. If $T(i)$ expires the node announces itself as CH and eliminates all candidatures for the role of CH on a calculated radius called RC (defined in Sect. 4.5). A node which receives the first announcement message from a CH before the expiration of its $T(i)$ becomes a normal node and cancels its application for the role of CH. A normal node sends a membership message to the nearest CH. DCOPA consists of two phases that make up a round. These are described in the following subsection.

4.4 DCOPA phases

4.4.1 Set up phase (nodes status distinction)

This is the phase of self election of CHs. At the beginning of each round, all nodes calculate a timer $T(i)$ using formula (14), including the residual energy and the distance from the BS. Each criterion can be associated with a weight depending on its degree of interest in a given application, which means that a weighted sum is minimized to have a minimum time to win the competition and reach the role of CH. The criteria and weights are normalized so that formula (18) is verified. Each node starts to decrement its $T(i)$ ($T(i) \leq \tau - \delta$). τ is the time of the self-election period of CHs. Simply, it means that a node has to have a competition time smaller than τ and δ is a small amount (time) that allows us to satisfy this condition.

If a node receives the first Control Message for Requesting to Join a Cluster Head (CtrlMsgRqJoinCH) before reaching zero, it cancels its timer and leaves the competition waiting for more CtrlMsgRqJoinCH until the end of the period reserved for the election of CHs (τ). A node can be in the RC of multiple CHs. If $T(i)$ expires, the node broadcasts the message CtrlMsgRqJoinCH to announce itself CH using the CSMA MAC protocol (Carrier-Sense Multiple Access Media Access Control). CtrlMsgRqJoinCH is sent on an RC calculated based on the optimum number of clusters K in the network.

The variables used in the definition of $T(i)$ are described in Table 2.

$$T(i) = \begin{cases} (\alpha E_i + \beta D_i)(\tau - \delta) & if\ i \in G \\ \tau - \delta & otherwise. \end{cases} \quad (14)$$

G is the set of nodes which were not CHs during the previous $(1/P)$ rounds, $P = \frac{K}{Nbr_{init}}$. Nbr_{init} is the initial number of nodes.

$$\alpha + \beta = 1 \quad (15)$$

E_i given in formula 16 and D_i given in formula 17, are defined as follow after the normalization process.

$$E_i = \left(\frac{E_{Max} - Er_i}{E_{Max}}\right) \quad (16)$$

$$D_i = \left(\frac{d_{itoBS} - d_{MintoBS}}{d_{MaxtoBS} - d_{MintoBS}}\right) \quad (17)$$

$$0 \leq \alpha \frac{E_{Max} - Er_i}{E_{Max}} + \beta \frac{d_{itoBS} - d_{MintoBS}}{d_{MaxtoBS} - d_{MintoBS}} < 1 \quad (18)$$

Formula (18) is checked as follows: the first objective is to minimize the distance from a node i to the BS (d_{itoBS}).

$$d_{MintoBS} \leq d_{itoBS} \leq d_{MaxtoBS} \quad (19)$$

$$0 \leq d_{itoBS} - d_{MintoBS} \leq d_{MaxtoBS} - d_{MintoBS} \quad (20)$$

after normalization we get:

$$0 \leq \frac{d_{itoBS} - d_{MintoBS}}{d_{MaxtoBS} - d_{MintoBS}} \leq 1 \quad (21)$$

Formula (21) is checked with the two formulas (19) and (20).

The second objective is to minimize the difference between the initial energy of a node i and its residual energy ($E_{Max} - Er_i$). Dead nodes are not taken into account ($Er_i > 0$).

$$0 \leq E_{Max} - Er_i < E_{Max} \quad (22)$$

After normalization we get:

$$0 \leq \frac{E_{Max} - Er_i}{E_{Max}} < 1 \quad (23)$$

Formula (23) is checked with formula (22).

Using formulas (15), (21) and (23), formula (18) is verified.

Example

if $Er_i \rightarrow E_{Max}$ and $d_{itoBS} \rightarrow d_{MintoBS} \Rightarrow N(i)$ has a high chance of being *CH*.

if Si $Er_i \rightarrow 0$ and $d_{itoBS} \rightarrow d_{MaxtoBS} \Rightarrow N(i)$ has a small chance of being *CH*.

Algorithm 1, explains the steps of the set up phase. The parameters $d_{MintoBS}$, $d_{MaxtoBS}$, RC are calculated once at the beginning of the first round. The variables used are presented in the following.

$N(i)$: node number i.

$N(i).Status$: node's status (*CH* or *N*).

$N(i).ClstrMbr$: equal to 1 if the node is part of a cluster or 0 otherwise.

$N(i).ListCH$: list (set) of *CHs* requesting a normal node.

period for clustering, where a normal node sends its Acknowledgment Control Message (CtrlMsgAckNd) to the nearest *CH*. The *CHs* listen for possible Acks during this first period. At the beginning of the second period, which is the Time for Waiting Data Messages of the Nodes (TimeWtDataMsgNds), each *CH* broadcasts its TDMA calendars for the cluster nodes in order to send their Data Message (DataMsgNd) at the beginning of this period. This period is reserved for intra-cluster data routing. Furthermore, the *CH* nodes stay tuned to receive data messages from nodes that have confirmed their membership to the cluster. A *CHs* aggregates the data (DataMsgAg) and sends it directly to the *BS* using the CSMA MAC protocol during the third period called Time for Sending Cluster Head Data Message to the *BS* (TimeSDataCHtoBS), which is dedicated to the data aggregation and communication with the *BS*.

Algorithm 1 : Set up Phase

Input
1:   $N(i).Status =' N'$,
2:   $N(i).ClstrMbr = 0$,
3:   $N(i).ListCH = \emptyset$,

Output
4:   $N(i).Status$,
5:   $N(i).ListCH$,
6: Compute $T(i)$,
7: **while** τ not expired **do**
8:      **if** $((T(i)$ expired$)$ and $(N(i).ClstrMbr == 0))$ **then**
9:          $N(i).Status =' CH'$,
10:         $N(i).ClstrMbr = 1$,
11:         Broadcast a CtrlMsgRqJoinCH on RC,
12:     **end if**
13:     **if** (Reception of CtrlMsgRqJoinCH from $N(j)$) $\wedge$ $(N(i).ClstrMbr == 0)$ **then**
14:         $N(i).ClstrMbr = 1$,
15:         $N(i).ListCH = N(i).ListCH \cup N(j)$,
16:         Cancel $T(i)$,
17:     **end if**
18:     **if** (Reception of CtrlMsgRqJoinCH from $N(j)$) $\wedge$ $(N(i).ClstrMbr = 1) \wedge (N(i).Status =' N')$ **then**
19:         $N(i).ListCH = N(i).ListCH \cup N(j)$,
20:     **end if**
21: **end while**

4.4.2 Steady state phase (clusters formation and routing)

Once the τ period reserved for the set up phase has expired, each node $N(i)$ will have its status. This phase is divided into three periods. The Time for Waiting Acknowledgments (Acks) of Nodes (TimeWtAckNds) is a reserved

Algorithm 2, explains the steps of the steady state phase. The variables used are presented in the following.

$N(i).NbrNdCH$: number of nodes joined to a given *CH*.

$N(i).ListNdCH$: list (set) of nodes joined to a given *CH*.

$N(i).NumCH$: *CH* number of a normal node.

Algorithm 2 : Steady state Phase

Input

1:    $N(i).NbrNdCH = 0$,
2:    $N(i).ListNdCH = \emptyset$,
3:    $N(i).NumCH = 0$,
4:    $N(i).Status$,

Output

5:    $N(i).NbrNdCH$,
6:    $N(i).ListNdCH$,
7: **if** $(N(i).Status ==' CH')$ **then**
8:    **while** (TimeWtAckMsgNds not expered) **do**
9:      **if** (Reception of CtrlMsgAckNd from $N(j)$) **then**
10:        $N(i).NbrNdCH = N(i).NbrNdCH + 1$,
11:        $N(i).ListNdCH = N(i).ListNdCH \cup N(j)$,
12:      **end if**
13:    **end while**
14:    Send TDMA Schdule for $N(i).ListNdCH$,
15: **else**
16:    $N(i).NumCH = j /\{N(j) \in N(i).ListCH$ is the nearest CH from $N(i)\}$,
17:    Send CtrlMsgAckNd for $N(N(i).NumCH)$,
18:    receive TDMA schedule from $N(N(i).NumCH)$,
19: **end if**
20: **if** $(N(i).Statut ==' CH')$ **then**
21:    **while** (TimeWtDataMsgNds not expered) **do**
22:      receive DataMsgNd (All nodes of $N(i).ListNdCH$),
23:    **end while**
24:    Data Agregation,
25:    Send MsgDataAg to the BS,
26: **else**
27:    Send DataMsgNd to the Node $N(N(i).NumCH)$,
28: **end if**

4.5 Calculate the radius of the clustering (RC)

Determining a *CH*'s RC is an important aspect of our contribution to reduce the announcement distance of a *CH*, find the distance over which all other candidatures for the *CH* role will be eliminated and divide the network into balanced clusters. For this effect, we will theoretically assume the shape of a cluster as having the shape of a circle. The radius R of a cluster in this case, considering K as the optimum number of clusters and M*Mm^2 as the deployment area, is given as $R = \frac{M}{\sqrt{\Pi K}}$, in other words R is the maximum distance between a *CH* and the nodes of the cluster.

If we assume the same conditions given in [12], a square of $100m * 100m$, $(x = 50m, y = 175m)$ for the *BS* position and the same radio parameters, we will obtain $K = 5$ using formula (11), in this case, the radius of each circle is $R \simeq 25m$ which makes four circles that do not extend beyond the square and without intersection (they are entirely contained within the square). The distance between two *CHs* is $2R \simeq 50m$. This is not applicable in our case since it is not a question of looking for the number of distinct circles of radius R in the surface of deployment. The circles can overflow the surface of the network, possessing zones of intersections and sparing the fact of having nodes which do not belong to any circle.

A *CH* node cancels the candidacy of other nodes on a radius $R = \frac{M}{\sqrt{\Pi K}}$, one or more nodes can be *CHs* located at the outer limit of the circle with the radius R, in this case, if $RC = R = \frac{M}{\sqrt{\Pi K}}$, the minimum distance between two *CHs* is $R + \epsilon$. We will not have a number of clusters close to K but a large number of *CHs* up to double $(2K)$, because the distance between two *CHs* that must be 2R is not respected. The schematic explanation is shown in the example given in Fig. 1.

In conclusion, to make two *CHs* separated by a minimum distance of 2R, we must have $RC = 2R = \frac{2M}{\sqrt{\Pi K}}$, which means in this case that the minimum distance between two *CHs* is $2R + \epsilon$, we will be able in this situation to have a number of clusters that is close to K as shown in the example given in Fig. 2.
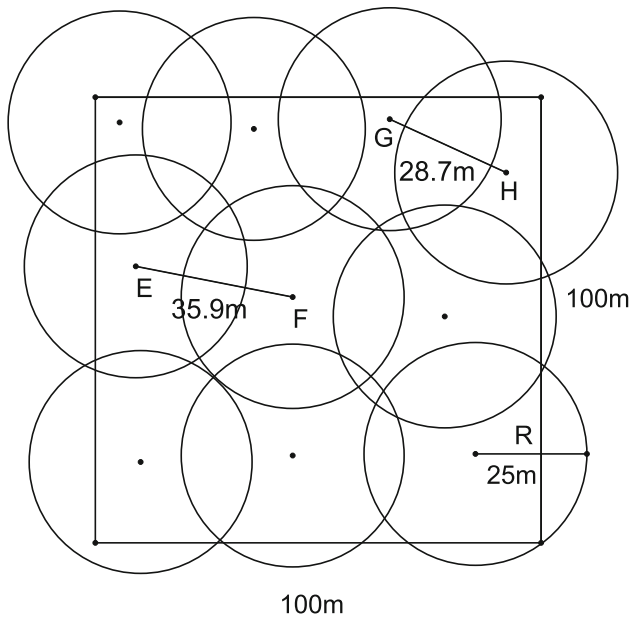
Fig. 1 Example: Number of clusters realized with a radius $R = RC = 25m$ in an area of 100m*100m
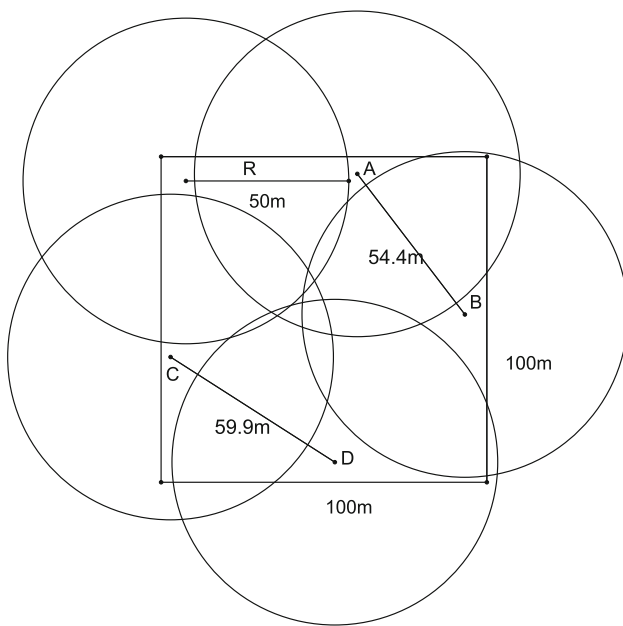
Fig. 2 Example: Number of clusters realized with a radius $R = RC = 50m$ in an area of 100m*100m

5 Simulation and performance analysis

5.1 Objectives

The aim of our simulations is to study the energy behaviour of the nodes and the network as a function of the rounds, which also includes studying the death rate of the nodes as a function of the rounds.

Table 3 Simulation parameters

Parameters	Values	Description
$M * M$	$100^2 \ m^2$	Area network
E_{Max}	0.5 j	Initial energy
$d_{MintoBS}$	75 m	Nearest point to BS
$d_{MaxtoBS}$	183 m	Furthest point to BS
$Sink_x$	50 m	Sink x-axis
$Sink_y$	175 m	Sink y-axis
$MsgCtrl$	25 bytes	Control message length
$DataMsg$	200 bytes	Data message length
K	5 clusters	Optimum clusters number [12]

5.2 States of the deployment environment and nodes

The area of deployment is M*M m^2, with N nodes deployed in a uniform random way. The nodes have the same initial energy. Two types of messages are considered, control messages and data messages. Table 3 summarizes the used values and parameters. The BS is outside the monitoring area. In the energy model used (introduced in 3.2), the relationship between energy and distance is very close, for this reason, we consider in our simulations that the weights are equals for each criteria ($\alpha = \beta = 0.5$). The assumptions used in these simulations are the same as those used in the reviewed papers and common among researchers. The following are the assumptions.

- The BS is unlimited in energy terms.
- The nodes are static.
- The nodes do not have equipment to know their positions.
- The batteries in the nodes are not replaceable.
- The node will fail only if its energy is completely consumed.
- All nodes can change their ranges according to their distance from the receiver(s).

5.3 Results and performance analysis

In the following, we will discuss all the comparisons and simulations performed using MATLAB[1]. In the same direction as all the research work done in the literature on energy optimisation in WSNs or WSN-based IoT applications, we will have to study practically the energy behaviours of the nodes and the network with regard to the variation and balancing of the energy consumption as well as the mortality rate of the nodes. Subsequently, confidence

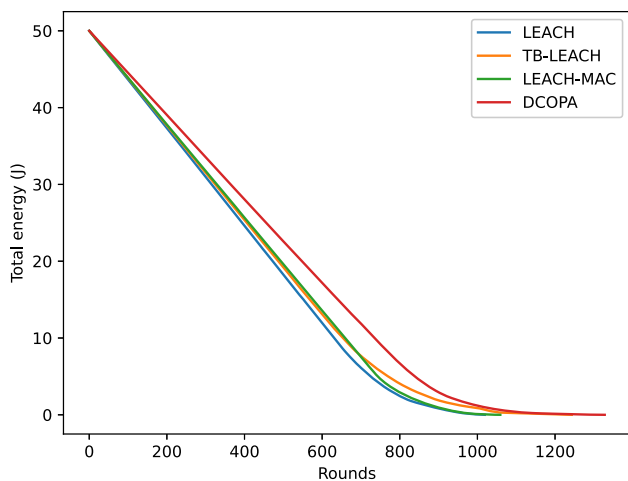---

[1] https://www.mathworks.com

Fig. 3 Total network energy as a function of rounds

intervals (95%) will be calculated for a certain percentage of mortality of the nodes in the network.

5.3.1 Total energy consumption

In this section, we evaluate the total energy consumption of all the nodes according to the rounds, or in other words, we explain it inversely as the total residual energy of all nodes. Figure 3 describes the total energy of the network as a function of rounds.

We notice that our contribution is characterized by a better management of the total energy of the network without drastic decrease compared to other protocols, namely LEACH, TB-LEACH and LEACH-MAC. The enhancement of the CH election process based on the examination of the residual energy of the nodes as well as their distance from the BS as the criteria that define the competition to access the role of CH. The announcement of the CHs is performed over a determined distance, but not for the whole network. The fact that the CHs are distributed throughout the whole monitoring area, due to the required minimum distance between two CHs. All the factors mentioned above have a very positive effect on our results. Table 4 serves as an example of the difference in total network energy between the four protocols considering two rounds. For round 600, we have a total of 17.18j for DCOPA, 13.58j for TB-LEACH, 13.08j for LEACH-MAC and 11.92j for LEACH, which means that DCOPA has a very satisfactory total energy management and saving,

which preserves it for a longer time and thus improves the network's life.

5.3.2 Node's average residual energy

In this comparison and evaluation part, we discuss the variation of the average residual energy per node as a function of the rounds. Figure 4 shows that the average residual energy of the nodes, when running DCOPA, is better compared to other protocols.

The decrease is linear until approximately round 800, after this, we notice a slight stability that lasts until the last node fails. This stability is explained by the mortality rate which is not really outstanding. The peaks of the TB-LEACH protocol between rounds 850 and 1100, are due to the number of living nodes considered as very reduced, which increases the average residual energy compared to the DCOPA protocol which has a higher number of living nodes between the two mentioned rounds. DCOPA's performance in terms of better average residual energy is achieved through moderate energy consumption in all stages of the protocol and by both types of nodes. This moderation is manifested by the better distribution of CHs in the network controlled by an optimized RC that allows nodes not to waste energy by joining the nearest CH with a distance less than or equal to RC and the fact that there is more chance that a node will act as a CH once in $(1/p)$ rounds except if the nodes are isolated or once there is a very small number of nodes in the network. Table 5 provides some detailed values contained in Fig. 4.

5.3.3 Average energy consumed per node

This part is focused on the analysis of the average energy consumption per node. Each protocol has a specific behavior that is more or less stable until near round 950. Afterwards we observe slightly different behaviors for several reasons, such as the decreased number of nodes, the acceleration of node failure and the degradation of their energies.

DCOPA ensures a balanced energy consumption between rounds and a slight difference in variation with a shorter interval. This ensures a balanced energy dissipation between nodes throughout the life of the network. The number of nodes per cluster and the number of CHs and their distribution across the network based on the minimum

Table 4 Example of global energy variation during two rounds

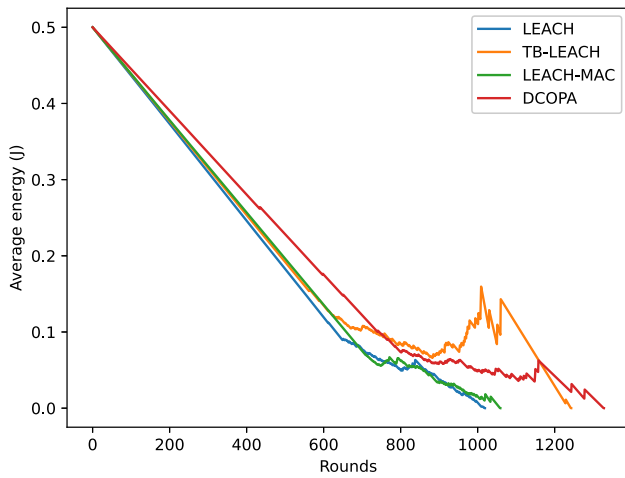| Protocols	LEACH		LEACH-MAC		TB-LEACH		DCOPA	
Rounds	600	800	600	800	600	800	600	800
Energy(j)	11.92	2.44	13.58	2.92	13.08	4.05	17.18	6.69

Fig. 4 Average residual energy per node as a function of rounds

required distance between them make the variation in the average energy consumption of the DCOPA nodes between rounds insignificant. Other protocols have a larger variation interval compared to DCOPA. Figure 5 and Table 6 consolidate our arguments.

5.3.4 Network average energy consumption

In what follows we address the average global consumption by all nodes of the network during each round. The protocol that preserves a certain balanced consumption with a slight variation is indeed DCOPA as shown in Fig. 6. All protocols have a consumption that varies in an interval, as shown in Table 7, until about round 700, where we see a gradual decrease for all four protocols as a result of the beginning of node loss.

The DCOPA protocol has a very small variation interval between rounds compared to other protocols, which means that it manages the total energy consumption in a balanced way across rounds. This property is among the most interesting in managing and balancing energy consumption to increase the lifetime of a given network. DCOPA distinguishes itself with this advantage thanks to the distributed aspect of the protocol, the optimization of the *CH* election process and the access to the role of the *CH* which is sensitive to the local criteria in terms of energy and distance to the *BS*.

5.3.5 Number of alive nodes

The evaluation of the results given in Fig. 7 concerns the most important parameter of the lifetime of a network which is the mortality rate of nodes as a function of rounds. As we can clearly see, the frequency of mortality in DCOPA is much lower compared to other protocols. We present in Fig. 8 the mortality rates which are the First Node who Dies (FND), Half of the Nodes who Die (HND) and the Last Node who Dies (LND).

We find that DCOPA is less efficient in terms of the metric of the first sensor who dies. This is the result of the re-election of some isolated nodes as *CHs* for many successive rounds until their total depletion. These nodes were not solicited by *CHs* nodes although their competition time is pushed towards the end of the time quantum reserved for the self-selection phase of *CHs* $(\tau - \delta)$, which increases their energy consumption in each round. Nevertheless, in the case where there are no isolated nodes we will have better results for the FND, as shown by the confidence interval (CI) in Fig. 21. For HND and LND, our proposed protocol presents interesting results with better performances. This is realized by the self-election of the *CHs* which is sensitive to the residual energy and the distance to the *BS*. This choice to give the chance to nodes close to the *BS* with high residual energy to be *CH* favors an efficient management of the nodes energy. Table 8 illustrates the FND, HND and LND values of the four protocols.

Figure 8, shows the results of a single run on the same network for DCOPA and all the protocols considered for the comparison.

5.3.6 Experimental study of the individual behaviour for the election of *CH* of a sensors Sample

In this experiment, we studied the behavior of a sensors' sample in terms of their election as *CH* (frequency). As we noted on the results obtained in the four Figs. 9, 10, 11 and 12 concerning the four protocols, the election as *CH* (value = 1 if *CH*, 0 otherwise) of the four nodes chosen in the case of the DCOPA protocol is periodic as well as balanced throughout their entire lifetime, unlike the other protocols where the election of these nodes sometimes accentuates and sometimes lingers. The performance of the DCOPA protocol is due to the multi-criteria function (rate
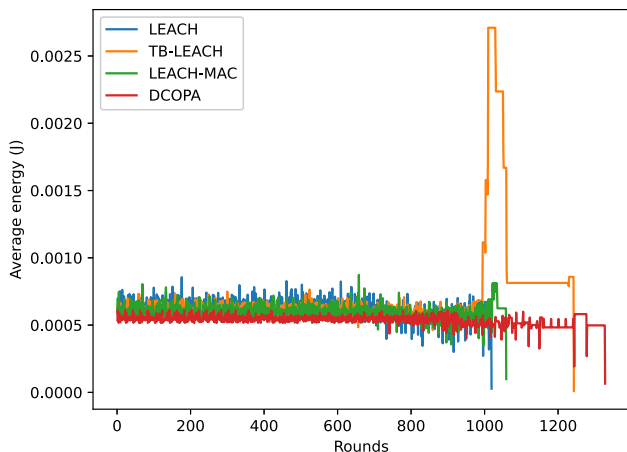
Table 5 variation example of average residual energy per node during two rounds

Protocols	LEACH		LEACH-MAC		TB-LEACH		DCOPA	
Rounds	600	800	600	800	600	800	600	800
Nodes	100	49	100	47	97	48	98	91
Energy(j)	0.119	0.049	0.135	0.062	0.134	0.084	0.175	0.073

Fig. 5 Average energy consumed by a node as a function of rounds

Table 6 Example min/max of average energy consumed by a node as a function of rounds

Protocols	LEACH	LEACH-MAC	TB-LEACH	DCOPA
Rounds	0–950	0–950	0–950	0–950
Min(j)	0.00030	0.00035	0.00046	0.00039
Max(j)	0.00085	0.00087	0.00076	0.00061
Nodes	100–15	100–14	100–17	100–31
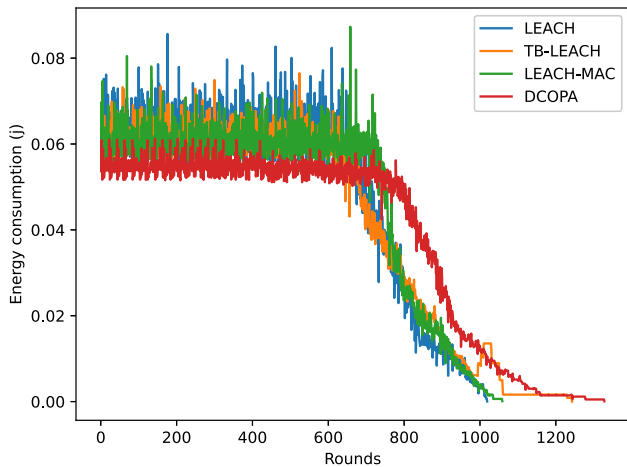
Fig. 6 Average network energy consumption as a function of rounds

Table 7 Example of global energy variation between two rounds

Protocols	LEACH	LEACH-MAC	TB-LEACH	DCOPA
Rounds	0–700	0–700	0–700	0–700
Min(j)	0.040	0.053	0.040	0.050
Max(j)	0.085	0.087	0.076	0.061
Nodes	100–79	100–100	100–72	100–97
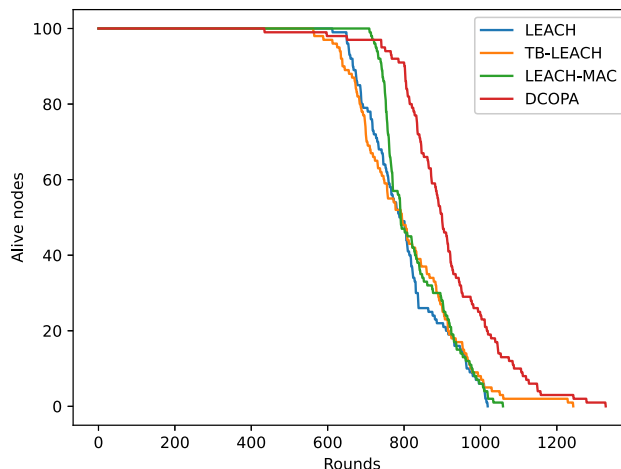
Fig. 7 Number of alive nodes as a function of rounds
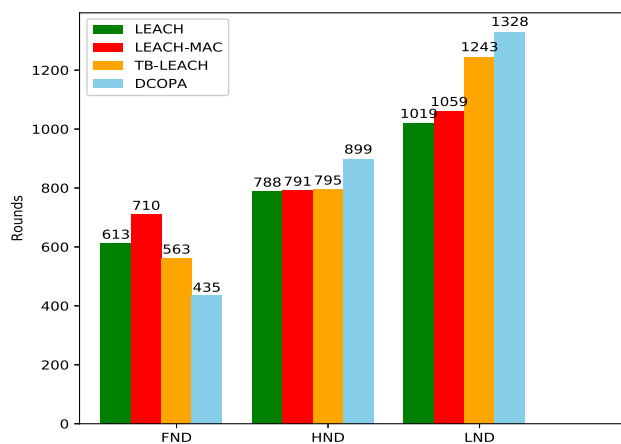
Fig. 8 lifetime: FND, HND and LND

Table 8 Number of alive nodes as a function of rounds

Protocols	LEACH	LEACH-MAC	TB-LEACH	DCOPA
FND	613	710	563	435
HND	788	791	795	899
LND	1019	1059	1243	1328

formula) that decides on the election as *CH* for a given round.

5.3.7 Experimental study of the individual behaviour of energy consumption of a sensors sample

Figures 13, 14, 15 and 16, show the energy consumption of four sensors which were randomly chosen during their lifetimes. We have used as unit of measurements micro-joule (*μJ*) for a better visualisation of the graphs. All
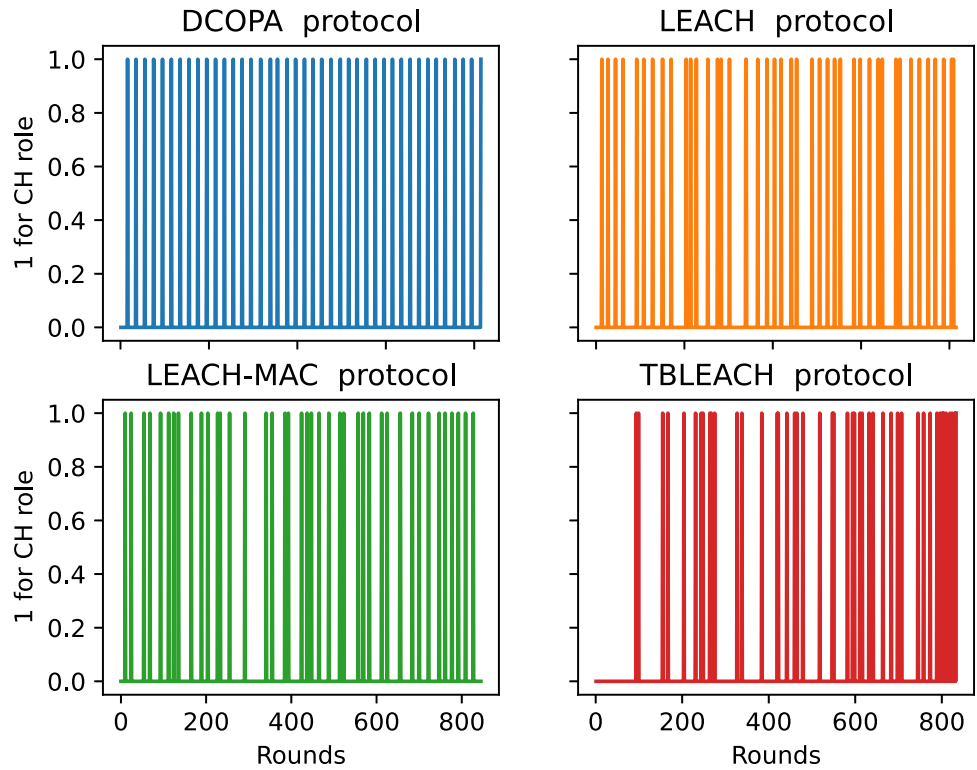
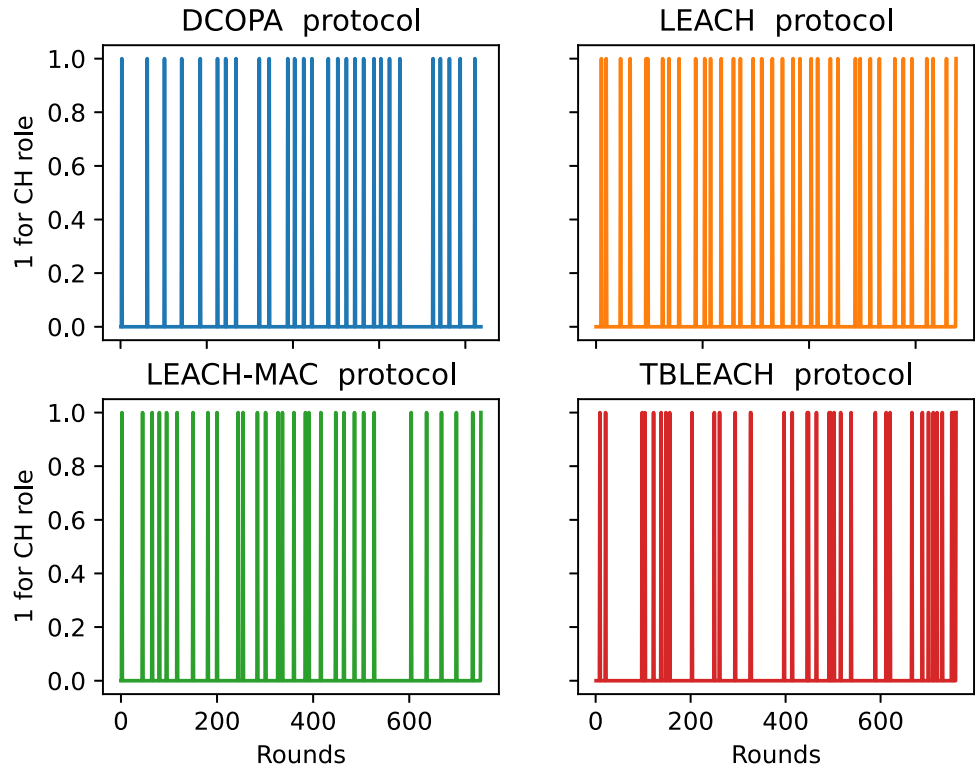**Fig. 9** Node 33 election for *CH* role
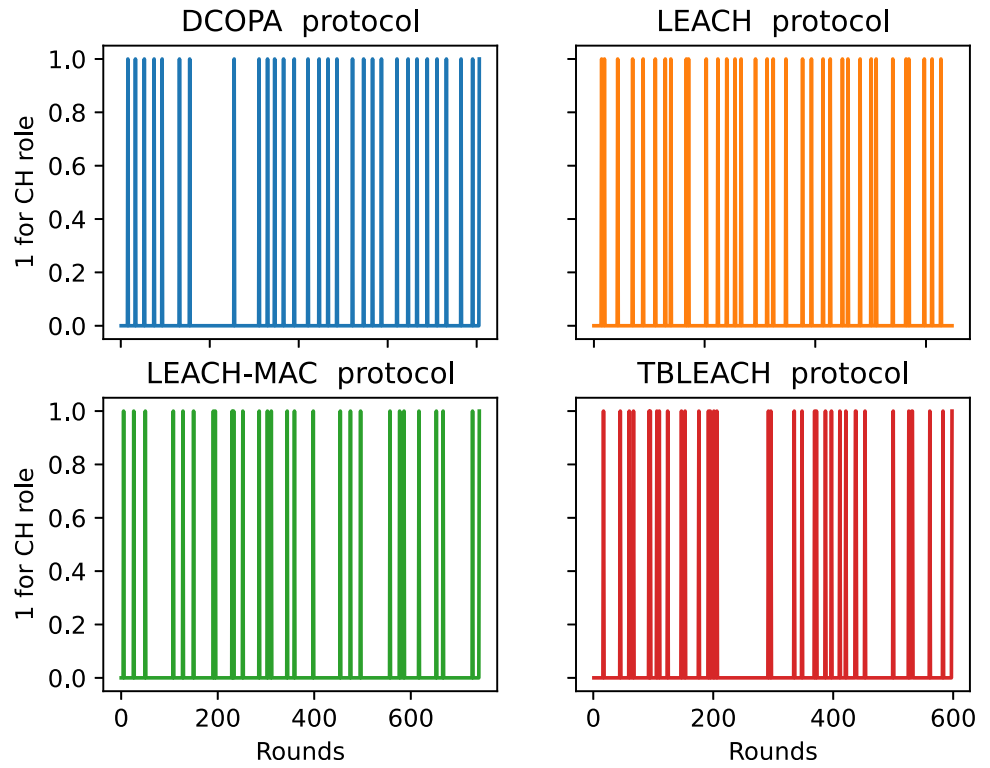
Fig. 10 Node 54 election for *CH* role
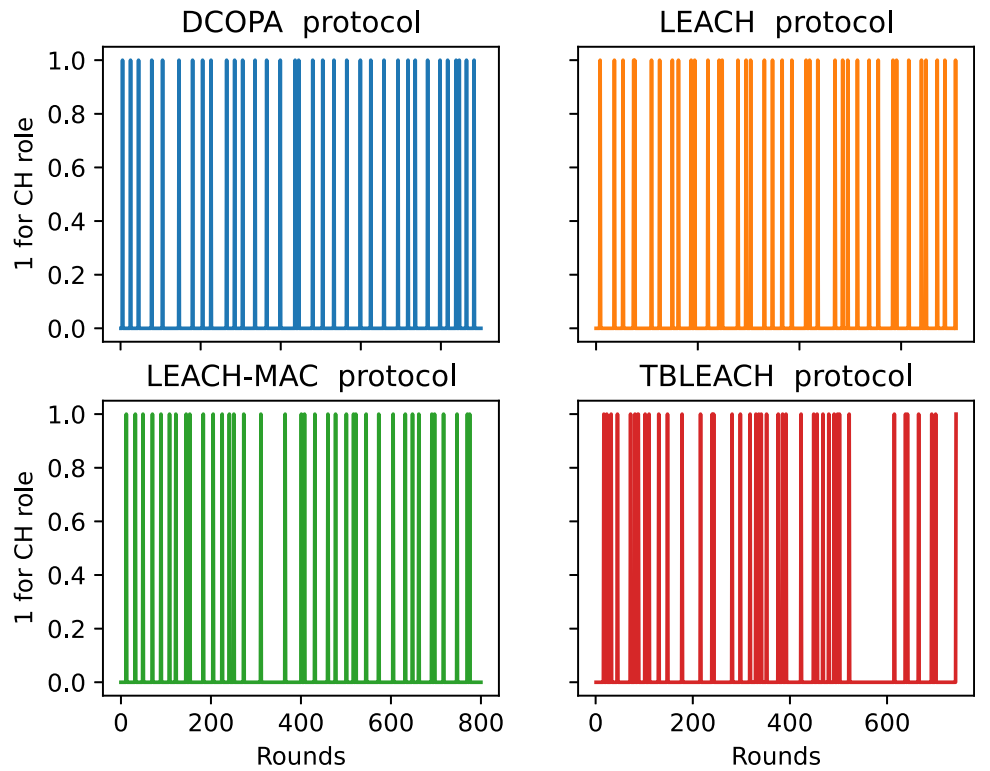
Fig. 11 Node 62 election for *CH* role


Fig. 12 Node 88 election for *CH* role

sensors will have consumption peaks when they are *CH* in given rounds. The performance of our protocol is visible from these figures, the energy consumption of the sensors
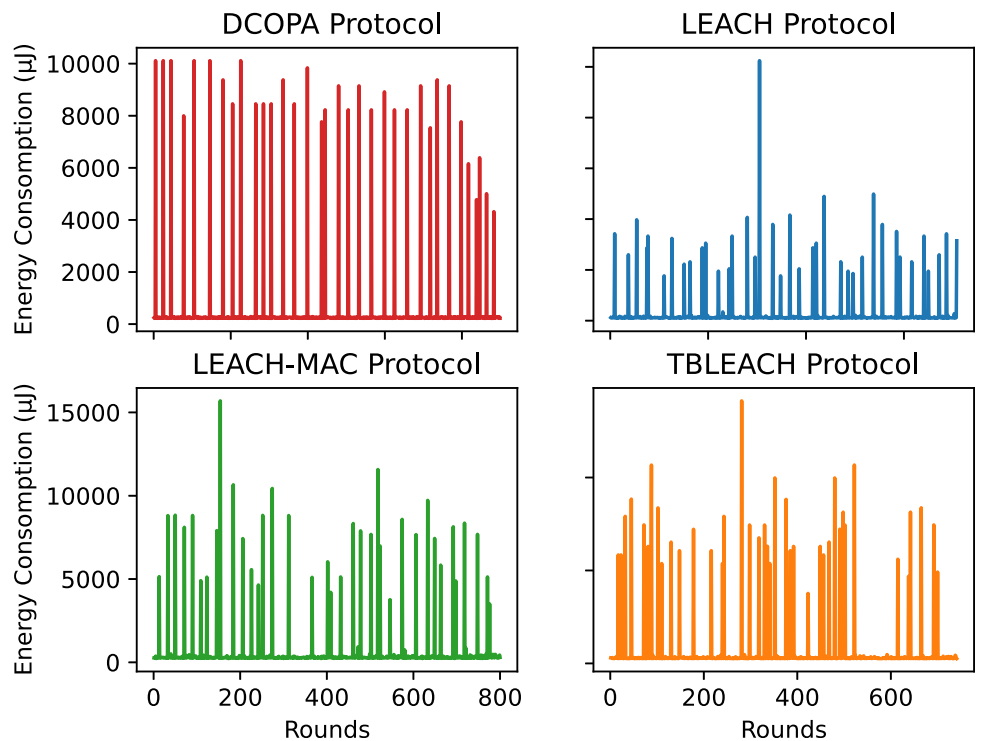
considered when they are *CH* are almost equivalent, this is due to the fact that a *CH* send the solicitations on a fixed radius is calculated which is the same for all other *CHs* in

Fig. 13 Node 33 Energy Consumption (μJ)
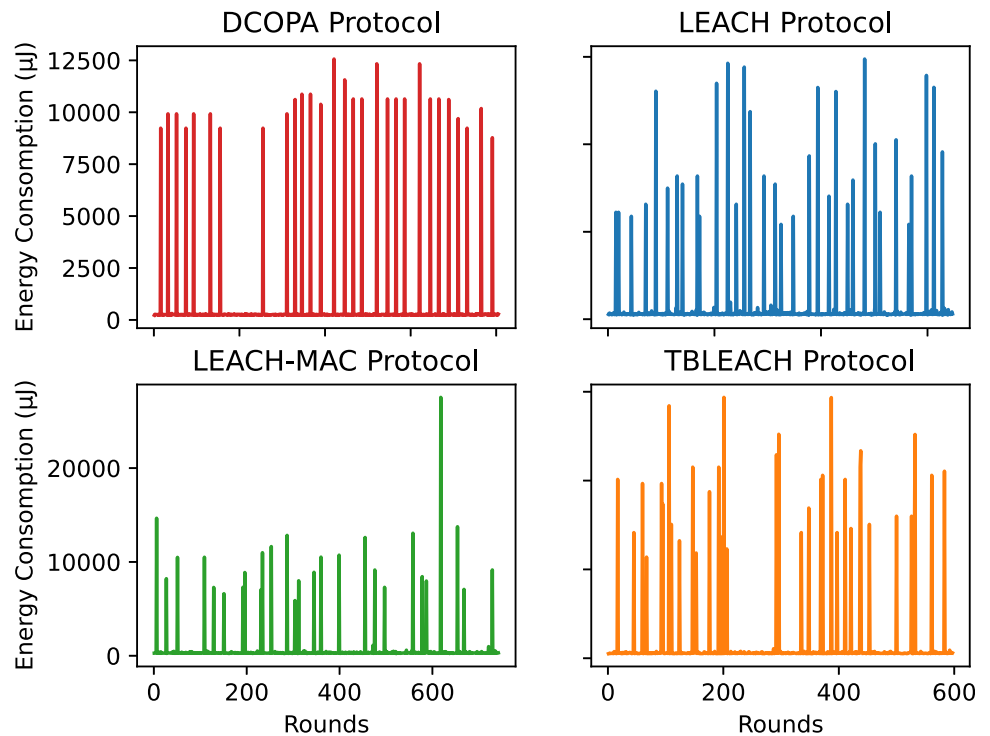
Fig. 14 Node 54 Energy Consumption (μJ)

the network as well as in all rounds. For the other protocols, we observed that there is a variation in consumption each time a sensor is elected as *CH*, this is due to the solicitation radius of the nodes of the cluster which varies as well as the number of normal nodes in the cluster.
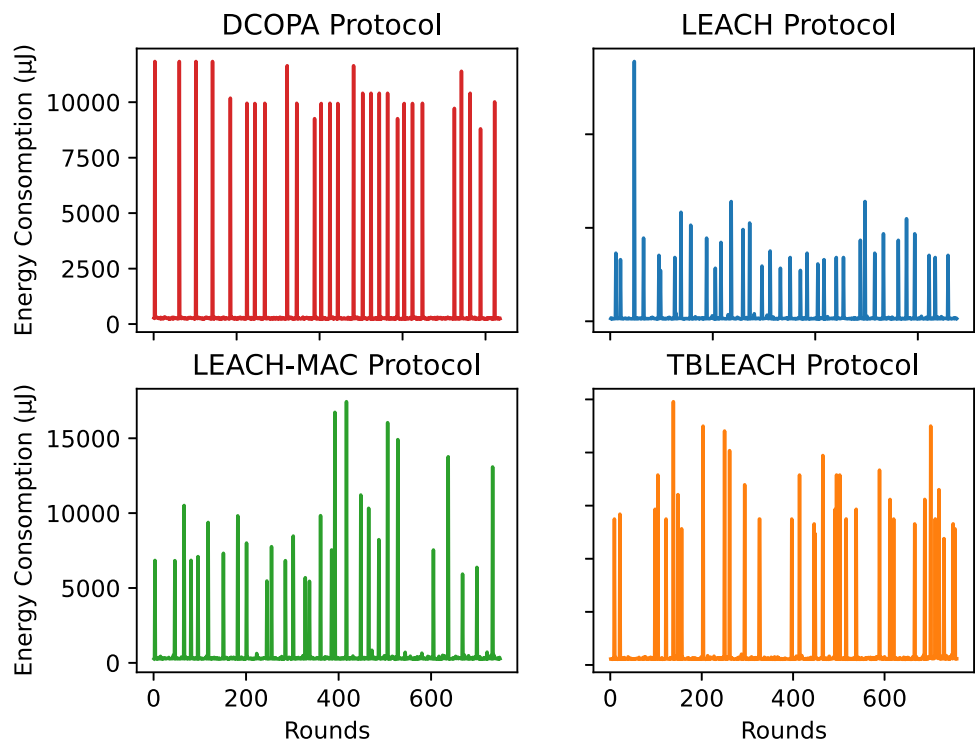
5.3.8 DCOPA performance by increasing the number of nodes

Verifying a protocol developed for a network by increasing the number of nodes (scalability) is an important and

Fig. 15 Node 62 Energy Consumption (μJ)
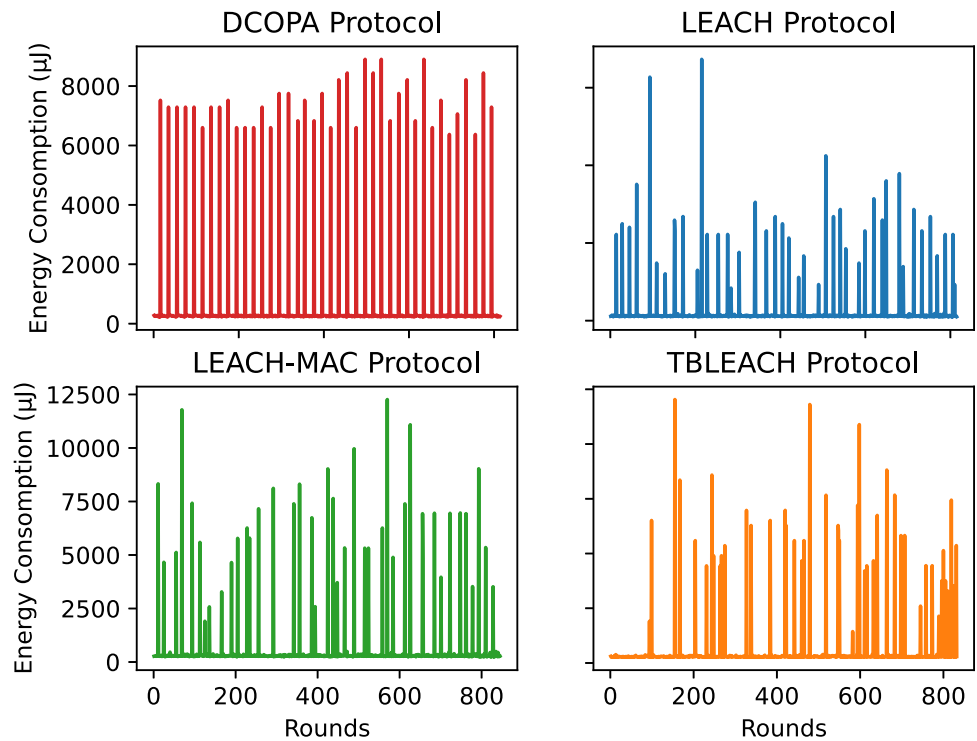
Fig. 16 Node 88 Energy Consumption (μJ)

imperative criterion to test the stability at the level of the performances in spite of the increase of the number of nodes. In our case we have tested the four protocols on a network of 120 nodes and then 150 nodes keeping the same dimensions of the network and the position of the *BS*. We

have retained two situations which are the variation of the global energy of the network as well as the mortality rate or the lifetime of the nodes and the network. As illustrated in Figs. 17, 18, 19 and 20, DCOPA keeps these performances in spite of the increase of the number of nodes in the
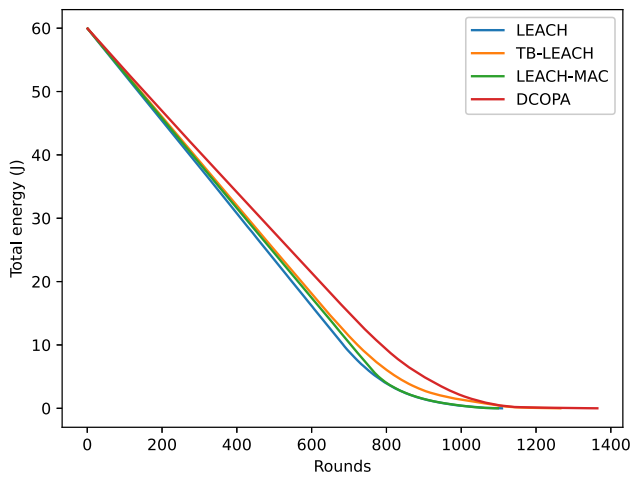
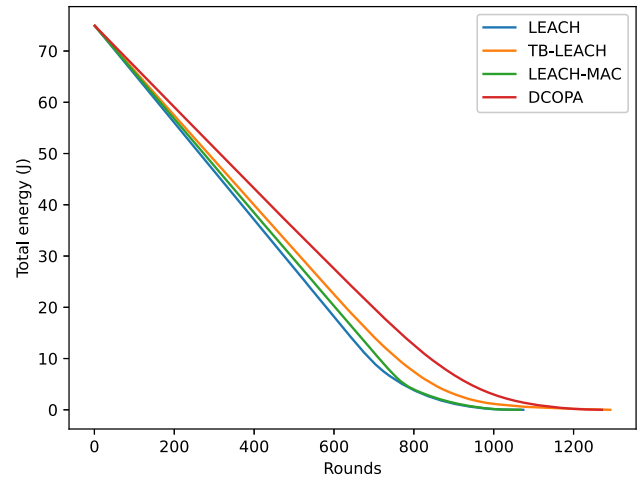**Fig. 17** Total network energy as a function of rounds (120 nodes)

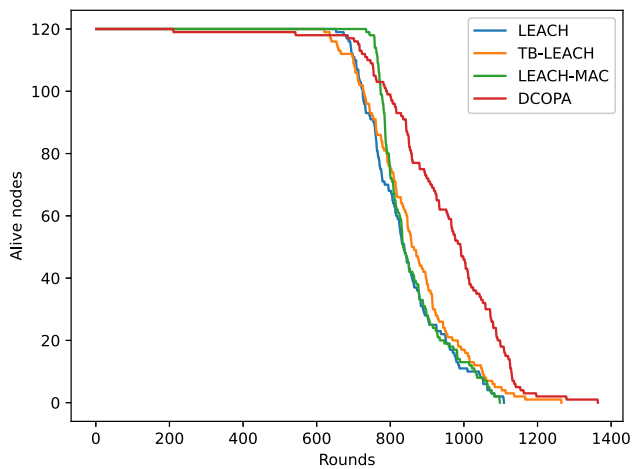**Fig. 19** Total network energy as a function of rounds (150 nodes)

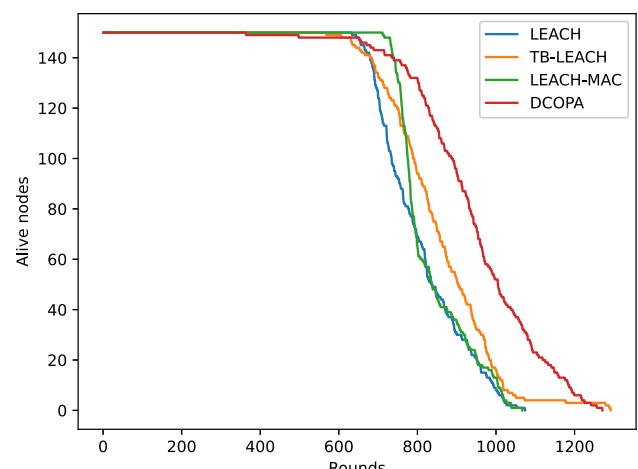**Fig. 18** Number of alive nodes as a function of rounds (120 nodes)

Fig. 20 Number of alive nodes as a function of rounds (150 nodes)

network. This stability is given by the fact that if there is an increase of the number of nodes the RC will be decreasing and so we will have a little more clusters to distribute the loads well between the *CHs*.

5.3.9 Confidence interval and mean values

To calculate the mean values of the life parameters taken into consideration, namely FND, HND and LND, we have generated thirty different networks on which the four protocols run. A mean value is calculated and represented in Fig. 21. Confidence intervals will then be established and shown on the same figure with a 0.95 % confidence. Table 9 displays the mean values and the confidence interval.
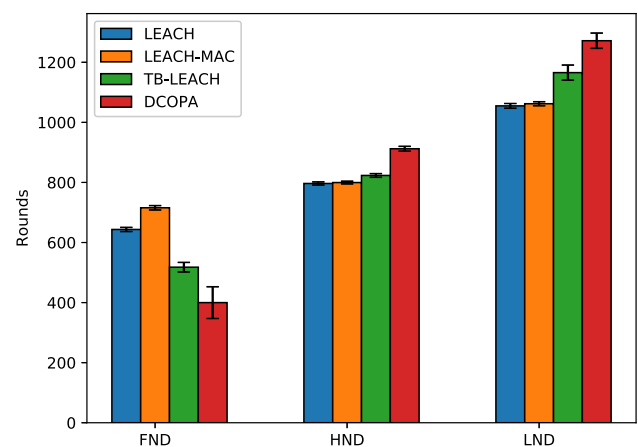

Fig. 21 Mean of the lifetime parameters and confidence interval

Table 9 Mean of the lifetime parameters and confidence interval

Protocols	LEACH		LEACH-MAC		TB-LEACH		DCOPA	
Properties	Mean	CI	Mean	CI	Mean	CI	Mean	CI
FND	643.43	6.90	715.63	7.33	517.83	15.90	399.9	52.86
HND	796.53	5.23	799.6	4.45	823.23	6.17	1165.53	8.03
LND	1054.86	7.90	1062.03	6.68	1165.53	25.22	1271.76	25.49

6 Conclusion

DCOPA is a distributed clustering algorithm for routing in energy-constrained WSN-based IoT systems. A timer $T(i)$ that triggers a competition between nodes to access the role of CH at the start of each round is used, considering the energy of the Node and its distance to the BS which are the local criteria. These two parameters are associated with predefined weights according to the interest we have associated with them. They depend on the type of IoT applications, the type of devices or nodes, the monitoring area, the type of criteria, the position of the BS and the environment in general. Nodes with a minimum weighted sum will be more likely to declare themselves as CHs with maximum residual energy and minimum distance to the BS. Our contribution solves the major shortcomings of LEACH, namely, the inefficient distribution of CHs across the deployment area, the variable number of CHs that does not respect the optimal number of CHs considered at the start, the large number of rounds with none CHs and the declaration of CHs that is carried out on a maximum distance that covers the whole network. Simulation results show that DCOPA performs well regarding lifetime compared to the basic LEACH and other LEACH enhancements which are TB-LEACH and LEACH-MAC. In future work we will continue to work on other CH selection criteria as well as multihop routing to improve lifetime performance even more. Other lifetime aspects such as connectivity and coverage will also be explored as comparison parameters as they are important in the performances of IoT networks.

Declarations

Conflict of interest Not applicable.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Li, S., Da Xu, L., Zhao, S.: The internet of things: a survey. Inf. Syst. Front. **17**(2), 243–259 (2015)
2. Evans, D.: The internet of things: how the next evolution of the internet is changing everything. CISCO White Pap. **1**(2011), 1–11 (2011)
3. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., Ayyash, M.: Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun. Surv. Tutor. **17**(4), 2347–2376 (2015)
4. Rault, T., Bouabdallah, A., Challal, Y.: Energy efficiency in wireless sensor networks: a top-down survey. Comput. Netw. **67**, 104–122 (2014)
5. Zaidi, S., Atiquzzaman, M., Calafate, C.T.: Internet of flying things (ioft): a survey. Comput. Commun. (2020)
6. Sung, Y., Lee, S., Lee, M.: A multi-hop clustering mechanism for scalable iot networks. Sensors **18**(4), 961 (2018)
7. Priyanga, M., Vimalraj, S.L.S., Lydia, J.: Energy aware multiuser & multi-hop hierarchical–based routing protocol for energy management in wsn-assisted iot, 701–705. IEEE (2018)
8. Sonthalia, M., Jha, A., Gupta, U., Thyagarajan, J.: A real time implementation of hierarchical routing protocol for iot based wireless sensor network, 512–516. IEEE (2019)
9. Behera, T.M., et al.: Residual energy-based cluster-head selection in wsns for iot application. IEEE Internet Things J. **6**(3), 5132–5139 (2019)
10. Sabet, M., Naji, H.R.: A decentralized energy efficient hierarchical cluster-based routing algorithm for wireless sensor networks. AEU-Int. J. Electron. Commun. **69**(5), 790–799 (2015)
11. Heinzelman, W.R., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. IEEE (2000)
12. Heinzelman, W.B., Chandrakasan, A.P., Balakrishnan, H.: An application-specific protocol architecture for wireless microsensor networks. IEEE Trans. Wirel. Commun. **1**(4), 660–670 (2002). https://doi.org/10.1109/TWC.2002.804190
13. Hosseinzadeh, M., Hemmati, A., Rahmani, A.M.: Clustering for smart cities in the internet of things: a review. Clust. Comput., pp 1–31 (2022)
14. Triantaphyllou, E.: In: Multi-criteria Decision Making Methods, pp. 5–21. Springer, Berlin (2000)
15. Shahraki, A., Taherkordi, A., Haugen, Ø., Eliassen, F.: Clustering objectives in wireless sensor networks: a survey and research direction analysis. Comput. Netw. **180**, 107376 (2020)
16. Abbasi, A.A., Younis, M.: A survey on clustering algorithms for wireless sensor networks. Comput. Commun. **30**(14–15), 2826–2841 (2007)
17. Liu, X.: Atypical hierarchical routing protocols for wireless sensor networks: a review. IEEE Sens. J. **15**(10), 5372–5383 (2015)

18. Pachlor, R., Shrimankar, D.: Lar-ch: a cluster-head rotation approach for sensor networks. IEEE Sens. J. **18**(23), 9821–9828 (2018)

19. Jia, D., Zhu, H., Zou, S., Hu, P.: Dynamic cluster head selection method for wireless sensor network. IEEE Sens. J. **16**(8), 2746–2754 (2016)

20. Roy, N.R., Chandra, P.: A note on optimum cluster estimation in leach protocol. IEEE Access **6**, 65690–65696 (2018)

21. Liao, Y., Qi, H., Li, W.: Load-balanced clustering algorithm with distributed self-organization for wireless sensor networks. IEEE Sens. J. **13**(5), 1498–1506 (2013)

22. Aslam, M., et al.: Survey of extended leach-based clustering routing protocols for wireless sensor networks, pp. 1232–1238. IEEE (2012)

23. Hani, R.M.B., Ijjeh, A.A.: A survey on leach-based energy aware protocols for wireless sensor networks. J. Commun. **8**(3), 192–206 (2013)

24. Rahayu, T.M., Lee, S.-G., Lee, H.-J.: Survey on Leach-Based Security Protocols, pp. 304–309. IEEE (2014)

25. Arora, V.K., Sharma, V., Sachdeva, M.: A survey on leach and other's routing protocols in wireless sensor network. Optik **127**(16), 6590–6600 (2016)

26. Singh, S.K., Kumar, P., Singh, J.P.: A survey on successors of leach protocol. IEEE Access **5**, 4298–4328 (2017)

27. Murata, T., Ishibuchi, H.: Performance Evaluation of Genetic Algorithms for Flowshop Scheduling Problems, pp. 812–817. IEEE (1994)

28. Agarwal, P.K., Procopiuc, C.M.: Exact and Approximation Algorithms for Clustering, SODA '98, pp. 658–667. Society for Industrial and Applied Mathematics, USA (1998)

29. Chen, J., Shen, H.: Meleach-l: More Energy-Efficient Leach for Large-Scale WSNs, pp. 1–4. IEEE (2008)

30. Chen, J., Shen, H.: Meleach an energy-efficient routing protocol for WSNs. Chin. J. Sens. Actuators **9**, 035 (2007)

31. Acharya, T., Chattopadhyay, S., Roy, R.: Energy-Aware Virtual Backbone Tree for Efficient Routing in Wireless Sensor Networks, pp. 96–96. IEEE (2007)

32. Junping, H., Yuhui, J., Liang, D.: A Time-Based Cluster-Head Selection Algorithm for Leach, pp. 1172–1176. IEEE (2008)

33. Ali, M.S., Dey, T., Biswas, R.: Aleach: Advanced Leach Routing Protocol for Wireless Microsensor Networks, pp. 909–914. IEEE (2008)

34. Hong, J., Kook, J., Lee, S., Kwon, D., Yi, S.: T-leach: the method of threshold-based cluster head replacement for wireless sensor networks. Inf. Syst. Front. **11**(5), 513 (2009)

35. Tong, M., Tang, M.: Leach-b: An Improved Leach Protocol for Wireless Sensor Network, pp. 1–4. IEEE (2010)

36. Beiranvand, Z., Patooghy, A., Fazeli, M.: I-leach: An Efficient Routing Algorithm to Improve Performance & to Reduce Energy Consumption in Wireless Sensor Networks, pp. 13–18. IEEE (2013)

37. Chen, H., Zhang, C., Zong, X., Wang, C.: Leach-g: an optimal cluster-heads selection algorithm based on leach. J. Softw. **8**(10), 2660–2667 (2013)

38. Salim, A., Osamy, W., Khedr, A.M.: Ibleach: intra-balanced leach protocol for wireless sensor networks. Wirel. Netw. **20**(6), 1515–1525 (2014)

39. Sasikala, A., Sangameswaran, N., Aravindh, P., et al.: Improving the energy efficiency of leach protocol using VCH in wireless sensor network. Int. J. Eng. Dev. Res **3**(2), 918–924 (2015)

40. Hussain, S., Matin, A.W.: Energy efficient hierarchical cluster-based routing for wireless sensor networks. Jodrey School of Computer Science Acadia University Wolfville, Nova Scotia, Canada, Technical Report, pp. 1–33 (2005)

41. Zhao, Z., Li, G., Xu, M.: An Improved Algorithm Based on Leach Routing Protocol, pp. 1248–1251. IEEE (2019)

42. Suwandhada, K., Panyim, K.: Aleach-plus: An Energy Efficient Cluster Head Based Routing Protocol for Wireless Sensor Network, pp. 1–4. IEEE (2019)

43. Mansura, A., Drieberg, M., Aziz, A.A., Bassoo, V.: Multi-energy Threshold-Based Routing Protocol for Wireless Sensor Networks, pp. 71–75. IEEE (2019)

44. Samant, T., Mukherjee, P., Mukherjee, A., Swain, T., Datta, A.: Leach-v: a solution for intra-cluster cooperative communication in wireless sensor network. Indian J. Sci. Technol. **9**(48) (2016)

45. Chakraborty, B., Mukherjee, P., Mukherjee, A., Swain, T.: V-leach-vd: a solution for higher energy utilization in low power communication in wireless sensor network. J. Adv. Res. Dyn. Control Syst. **9**(SI–14), 581–591 (2017)

46. Suman, J., Shyamala, K., Roja, G.: Improving Network Lifetime in WSN's Based on Maximum Residual Energy, pp. 1–5. IEEE (2021)

47. Panchal, A., Singh, L., Singh, R.K. :Rch-leach: Residual Energy Based Cluster Head Selection in Leach for Wireless Sensor Networks, pp. 322–325. IEEE (2020)

48. Esposito, C., Ficco, M., Castiglione, A., Palmieri, F., Lu, H.: Loss-tolerant event communications within industrial internet of things by leveraging on game theoretic intelligence. IEEE Internet Things J. **5**(3), 1679–1689 (2018)

49. Chang, L., Li, F., Niu, X., Zhu, J.: On an Improved Clustering Algorithm Based on Node Density for WSN Routing Protocol. Clust. Comput., pp. 1–13 (2022)

50. Mtopi, Y.B.C., et al.: Multihop Optimal Time Complexity Clustering for Emerging IoT Applications. Clust. Comput., pp. 1–17 (2022)

51. Dietrich, I., Dressler, F.: On the lifetime of wireless sensor networks. ACM Trans. Sens. Netw. (TOSN) **5**(1), 1–39 (2009)

Foudil Mir received a magister degree on networks and distributed systems at the doctoral school on networks and distributed systems, university of Bejaia, Algeria. He is currently an assistant lecturer of computer science at the same university and a member of the Medical Computing Laboratory (LIMED). His research activities are focused on ad hoc networks, wireless sensors networks (WSNs) for the internet of things applications, algorithms and clustering approaches and energy optimization in WSNs hierarchical routing protocols.

Farid Meziane obtained his Ph.D in computer science from the university of Salford on his work on producing formal specifications from natural language requirements. He is currently holding a chair in data science and the head of the data science research centre at the university of Derby, United Kingdom. He specialises in extracting, organising and interpreting data and knowledge under all its form from various sources including quantitative data, text and images.