The Journal of Supercomputing
https://doi.org/10.1007/s11227-024-06426-w

Unequal-radius clustering in WSN-based IoT networks: energy optimization and load balancing in UDCOPA protocol

Foudil Mir[1] · **Farid Meziane**[2]

Accepted: 2 August 2024

Abstract

The internet of things (IoT) is an exponentially growing network of physical objects equipped with sensors, software and network connectivities to collect, process, transmit and receive data. Wireless sensor networks (WSNs) play an essential role in supporting the IoT. These networks, made up of nodes with the ability to monitor their environment, enable the collection and transmission of specific data in real time, offering enhanced applications and services within IoT networks. This symbiosis between WSN and IoT can be defined as WSN-based IoT. The complexity of WSN-based IoT lies in the effective management of these varied devices, each with its own distinct capabilities. Clustering is a popular technique for reducing the communication load, conserving energy, aggregating data and optimizing the performance of WSN-based IoT systems. Once the cluster heads (CHs) are chosen, conventional clustering algorithms typically use a single radius of clustering (RC) to group devices into multiple clusters. However, this approach may not be optimized for WSN-based IoT networks, as devices may have different features, for example, the residual energy (R_{Enrg}) and the distance to the base station (DistBS). In a previous work, we proposed the DCOPA (a distributed clustering based on objects performances aggregation for hierarchical communications in IoT applications) protocol for clustering in WSN-based IoT networks. DCOPA applies the same clustering algorithm to the elected CHs, without considering their distinctions in terms of R_{Enrg} and DistBS. The proposed new approach, called unequal-DCOPA (UDCOPA), allows us to define for each CH its adaptive radius of clustering (ARC) which will be sensitive to the local criteria of R_{Enrg} and DistBS of the CH concerned. The ARC is modeled as a multicriteria system applied to each CH. Simulation results show that our new UDCOPA approach outperforms DCOPA and LEACH protocols for energy management, load balancing, scalability and network lifetime. UDCOPA increases lifetime by (62.61%) over LEACH and by (32.72%) over DCOPA.

Extended author information available on the last page of the article

Published online: 29 August 2024

Keywords WSN-based IoT · DCOPA · UDCOPA · Unequal clustering · Data communication · Load balancing · Energy efficiency · Multicriteria approach

1 Introduction and motivation

1.1 Introduction

WSN-based IoT networks are now indispensable in diverse fields such as the environment, logistics, industry, home and urban automation, disaster management and medical care, as well as in the security and monitoring of critical infrastructures [1, 2]. These two closely related fields, WSN and IoT, have witnessed exponential growth in the number of research projects in recent years. Clustering [3–5], a technique commonly used in WSN-based IoT networks for energy optimization and lifetime enhancement, involves grouping nodes into multiple clusters to optimize well-screened network performance. Most of the work in the literature has focused on clusters of equal size. However, in reality, different nodes in WSN-based IoT networks may have different processing power, R_{Enrg}, DistBS, communication range and several other criteria. Our contribution focuses on the clustering of unequal sizes in WSN-based IoT. The aim is to explore how this approach can optimize the energy performance of WSN-based IoT networks by taking into account existing disparities between nodes, namely R_{Enrg} and DistBS. The DCOPA protocol [6, 7] is a distributed clustering algorithm based on multicriteria decision making [8]. CHs are elected through a competition in which all sensors participate on the basis of a timer $T(i)$ calculated according to two local criteria: the node's R_{Enrg} and its DistBS. $T(i)$ is obtained using multicriteria aggregation with predefined weights associated with the two selected criteria. DCOPA considerably ameliorates the shortcomings of the LEACH (low-energy adaptive clustering hierarchy) protocol [9], notably the fact that a CH is advertised by sending a message over a maximum distance that covers the entire network, and the abundance of randomness in the process of electing CHs. The aim of this work is to analyze and further improve the DCOPA protocol in such a way as to increase network lifetime with balanced clustering and energy efficiency. DCOPA uses the R_{Enrg} and the DistBS of a node as two primary criteria that define its qualification to win the CH role by calculating a $T(i)$. Each criterion is assigned a specific weight to express its relevance. Our contribution consists in analyzing the DCOPA protocol and, more particularly, in examining the clustering process, focusing on the RC used by elected CHs when announcing their elections. A major drawback of this protocol is that the same RC is applied to all CHs, independently of differences between them in terms of their local criteria such as R_{Enrg} and DistBS. This unified approach, used by the DCOPA protocol, does not consider the specific features of each CH, thus inevitably triggering failures of nodes lacking sufficient energy for clustering over a very considerable radius such as that required by the protocol, which is constant for all CHs throughout the network's lifetime. The DistBS of a CH was also not taken into account when broadcasting its candidacy. To address this limitation, our study aims to explore an alternative approach that dynamically adjusts the RC according to the local criteria of a node elected as CH. Adopting our new adaptive approach, we aim to model the computation of the RC using the multicriteria approach to associate two essential criteria of each CH to decide its RC while assigning weights depending on their interests in the context of the application. As a result, we will design a distributed multicriteria approach for unequal RC in WSN-based IoT networks to address a major drawback of the DCOPA protocol and highlight the importance of the unequal radius approach for clustering in WSN-based IoT. By exploiting the differences between nodes in a WSN-based IoT network, this approach opens up new potential for optimizing the energy efficiency of nodes and the entire network, boosting network lifetime and balancing the load between CHs nodes according to their individual performance.

1.2 Motivation

This section details the fundamental motivations that were behind our research approach.

- *Performance optimization*: The main motivation behind our work is to improve the performance of the DCOPA protocol in the clustering operation. Currently, the same clustering algorithm is used across all CHs, without considering their disparities in terms of R_{Enrg} and DistBS.
- *Energy management*: Another key aim of our study is to optimize energy consumption and network lifetime. This is achieved by empowering each CH to adjust its RC according to its specific level of performance.
- *Reduction of long-distance communications*: By integrating DistBS and R_{Enrg} into the setting of each CH's RC, our UDCOPA [10] approach aims to attenuate long-distance communications, thus ensuring a more efficient use of network resources.
- *Multicriteria and distributed nature of the contribution*: We adopt a multicriteria approach to modeling the RC calculation. This approach simultaneously considers R_{Enrg} and DistBS as criteria, providing for more refined and adaptive decision-making.

By pursuing these research motivations, our goal is to enhance performance, energy efficiency and communication quality within WSN-based IoT networks, through an unequal clustering approach based on adaptive RC.

1.3 Research questions

- Limitations of the DCOPA protocol:
    - What are the potential limitations of the DCOPA protocol when the RC is kept unchanged for all CH nodes throughout the lifetime of the network?
    - How might these limitations impact on energy performance, load balancing and network scalability?

- Improving the DCOPA protocol:
  - How can we improve the DCOPA protocol by introducing RC adjustment to an ARC aware of the context and local criteria of an elected CH?
  - How can this adaptation be adjusted according to the level of R_{Enrg} and the DistBS of each CH?
- Impact of ARC on energy performance, load balancing and network lifetime:
  - What will be the impact of introducing ARC into the DCOPA protocol clustering process?
  - How will ARC influence critical parameters such as energy consumption, node death rate, load balancing, cluster distribution and geographic coverage?

These research questions have oriented our approach to enhancing the DCOPA protocol by integrating adaptive and adjustable RC, enabling greater flexibility and dependability in response to CH node criterion conditions.

The following is the structure of the rest of the paper. Related work is given in Sect. 2. In Sect. 3, we explain the concept of the optimum number of clusters and the energy model that we will employ in our simulations. Section 4 outlines the theoretical and formal foundations of the DCOPA protocol. Section 5 is reserved for the development of our new UDCOPA approach. In Sect. 6, we will illustrate the performance evaluation of the proposed approach. The final section is devoted to conclusions of our contribution.

2 Related works

This section is dedicated to a review of the few related works focusing on the concept of unequal clustering in WSN-based IoT. We discuss some algorithms and protocols proposed in the literature to dynamically adapt and adjust the RC according to some specific criteria which are relevant for WSN-based IoT networks. By studying these works, we are clearly able to position our proposed approach and demonstrate the effectiveness of our contribution in the area of unequal clustering. In this section, it is essential to examine some protocols for building equal-sized clusters in WSN-based IoT networks. We start with the protocol considered in the literature to be the founder of clustering in WSN. Heinzelman et al. [9] proposed LEACH, a distributed protocol for dynamic, probabilistic clustering. The protocol uses a threshold calculated by each node in the network according to a number of parameters. During the setup phase, nodes calculate their $T(i)$ value and declare themselves as CHs by sending an advertisement CH message (*ADV-CH*) into the network if a random number generated (between 0 and 1) is less than $T(i)$. The steady state phase is dedicated to cluster formation and data communication. Ordinary nodes choose the nearest CH. The CHs aggregate the data and transmit it to the base station (BS). In a large-scale context, the authors in [11] have applied the concept of regression to a variant of LEACH in order to predict the lifetime of a large-scale network based on a network with a reduced number of nodes. With the DCOPA protocol (a distributed clustering based on objects performances aggregation for hierarchical communications in IoT applications) proposed by Mir et al. [6], the network nodes organize a competition between themselves by calculating a timer $T(i)$ as a function of the R_{Enrg} and the DistBS for the election of CHs. Each node decrements its $T(i)$. If $T(i)$ reaches 0, the node declares itself CH and cancels the candidacies of the other nodes for the role of CH on an RC radius who will drop out of the competition and send a joining message to the nearest CH. Unequal clustering has been shown to be more effective than equal clustering, as assigning the same RC to all CHs in the network, without taking into account their respective performance, inevitably leads to load imbalance and the exhaustion of some CHs with little R_{Enrg}. Determining this RC for CHs is a complex issue when designing and modeling unequal clustering protocols. Haleem et al. in [12] introduced and evaluated a novel grid-based hybrid network deployment (GHND) framework for WSN. The framework is designed to enhance energy efficiency and load balancing through a merge and split technique that ensures even distribution of sensor nodes across the network. Specifically, the method constructs a grid where low-density zones are merged and high-density zones are split to address the hotspot problem. The authors in [13] proposed an improved grid-based hybrid network deployment (IGHND) scheme for zone head (ZH) selection and reselection in WSN-based IoT. The framework enhances energy efficiency and network stability by considering multiple criteria for ZH selection, including energy level, distances from neighboring nodes and the zone center, the number of times a node has been ZH, and whether a node is merged. These criteria are evaluated using the analytical network process (ANP), a multicriteria decision-making tool. The authors in [14] proposed the DSBCA (Balanced clustering algorithm with distributed self-organization) protocol, which focuses on balanced clustering with distributed self-organization. The cluster size is adjusted based on the distance from the BS and node density, employing a larger cluster radius for greater distance and lower density, and vice versa. DSBCA aims to achieve cluster load balancing while minimizing communication costs. The main idea of [15] is the introduction and evaluation of COSBIoT, a centralized algorithm designed to optimize energy-efficient maximum area coverage in sensor-based IoT (SBIoT) frameworks. By leveraging Miquel triangles (MT) and anticomplementary triangles (ACT), COSBIoT schedules sensor nodes in sleep-awake cycles to effectively cover dense and randomly deployed networks. The authors in [16] proposed a novel scheme for clustering and event detection in sensor-based IoT (SBIoT) frameworks that incorporate user context. This approach divides the deployment region into sub-regions, forming sensing clusters based on user-defined parameters, which enables accurate and context-aware event detection. Additionally, the scheme utilizes communication clusters and compressive data gathering to enhance energy efficiency. The authors in [17] proposed a hybrid cluster-based smart random walk (CBSRW) routing technique for WSN aimed at enhancing reliable data transmission and network lifetime. By integrating random walk (RW) with clustering methodologies, the CBSRW approach facilitates multihop data gathering and aggregation, effectively reducing redundant data transmission. Proposed in [18], the EADUC (Energy-aware distributed unequal clustering) protocol suggests the formation of clusters with unequal sizes by assigning varying radius of competition to the nodes. Smaller clusters are favored for CHs closer to the BS to conserve energy for intercluster data communications. The EECS (Energy efficient clustering scheme) protocol, presented in [19], proposes a clustering method incorporating competition between candidate CHs. A node chooses the CH based on its energy and the load balancing of the CHs with two distance factors. The CH selection decision is guided by a cost function, combining intra and intercluster distances with a weighting factor. The aim is to achieve a balance between intracluster energy consumption and intercluster communication load. An in-depth taxonomy, comparison, analysis and discussion of unequal clustering protocols is reported in [20], investigating scalability, energy efficiency and load balancing. The lack of exploration of unequal clustering protocols was expressed by the authors, despite their potential for load balancing, hotspot remediation and energy efficiency. Unequal clustering protocols are categorized according to cluster size control methods based on (i) node degree like power aware dynamic clustering protocol (PADCP) [21], dynamic transmission power control method (DTPC) [22] and dynamic load balanced clustering-problem (DLBCP) [23], (ii) relay load like: energy efficient unequal clustering (EEUC) [24], cross layer unequal clustering routing algorithm (CUCRA) [25], UCS [26], EEC-SCH [27], UBUC [28] and fuzzy logic-based relay load balancing unequal clustering protocol EAUCF [29] and (iii) the combination of both as: IEEUC [30], IFUC [31], PSO-UFC [32] and EDDUCA [33]. The authors in their review, as detailed in [34], have categorized unequal clustering protocols into three basic classes: probabilistic, deterministic, and preset. The pobabilistic category is further subdivided into random, including protocols such as PRODUCE [35], EDUC [36], LUCA [37] and Hybrid, including protocols such as EEUC [24] and UCR [38]. The deterministic category is structured as weight, encompassing approaches such as, ACT [39], EADUC [18], CUCA [40], fuzzy with protocols such as FUCP [41], FBUC [42], DUCF [43], heuristic including SMEBUC [44], GAEEP [45] and compound like EDDUCA [33]. Finally, the preset category, with the UCS protocol [26] as an example. Pravin et al. [46] developed a protocol to improve energy efficiency, energy balance and network scalability in IoT-enabled WSN using a genetic algorithm (GA) for CHs selection, considering factors likedensity, stability, node energy and capacity. The protocol featured a stochastic CHs selection model (SCHSM) and included multiple movable sink nodes. Sankar et al. [47] introduced a novel fuzzy-based Harris–Hawks optimization (FHHO) algorithm for optimal CHs selection in IoT networks, aiming to address the energy scarcity problem and extend network lifetime. The FHHO algorithm takes into account residual energy and the distance between the sink and the node, using fuzzy logic to evaluate the fitness function. The authors in [48] developed a new approach for optimizing CHs selection and the path of data transmission in IoT-enabled smart agriculture systems. This novel approach integrates whale optimization algorithm (WOA), modified fuzzy logic and enhanced crow swarm optimization (ECSO). Fuzzy logic evaluates parameters like distance, trust, energy, overhead and node degree for CH selection, while ECSO optimizes the data transmission path. Recent research introduces an energy-efficient data packet aggregation scheme for long range (LoRa) communication in low-power wide area network (LPWANs) to address high-energy consumption. The proposed load balancing algorithm [49] improves network performance and energy savings. The load

Table 1 ECM parameters

Parameters	Values	Description
E_{elec}	50 nJ/bit	Required energy to run electronic circuit
E_{mp}	0.0013 pJ/bit/m^4	Multi path propagation
E_{fs}	10 pJ/bit/m^2	Free space propagation
E_{DA}	5 nJ/bit/signal	Required energy for data aggregation

balancing model used in the study focuses on optimizing data transmission in LoRa networks by strategically managing how data packets are aggregated and transmitted among nodes. Compared to conventional star connected LoRa and other protocols, this approach significantly enhances energy efficiency, prolongs network lifetime and increases stability.

3 Optimum number of clusters and the energy model

3.1 Energy consumption model

The energy consumption model (ECM) from Heinzelman et al. [50] is used in our simulations. It measures the energy required for transmission as a function of message size (l) and communication distance (d). The reception energy is determined by the size of the received message (l). Data aggregation energy is covered by this model. When transmitting, the energy consumed is given by $E_{Tx}(l, d)$ shown in Formula 2. Depending on the distance d, two power control settings and two channel models are applied as follows (d_0 is specified in Formula 3):

- Utilize the free space power amplifier E_{fs} and the channel corresponding to the free space model (d^2 power loss) when $d < d_0$.
- Utilize the multipath power amplifier E_{mp} and the channel corresponding to the multipath fading model (d^4 power loss) when $d \geq d_0$.

$$E_{Tx}(l, d) = E_{Tx-elec}(l) + E_{Tx-amp}(l, d) \tag{1}$$

$$E_{Tx}(l, d) = \{\{ \begin{matrix} llE_{\text{elec}} * l + E_{fs} * l * d^2 & if\, d < d_0 \\ E_{\text{elec}} * l + E_{mp} * l * d^4 & if\, d \geq d_0. \end{matrix} \tag{2}$$

$$d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}} \tag{3}$$

The definitions for E_{elec}, E_{mp} and E_{fs} are given in Table 1. The reception energy $E_{Rx}(l, d)$ for l bits is defined in Formula 4

$$E_{Rx}(l, d) = E_{Rx-elec}(l) = E_{\text{elec}} * l \tag{4}$$

The CHs nodes aggregate data using an energy represented as E_{DA} for transmission to the BS, as indicated in Table 1.

3.2 The optimum number of clusters (K_{opt})

The K_{opt} (see Formula 5) defined by Heinzelman et al. [50] is computed according to several network and radio parameters (see Table 1). The demonstration is clearly outlined in [50]. The deployment of the network is on a surface of M*M m^2, N is the initial number of nodes and d_{toCH} is the average distance of the CHs from the BS.

$$K_{\text{opt}} = \frac{\sqrt{N}}{\sqrt{2\Pi}} \sqrt{\frac{E_{fs}}{E_{mp}}} \frac{M}{d_{\text{toCH}}^2} \tag{5}$$

4 Theoretical and formal basis of the DCOPA protocol

DCOPA is a distributed protocol conceived for clustering in WSN-based IoT. A competition is initiated between network nodes by calculating $T(i)$, which is treated as a timer, according to the R_{Enrg} and the DistBS for self-selection of CHs in a current round. $T(i) \in]0, \tau - \delta[$, is less than the time allocated to the CH election period, which is (τ). $\delta \in]0, 0.1[$, is a very small time that ensures that a node does not declare itself as a CH outside the CH designation period. DCOPA is scheduled in two phases. In the setup phase, each node decrements its $T(i)$ (see Formula 6) at the beginning of each round. If $T(i)$ is equal to zero, the node decides to act as a CH and broadcasts an *ADV_CH* on an RC radius to its neighbors, who then drop their candidacy for the CH position and look forward to further *ADV_CH* from other CHs. The steady state phase consists of three periods: (i) Normal nodes send acknowledgment control messages to the nearest CH, (ii) CHs broadcast a TDMA (Time division multiple access) calendar for cluster nodes to transmit data messages. This time period is used for data routing inside the cluster and (iii) the CHs aggregate the data and forward it to the BS using the MAC CSMA (carrier-sense multiple access media access control) protocol. Table 2 shows the variables used in the description of $T(i)$.

$$T(i) = \begin{cases} (\alpha E_i + \beta D_i)(\tau - \delta) & \textit{if } i \in G \\ \tau - \delta & \text{otherwise} \end{cases} \tag{6}$$

$$\alpha + \beta = 1 \tag{7}$$

The following points should be noted: E_i (Formula 8) and D_i (Formula 9) are defined after the normalization procedure.

$$E_i = \left(\frac{E_{Max} - Er_i}{E_{Max}} \right) \tag{8}$$

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 R_{Enrg} 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
N_{init}	The initial number of nodes
G	The set of nodes which were not *CHs* during the previous ($1/P$) rounds, P=($K_{\mathrm{opt}}/N_{init}$)

$$D_i = \left(\frac{d_{itoBS} - d_{MintoBS}}{d_{MaxtoBS} - d_{MintoBS}} \right) \tag{9}$$

$$0 < \alpha \frac{E_{Max} - Er_i}{E_{Max}} + \beta \frac{d_{itoBS} - d_{MintoBS}}{d_{MaxtoBS} - d_{MintoBS}} \leq 1 \tag{10}$$

Formula (10) is verified and demonstrated in [6].

5 Contribution: the unequal-DCOPA protocol

5.1 DCOPA protocol and equal clustering limitation

A particular drawback of the DCOPA protocol is the restriction of using the same RC for all CHs, as exemplified in Fig. 1. This is a drawback due to the dissimilar R_{Enrg} between CHs and their DistBS. Broadcasting an advertising message (*ADV_CH*) by a CH using a RC that is the identical for all nodes, whatever their amount of energy, and whatever their distance from the BS, can never be an advantage. This equality of RC across CHs, even though they differ in criteria such as energy and DistBS, results in premature failures of certain CHs. As a consequence, nodes with low energy may not be in a position to efficiently cover their clustering area on the one hand and have a large number of nodes and thus a large number of control messages and data on the other. In addition, nodes located far from the BS may have the same intracluster activities as a CH that is very close to the BS and in addition have a long transmission distance to the BS once the data are aggregated, which inevitably leads to their accelerated failures compared to other CHs in the network.

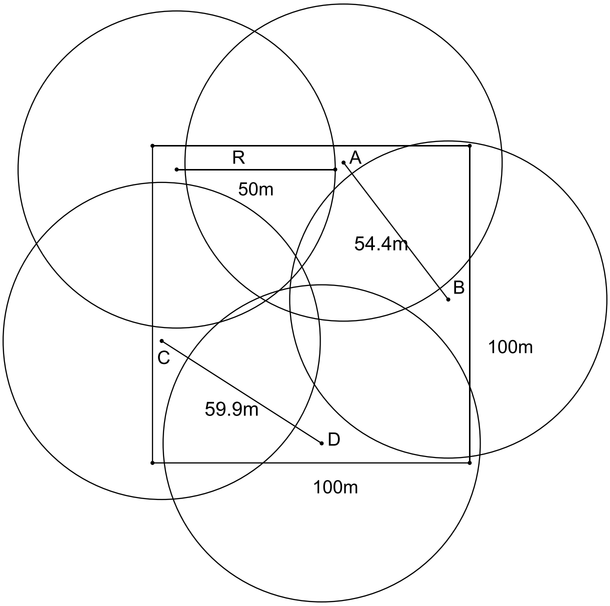


Fig. 1 RC used in DCOPA protocol [6]

5.2 Contribution description

To avoid the inequity of assigning the same RC to the CHs despite their very different capabilities, it would be appropriate to take into account the differences in R_{Enrg} and DistBS when broadcasting the *ADV_CH* by the CHs in the network. The present contribution is a continuation of the work undertaken in [10], which establishes the conceptual and methodological basis for this proposal. It could be envisaged to define variable or adaptive RC based on these criteria, enabling a more efficient allocation of resources and better adaptation to the context of each CH. An adaptive approach that dynamically adjusts RC according to the individual characteristics of CHs could help improve mortality rates, node and network energy and network lifetime. We have modeled the problem as a multicriteria optimization system, where we attempt to find the ARC taking into account two different criteria. We assign weights to the two criteria to define their degree of relative importance. These weights can be defined according to the specific needs of the application. For example, if energy conservation is a top priority, a higher weight can be assigned to R_{Enrg} than to DistBS. The weighted sum approach is chosen to combine the two criteria into a single objective function. The objective function would take the form of a linear combination of the two criteria, where each criterion is multiplied by its corresponding weight. Using this approach, DCOPA can make more intelligent and adaptive decisions regarding the calculation of RC by each CH.

5.3 UDCOPA protocol architecture

UDCOPA consists of two phases, the setup phase and the steady state phase. In the setup phase, each node begins to decrement its $T(i)$ (see Formula 6) at the starting point of each round. If $T(i)$ reaches zero, the node declares itself as CH and diffuses an *ADV_CH* for its neighbors who are within an ARC, see Formula 13, instead of the RC implemented in DCOPA, which is identical for all CHs and across the whole lifetime of the network. Nodes receiving this *ADV_CH* message drop their applications for the CH role and wait for other solicitations by *ADV_CH* from additional CHs if it is included in their ARC. The steady state phase is exactly the same as in the DCOPA protocol.

5.3.1 Adaptive radius of clustering

Before the *ADV_CH* message is diffused by a CH, the ARC is calculated using the weighted sum, which is a method applied in multicriteria analysis, of the two criteria of R_{Enrg} and DistBS, assigning weights that will be defined according to the WSN-based IoT application or user needs. The ARC takes the maximum value equal to RC which is calculated, specified and demonstrated in [6]. RC is calculated as a function of the surface area of the monitoring zone and the optimum number of culsters $K = K_{\text{opt}}$ (see Formula 5).

$$RC = \frac{2M}{\sqrt{\Pi K}} \tag{11}$$

ARC will be maximized if the R_{Enrg} is maximized and the *DistBS* is maximized. This means that we will get clusters with a large radius once the CH's R_{Enrg} and DistBS are large. As a result, nodes with a very small amount of energy will pose the problem of obtaining a very small ARC that tends toward zero, which is a real drawback, as once the nodes in the network lose a large part of their energy, we will have a very large number of CHs communicating directly to the BS. Consequently, the network can lose a very large number of nodes, which has an impact on its lifetime. To overcome this drawback, we have set a minimum value for this radius, called *RC_Min*, given in Formula 12.

$$RC\_Min = (RC * F_R) = \left(\frac{2M}{\sqrt{\Pi K}} * F_R\right) \tag{12}$$

F_R represents the Factor of Reduction of the chosen RC, considered as a percentage (examples: 25%, 30%, 50%, etc.).

Table 3 Parameters of the Formula 13

Parameters	Meaning
$d_{CH_i toBS}$	The distance of the CH_i from the BS
E_{CH_i}	The R_{Enrg} of the CH_i
RC	The Radius of Clustering (see Formula 11)
RC_Min	The minimum value of the RC (see Formula 12)
F_R	The Factor of Reduction
θ	The weight of energy criterion
ω	The weight of the distance criterion

5.3.2 Calculation of ARC in the UDCOPA protocol

The ARC on which a CH_i sends its ADV_CH message is determined by Formula 13, explained in what follows.

$$\begin{aligned} ARC_{CH_i} &= ((\theta E_{CH_i} + \omega D_{CH_i})(RC - RC\_Min)) + RC\_Min \\ &= \left((\theta E_{CH_i} + \omega D_{CH_i})\left(\left(\frac{2M}{\sqrt{\Pi K}}\right) - \left(\frac{2M}{\sqrt{\Pi K}} * F_R\right)\right)\right) + \left(\frac{2M}{\sqrt{\Pi K}} * F_R\right) \\ &= \left((\theta E_{CH_i} + \omega D_{CH_i})\left(\frac{2M}{\sqrt{\Pi K}}(1 - F_R)\right)\right) + \left(\frac{2M}{\sqrt{\Pi K}} * F_R\right) \end{aligned} \tag{13}$$

$$\theta + \omega = 1 \tag{14}$$

The parameters used in Formula 13 are given in Table 3. Following the normalization process, E_{CH_i} and D_{CH_i} are defined as follows: (the variables $d_{MaxtoBS}$, $d_{MintoBS}$ and E_{Max} are described in Table 2):

$$E_{CH_i} = \left(\frac{Er_{CH_i}}{E_{Max}}\right) \tag{15}$$

$$D_{CH_i} = \left(\frac{d_{CH_i toBS} - d_{MintoBS}}{d_{MaxtoBS} - d_{MintoBS}}\right) \tag{16}$$

$$0 < \theta \frac{Er_{CH_i}}{E_{Max}} + \omega \frac{d_{CH_i toBS} - d_{MintoBS}}{d_{MaxtoBS} - d_{MintoBS}} \leq 1 \tag{17}$$

Formula 17 is verified as follows:

- The primary objective is to maximize the distance from a CH_i to the BS, denoted as $d_{CH_i toBS}$.

$$d_{MintoBS} \leq d_{CH_i toBS} \leq d_{MaxtoBS} \tag{18}$$

$$0 \le d_{CH_i toBS} - d_{MintoBS} \le d_{MaxtoBS} - d_{MintoBS} \tag{19}$$

Following normalization, the result will be:

$$0 \le \frac{d_{CH_i toBS} - d_{MintoBS}}{d_{MaxtoBS} - d_{MintoBS}} \le 1 \tag{20}$$

- The second objective is to maximize the R_{Enrg}, for a CH_i, represented as Er_{CH_i}.

$$0 < Er_{CH_i} \le E_{Max} \tag{21}$$

Following normalization, the result will be:

$$0 < \frac{Er_{CH_i}}{E_{Max}} \le 1 \quad \Rightarrow \quad 0 < E_{CH_i} \le 1 \tag{22}$$

Using Formulas 14, 20, and 22, the verification of Formula 17 is completed. Under conditions where $E_{CH_i} \to E_{\mathrm{Max}}$ and $d_{CH_i\mathrm{toBS}} \to d_{\mathrm{MaxtoBS}}$, the (ARC) converges to the RC. This indicates that when the energy resource of a CH is high and it is located at a high distance from the BS, the ARC mechanism favors a configuration that maximizes the number of member nodes with the advertisement of its election over a sizeable radius. On the other hand, when the residual energy of a CH decreases ($E_{CH_i} \to 0$) and its DistBS is minimal ($d_{CH_i\mathrm{toBS}} \to d_{\mathrm{MintoBS}}$), the ARC converges toward (*RC_Min*) by favoring a minimal number of member nodes. This underlines the adaptability of the protocol, dynamically adjusting the RC according to energy constraints and DistBS.

6 UDCOPA: performance evaluation

6.1 Simulation assumptions

In our simulations, we made several very important assumptions about the features of the BS and the network nodes. We assumed that the BS has no energy constraints, allowing it to operate without any power restrictions. Nodes, however, are equipped with storage batteries that cannot be recharged or renewed. In addition, nodes are not able to move, nor do they have the technological equipment needed to know their positions. Finally, nodes have the ability to adapt their transmission range according to their distance from the receiver(s) and will stop working if, and only if, their energy is fully depleted.

6.2 Simulation environment

The coverage area is a square of side *M* meters and area M*M m^2. N nodes are randomly and uniformly dispatched. All nodes are initially supplied with an equal amount of energy. Two types of messages are exchanged: control messages, which are used to structure the network, and data messages, which include the data

Table 4 Criterion weights for both cases

Case	Function	Criterion	Value	$F_R(\%)$	Description
1	$T(i)$	E_i	$\alpha = 0.3$	$(1/3) = 33\%$	$R_{\text{Enrg}}(N(i))$
		D_i	$\beta = 0.7$		$DistBS(N(i))$
	ARC_{CH_i}	E_{CH_i}	$\theta = 0.3$		$R_{\text{Enrg}}(CH_i)$
		D_{CH_i}	$\omega = 0.7$		$DistBS(CH_i)$
2	$T(i)$	E_i	$\alpha = 0.5$	$(1/2) = 50\%$	$R_{\text{Enrg}}(N(i))$
		D_i	$\beta = 0.5$		$DistBS(N(i))$
	ARC_{CH_i}	E_{CH_i}	$\theta = 0.3$		$R_{\text{Enrg}}(CH_i)$
		D_{CH_i}	$\omega = 0.7$		$DistBS(CH_i)$

Table 5 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
$K = K_{\text{opt}}$	K_{opt}, see Formula 5	Optimum clusters number [50]

collected by the nodes for the specific application. Simulations are carried out using MATLAB[1] to measure the protocol's efficiency. The BS, in charge of data collection and processing, is outside the area under surveillance. The weights of the criteria, used in $T(i)$ and ARC_{CH_i}, that we have chosen are listed in the Table 4. Table 5 defines the utilized parameters. We used a computer equipped with an Intel(R) Core(TM) i5-10400 CPU running at a frequency of 2.90 GHz. This processor, from the 10th generation Intel Core family, offers robust performance for computationally intensive tasks, which are essential for running complex simulations. In addition, the computer is equipped with 16 GB of RAM memory, enabling it to efficiently manage large data volumes and run several simulations in parallel without compromising performance.

We have chosen to compare our UDCOPA protocol, which is an improvement on the DCOPA protocol, with LEACH for several reasons. Firstly, LEACH is widely recognized in the literature as the leading clustering protocol for WSN. It is often used as a baseline reference for the evaluation of new protocols, due to its reputation as a distributed and probabilistic protocol. Secondly, in addition to LEACH, we also carried out a comparison with DCOPA, which is a more recent protocol that relies

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

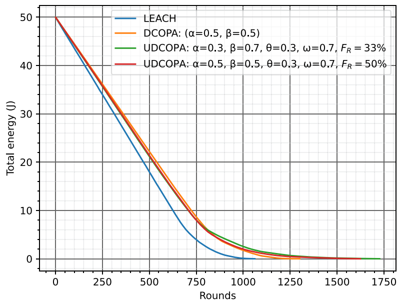
Table 6 Lifetime metrics results

Protocols	FND	HND	LND
LEACH	605	771	1062
DCOPA ($\alpha = 0.5, \beta = 0.5$)	629	858	1301
UDCOPA ($\alpha = 0.3, \beta = 0.7, \theta = 0.3, \omega = 0.7, F_R = 33\%$)	419	782	1727
UDCOPA ($\alpha = 0.5, \beta = 0.5, \theta = 0.3, \omega = 0.7, F_R = 50\%$)	485	814	1624

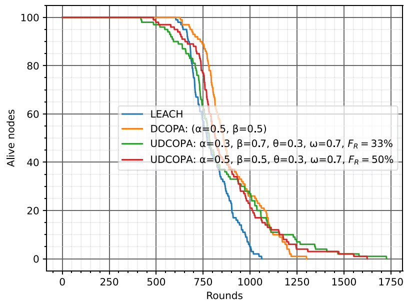
on multicriteria aggregation for CH election. DCOPA was compared with other protocols such as TB-LEACH and LEACH-MAC and demonstrated better performance than the latter. So, by showing that our UDCOPA protocol outperforms DCOPA, we can indirectly conclude that UDCOPA is superior not only to LEACH, but also to TB-LEACH [51] and LEACH-MAC [52].

6.3 Network nodes mortality rate

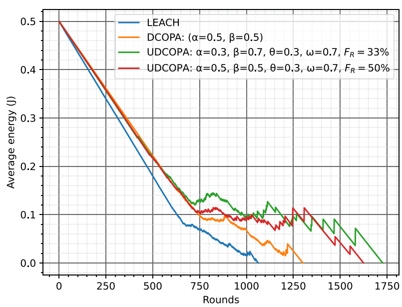
Careful analysis of Fig. 2b, clearly reveals that the *LEACH* protocol exhibits more severe mortality compared to the UDCOPA and DCOPA protocols. Initially, the



(a) Total network energy as a function of rounds



(b) Number of alive nodes as a function of rounds



(c) Average residual energy per node as a function of rounds

Fig. 2 Energy performance and mortality rate (100 nodes)

UDCOPA protocol, in both simulated cases, and the DCOPA protocol rival each other in node loss rate, with approximately equivalent performance until around the 1100 round. From this round onwards, DCOPA shows a more accelerated and severe rate of mortality than UDCOPA, leading to total node extinction in round 1301. In contrast, the UDCOPA protocol maintains a more stable loss with a less accelerated rate than DCOPA, resulting in prolonged resistance up to round 1727 for Case (1): (UDCOPA: $\alpha = 0.3, \beta = 0.3, \theta = 0.3, \omega = 0.7, F_R = 33\%$) and round 1624 for Case (2): (UDCOPA: $\alpha = 0.5, \beta = 0.5, \theta = 0.3, \omega = 0.7, F_R = 50\%$). It is important to state that the UDCOPA protocol is characterized by its ability to maintain a more stable node loss than DCOPA. This makes it more resilient to sudden decreases in the number of nodes in the network. The lifetime metrics selected are given in Table 6. These are the first node dead (FND), the half nodes dead (HND) and the last node dead (LND).

6.3.1 Interpretation 1

It is important to emphasize that the UDCOPA protocol stands out for its superior performance in terms of node failure rate, with a steady and uniform decrease in the number of nodes over the lifetime of the network. UDCOPA shows no sharp drop-off or significant failures over a small number of rounds. This indicates highly improved and balanced node energy management. The nature of the radius ARC is behind the improvement in node and network lifetime metrics. The radius is characterized by its contextual awareness of a CH's criteria. It is calculated on the basis of an objective function based on multicriteria optimization (aggregation) that combines two criteria: the CH's energy capacity and its DistBS, taking into account their respective weights according to their degree of importance in the application or network context. As a result, the CHs in the network will have different ARC depending on their individual performance. This approach creates an unequal clustering where CHs with high energy and long DistBS get a large ARC. Inversely, CHs build clusters on a very small ARC to avoid long and energy-intensive communications, which could put an end to their energies. This property can be described as a distribution of the number of nodes or the load according to the individual capacity of each CH. In summary, ARC optimizes the energy use of nodes and CHs, ensuring a balanced consumption and distribution of normal nodes over all CHs in a given round.

6.4 Total network energy

Figure 2a shows the results for total network energy as a function of rounds. The results clearly show that UDCOPA is characterized by better energy conservation than the DCOPA and LEACH protocols. Total energy management for UDCOPA and DCOPA is particularly adjacent, indicating almost identical energy management up to around round (750). However, from round (750) on, the UDCOPA protocol distinguishes itself by showing better energy conservation than the DCOPA protocol, and thus maintains stability up to round (1727) for Case (1) and (1624) for Case (2) (see Table 4).

6.4.1 Interpretation 2

The performance reported regarding the total network energy indicates that UDCOPA and DCOPA significantly outperform the LEACH protocol, mainly due to its many drawbacks, notably the RC covering the entire network. This accelerates the exhausting process of CHs and nodes. The energy efficiency of UDCOPA compared with DCOPA lies in the fact that when nodes lose energy, UDCOPA compensates by reducing the radius according to R_{Enrg} and DistBS. This strategy distributes the load between CHs and avoids long-distance communications. In contrast, the DCOPA protocol continues to operate with the same RC radius despite the energy loss of the nodes and does not take their DistBS into account. The UDCOPA protocol achieves energy balancing and stability, demonstrating the success of the radius adaptation mechanism of ARC, which is sensitive to variations in DistBS and CH energy. On the other hand, the DCOPA protocol shows a certain limitation due to the constant RC, which is not sensitive to the criteria of a CH, and to the decrease in the R_{Enrg} of the nodes over the rounds.

6.5 Average residual energy per node

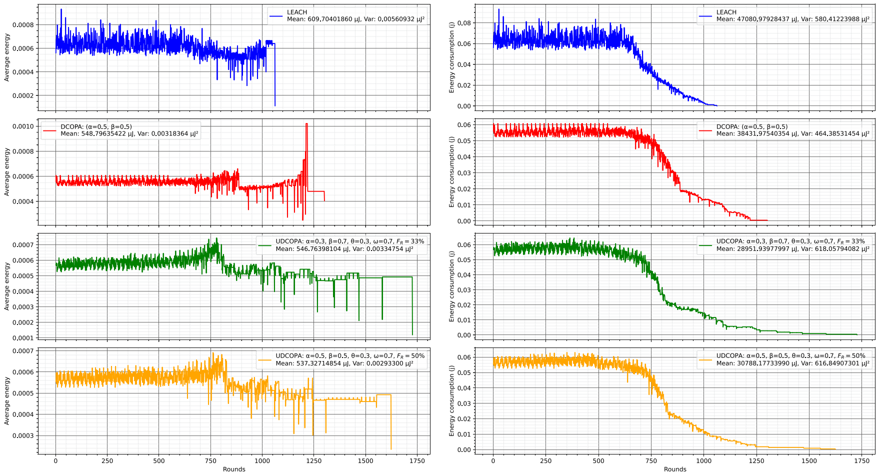
The results for average R_{Enrg} per node as a function of rounds are shown in Fig. 2c. The LEACH protocol shows a somewhat abrupt degradation of its average R_{Enrg} per node compared with other protocols. By contrast, the DCOPA and UDCOPA protocols show very similar values up to the round (700). Beyond this stage, the DCOPA protocol experiences a severe decrease in the average energy of network nodes [from round (750) with (0.1 J) to round (1301) with (0J)], while the UDCOPA protocol maintains almost linear stability over a significant number of rounds (from round 750 with (0.12 J) to around (1500) (0.09–0.06 J)). From this point on, a decay is observed until the entire network is lost.

6.5.1 Interpretation 3

The drastic decrease in the average R_{Enrg} of the nodes observed in the LEACH protocol can be attributed to several drawbacks known in the literature. These include the random election of CHs, the lack of consideration of the energy factor in the CH election process, the variability of the number of CHs in each round and the high RC covering the entire network. These factors are directly associated with a considerable reduction in average node energy from one round to the next. On the other hand, the DCOPA and UDCOPA protocols demonstrate very adjacent performance up to a certain round, approximately at (750). From this point on, the nodes start to lose energy significantly. In the case of DCOPA, nodes maintain their RC as initially defined, while for UDCOPA, nodes adjust their RC to form a significantly reduced adjusted radius (ARC). This very positive adaptation has a significant impact on

Table 7 Variance and mean of average energy consumption per node

Protocol	Variance (μJ^2)	Mean (μJ)
LEACH	0.0056	609.7040
DCOPA	0.0032	548.7964
UDCOPA Case (1)	0.0033	546.7640
UDCOPA Case (2)	0.0029	537.3271



(a) Average energy consumed by a node as a function of rounds

(b) Average network energy consumption as a function of rounds

Fig. 3 Average energy consumption per node and for the entire network (100 nodes)

energy conservation, as the adjusted ARC allows CHs to communicate over optimal distances for cluster construction. Similarly, normal nodes transmit their data over distances smaller than the RC of DCOPA, thus minimizing the energy required for communication.

6.6 Average energy consumption per node

The results for average energy consumption per node as a function of rounds are plotted in Fig. 3a and the results summarized in Table 7. We will carry out an in-depth analysis of the variances and means of four distinct representations, with the aim of comparing them rigorously. Analysis of the findings shows significant differences between the LEACH, DCOPA, UDCOPA Case (1) and UDCOPA Case (2) protocols in terms of the variance and mean of the average energy consumed per node as a function of rounds. When we focus on the variance values, which measure the dispersion of the data in relation to the mean, we see

Table 8 Variance and mean of energy consumption for the entire network

Protocol	Variance (μJ^2)	Mean (μJ)
LEACH	580.4122	47080.9793
DCOPA	464.3853	38431.9754
UDCOPA Case (1)	618.0579	28951.9398
UDCOPA Case (2)	616.8491	30788.1773

that LEACH has the highest variance (0.0056), followed by DCOPA (0.0032), UDCOPA Case (1) (0.0033) and finally UDCOPA Case (2) (0.0029). This difference in dispersion indicates that LEACH has greater variability in energy consumption than the other protocols. Concerning means, which represent the central tendency of the data, we observe that LEACH has the highest mean (609.7), followed by DCOPA (548.8), UDCOPA Case (1) (546.8) and UDCOPA Case (2) (537.3). This indicates that LEACH has a higher average energy consumption per node than other protocols. In summary, although LEACH has a higher average energy consumption, it is important to take into account the variability measured by the variance. DCOPA, UDCOPA Case (1) and UDCOPA Case (2) show lower levels of variability, reflecting more stable energy consumption. These results underline the importance of considering both central tendency and data dispersion for a comprehensive evaluation of protocol performance in the context of energy consumption.

6.7 Average energy consumption by the complete network

Examining the variability of the average energy consumption of the whole network as a function of rounds, see Fig. 3b. The results, as summarized in Table 8, reveal significant differences between protocols. The variances, measuring the dispersion of values around the mean, are more pronounced for UDCOPA Case (1) (618.06) and UDCOPA Case (2) (616.85) than for LEACH (580.41) and DCOPA (464.39). This variability indicates a greater fluctuation in node energy consumption for the UDCOPA protocols. In terms of means, LEACH has the highest value (47, 080.98), indicating high average energy consumption compared with the other protocols. DCOPA has a lower average (38, 431.98), followed by UDCOPA Case (2) (30, 788.18) and UDCOPA Case (1) (28, 951.94).

6.7.1 Interpretation 4

LEACH takes last rank because of certain limitations observed in its mode of operation. LEACH uses a fixed RC for cluster formation, and this radius covers the entire network region. In addition, LEACH does not take node-specific criteria into account when selecting nodes as CHs. This approach can lead to problems, particularly when nodes with energy very close to zero are elected as CHs. Not considering node criteria can lead to the selection of CHs less suited to their context, which can result in premature failure of these nodes and negatively affect network energy performance.

In second-to-last place, we have DCOPA, which offers an improvement on LEACH by aggregating node criteria to compete for the CH role. However, DCOPA has an important limitation in terms of its fixed broadcast radius RC. The RC is used to define the solicitation message range of normal nodes in order to include them in a CH's cluster. Unfortunately, this fixed RC remains constant throughout the lifetime of the network, which can impact the results of average energy consumption per node and per network over the rounds. This limitation becomes more of a factor, especially when network nodes have lost a significant amount of their energy. In the first place, we find UDCOPA, which represents an improved version of DCOPA by introducing significant adjustments. UDCOPA takes into account nodes' criteria in their election as CHs, offering a more adaptive approach to the various conditions of the network. Unlike DCOPA, UDCOPA introduces the notion of the ARC, which is dynamic and adjustable according to the CH's performance. This innovation enables a CH to adapt its communication radius to its specific context, thus optimizing its communications and energy consumption.

The flexibility provided by the ARC contributes to better adaptation to changing network conditions, thus enhancing the overall energy performance of the UDCOPA protocol. It is important to note that the number of rounds has a significant influence on the results obtained. Greater variability in the number of rounds can lead to increased data dispersion, thus impacting the variance of observations. In our context, the disparity in the number of rounds between UDCOPA (1727), DCOPA (1301) and LEACH (1062) could explain the higher variance observed in UDCOPA. This difference is explained by the fact that an increase in the number of rounds leads to a more extensive collection of observations, potentially introducing greater variability into the data analyzed.

7 Analysis of clusters load balancing and distribution of CHs and their member nodes

In what follows, we will look at metrics that are of crucial importance in WSN-based IoT network clustering protocols. These metrics play a key role in performance evaluation. The metrics examined include distribution of nodes, CHs and clusters. These metrics evaluate how nodes and CHs are distributed in clusters, as well as the shape and geographical extent of the clusters formed (the geographical configuration). These objectives are developed with a view to designing clustering protocols that minimize energy consumption by guaranteeing a balanced distribution of nodes and geographical space (area). The analysis of the following points relates to the load balancing metrics of a clustering protocol.

- Distribution of nodes and CHs: evaluating how nodes are distributed in their clusters and the distribution of CHs in the network is crucial. An unequal distribution of nodes can lead to energy overload for some CHs, compromising the life of the network and CHs overloaded with member nodes.

- Shape and geographical extent of clusters: the shape of the clusters formed and their geographical extent is essential for energy efficiency. Clusters with a well-balanced surface area enable more efficient energy management, reducing the need for long-distance communication. A very large cluster surface area implies the possibility of having so many member nodes far removed from their CHs.

These metrics aim to highlight the development of energy-efficient clustering protocols to meet the challenges posed by WSN-based IoT networks and thus contribute to optimizing their lifetime.

7.1 Specific metrics studied

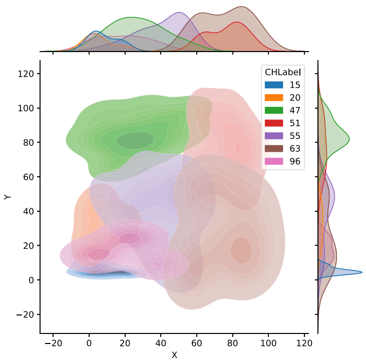
The following metrics are applied to four case studies, namely the LEACH, DCOPA and UDCOPA protocols with two parameter variations.

- Bidimensional distribution of formed clusters with density contours.
- Bidimensional distribution of nodes within each formed cluster.
- Distribution of nodes and CHs within the formed clusters, including the RC for each CH.
- Distribution of CHs with their RC.
- Elected CHs along with their corresponding RC, the count of their nodes and the DistBS.

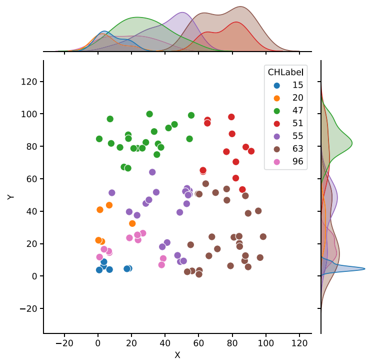
This analysis aims to compare the performance of these protocols, highlighting the distribution and load balancing that are extremely important for any clustering protocol. Our load balancing model takes into account several key aspects for optimizing energy efficiency in IoT-based WSN. The distribution of nodes and CHs is analyzed to ensure even distribution across the network. The shape and geographical extent of clusters are considered. We use bidimensional distributions to visualize the density contours of the clusters formed and the distribution of nodes within each cluster. In addition, our model takes into account the ARC coverage for each CHs. Compared to the load balancing model presented in [49], the focus is on optimizing the aggregation and data transmission between nodes.

7.1.1 Interpretation 5

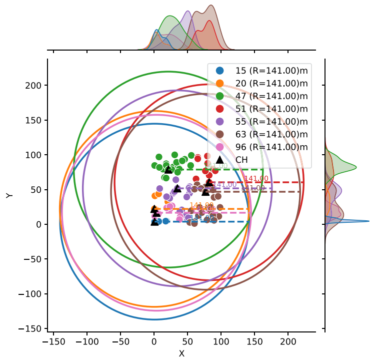
1. LEACH: LEACH protocol performance is detailed in Fig. 4. However, the clusters formed do not show satisfactory homogeneity or balancing, as illustrated in Fig. 4a, b. There are marked disparities in both cluster size and shape, with clusters of highly variable dimensions. This is attributable to the purely random nature of the CHs election. The RC, having been determined to encapsulate the entire



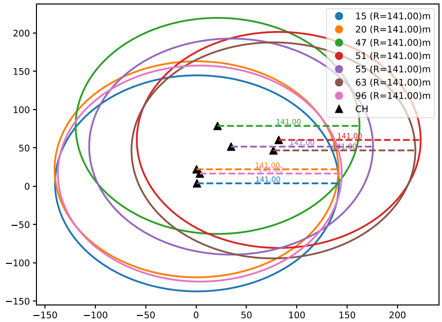
(a) Bidimensional distribution of formed clusters with density contours



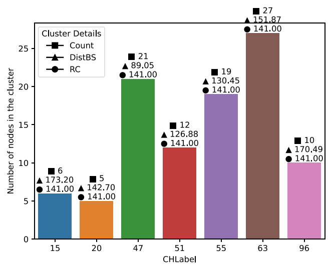
(b) Bidimensional distribution of nodes in each formed cluster



(c) Distribution of nodes and CHs within the formed clusters, including the RC for each CH

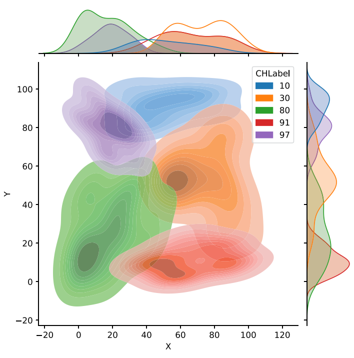


(d) Distribution of CHs with their RC

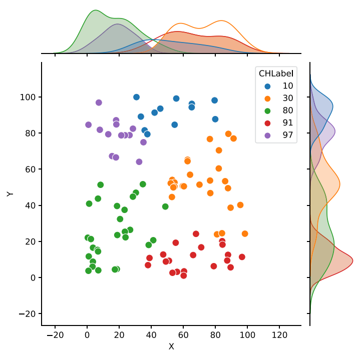


(e) Clusters metrics : Number of Nodes, Elected CH, DistBS and RC

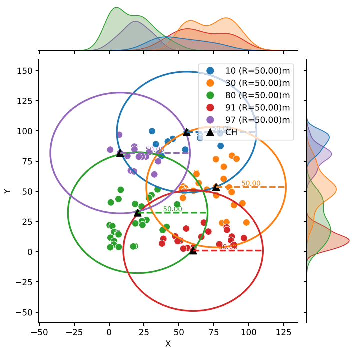
Fig. 4 Clustering metrics: distribution and load balancing in the LEACH protocol (100 nodes)



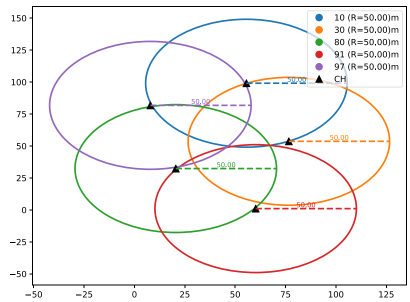
(a) Bidimensional distribution of formed clusters with density contours



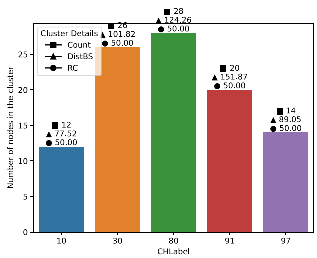
(b) Bidimensional distribution of nodes in each formed cluster



(c) Distribution of nodes and CHs within the formed clusters, including the RC for each CH



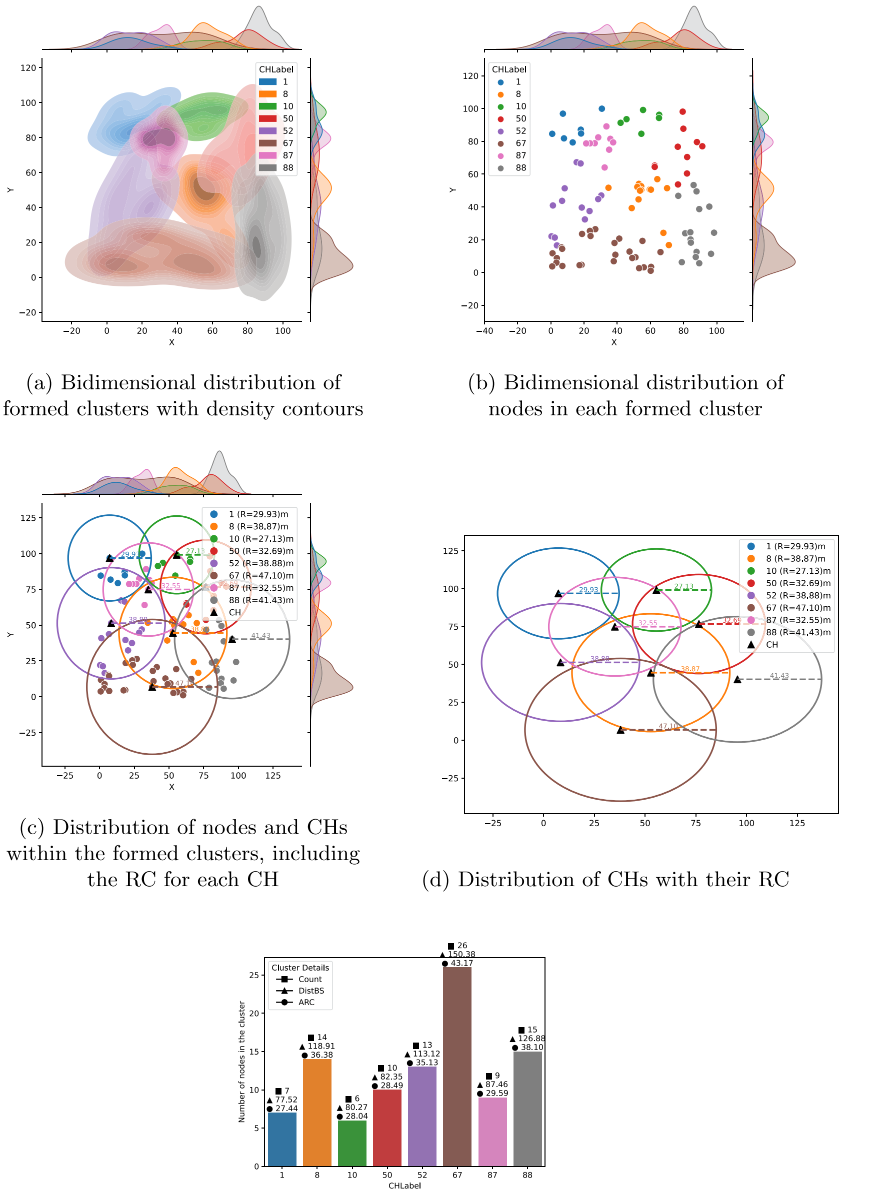
(d) Distribution of CHs with their RC



(e) Clusters metrics : Number of Nodes, Elected CH, DistBS and RC

Fig. 5 Clustering metrics: distribution and load balancing in the DCOPA protocol (($\alpha = 0.5, \beta = 0.5$), 100 nodes) network, turns out to be quite extensive. A relevant observation from Fig. 4c, d highlights the distribution property of the CHs in the network. It is clear that some CHs are very close to each other, negatively impacting their distribution. Figure 4e reinforces this observation by revealing significant disparities in terms of the number of member nodes within clusters, compromising load balancing. Taken together, these observations point to critical aspects inherent to the LEACH protocol, notably the random election of CHs, the absence of criteria to guide CHs selection and the full coverage of the network by the RC. These shortcomings have adverse consequences on the distribution of both CHs and nodes, directly compromising the guarantee of efficient load balancing.

2. DCOPA ($\alpha = 0.5, \beta = 0.5$): The performance of the DCOPA protocol is illustrated in Fig. 5. DCOPA is distinguished by a more efficient and geographically balanced distribution of clusters, as shown in Fig. 5a, b. Figure 5c, d highlights a more even distribution of CHs in the network, with a defined and respected minimum distance between each pair of CHs. This feature is crucial to ensure a balanced distribution of CHs throughout the network. Furthermore, Fig. 5e confirms that the DCOPA protocol excels in terms of load balancing, with clusters featuring a more uniform distribution of the number of member nodes (12, 14, 20, 26, 28). The DCOPA protocol outperforms the LEACH protocol in terms of CH distribution and more efficient load balancing, due to an RC calculation approach and competition between nodes for the role of CHs based on a multicriteria approach involving the R_{Enrg} of nodes as well as their DistBS.
3. UDCOPA ($\alpha = 0.3, \beta = 0.7, \theta = 0.3, \omega = 0.7, F_R = 33\%$): The performance of the UDCOPA protocol, with the specified parameters, is shown in Fig. 6. This approach proposes a new clustering method known as unequal clustering. Cluster sizes, as shown in Fig. 6a, b, vary, but in a homogeneous way. Looking at these results, we see large clusters in regions far from the BS, medium-sized clusters in the center and small clusters in regions closer to the BS. This feature is not constant across all rounds, as two criteria define the RC. In our case, the energy is almost the same considering we are in the first round, so it is the DistBS criterion that predominates. Figure 6c, d highlights a homogeneous distribution of CHs, with variable RC, as explicitly illustrated. The CHs respect their minimum distances, calculated according to the performance of each node. Figure 6d highlights a crucial property: the RC differs from one CH to another, depending on its performance, thus ensuring the formation of large or small clusters. In this context, the UDCOPA protocol guarantees real load balancing, despite disparities in terms of number of nodes and geographical extent, these factors being controlled by the individual performance of each node. UDCOPA's advanced performance results from the introduction of a new notion: ARC to ensure unequal clustering, sensitive to the context of a CH node's energy performance and its DistBS.
4. UDCOPA ($\alpha = 0.5, \beta = 0.5, \theta = 0.3, \omega = 0.7, F_R = 50\%$): The performance of the UDCOPA protocol, configured with the specified parameters, is shown in Fig. 7. A further run of the UDCOPA protocol was carried out, this time with a variation in parameters. Figure 7a–e provides an overview of the distribution and load balancing performance resulting from this new configuration, particularly with regard to the minimum ARC value (50%) and the weights assigned to the criteria



(a) Bidimensional distribution of formed clusters with density contours

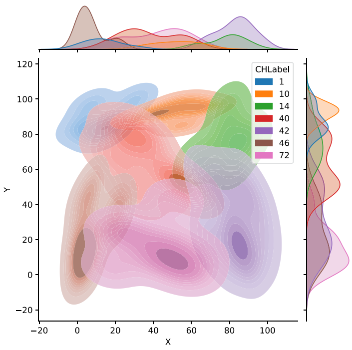
(b) Bidimensional distribution of nodes in each formed cluster

(c) Distribution of nodes and CHs within the formed clusters, including the RC for each CH

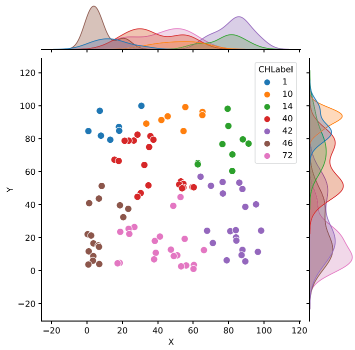
(d) Distribution of CHs with their RC

(e) Clusters metrics : Number of Nodes, Elected CH, DistBS and RC

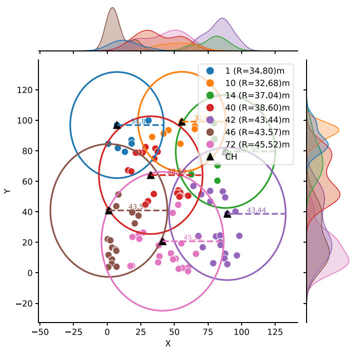
Fig. 6 Clustering metrics: distribution and load balancing in the UDCOPA protocol [($\alpha = 0.3, \beta = 0.7, \theta = 0.3, \omega = 0.7, F_R = 33\%$), 100 nodes]



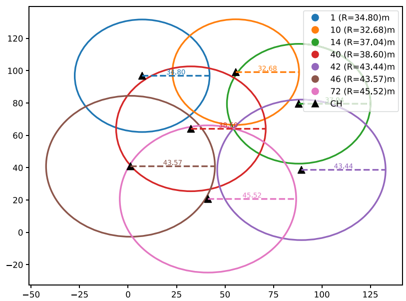
(a) Bidimensional distribution of formed clusters with density contours



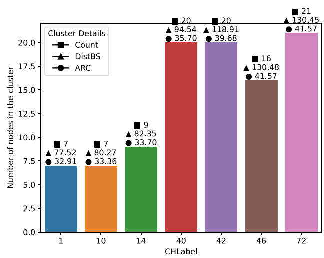
(b) Bidimensional distribution of nodes in each formed cluster



(c) Distribution of nodes and CHs within the formed clusters, including the RC for each CH



(d) Distribution of CHs with their RC



(e) Clusters metrics : Number of Nodes, Elected CH, DistBS and RC

Fig. 7 Clustering metrics: distribution and load balancing in the UDCOPA protocol [($\alpha = 0.5, \beta = 0.5, \theta = 0.3, \omega = 0.7, F_R = 50\%$), 100 nodes]

Table 9 K_{opt} values for different numbers of nodes

Number of nodes	K_{opt} (see formula 5)	Rounded K_{opt}
100	4.84	5
150	5.93	6
200	6.85	7
250	7.65	8
300	8.39	8
350	9.06	9
400	9.68	10
450	10.27	10
500	10.83	11

used in calculating the competition timer or ARC. Performance in this scenario is influenced by the same factors discussed in the first UDCOPA configuration.

The load balancing and distribution of CHs within the UDCOPA protocol can be described as highly efficient, or even true balancing and efficient distribution. Both properties are sensitive to the context of CH energy performance and DistBS, with RC adaptation or adjustment depending on these two criteria. This leads to the design of a new ARC aimed at assigning a load adapted to the capacity of the network CHs.

8 Scalability analysis

Scalability is of key importance in WSN-based IoT networks. Its primary aim is to ensure that the network adapts effectively and efficiently to the growing number of nodes and extended coverage space. The continuous evolution of WSN-based IoT applications requires expansion capacity without affecting network performance, particularly in terms of energy management. We undertook the realization of several scenarios with different settings for UDCOPA, compared to a single configuration for DCOPA. Our choice was to study scalability by looking at two key lifetime metrics, namely FND and LND. These two metrics were measured for several networks, ranging from 100 nodes to 500, with an increment of 50 nodes at each time. The simulation parameters are indicated in Table 5.

The value of K_{opt} varies according to several factors, see Formula 5. In our case, as we increase the number of nodes in the network, K_{opt} increases as the number of nodes increases. Table 9 shows the values of K_{opt} as a function of the number of nodes in the network.

Density is a crucial aspect of clustering. In our study, we observed an increase in node density as the number of nodes in the network increased, while the deployment area remained unchanged. Thus, the K_{opt} values, as presented in Table 9, also remain

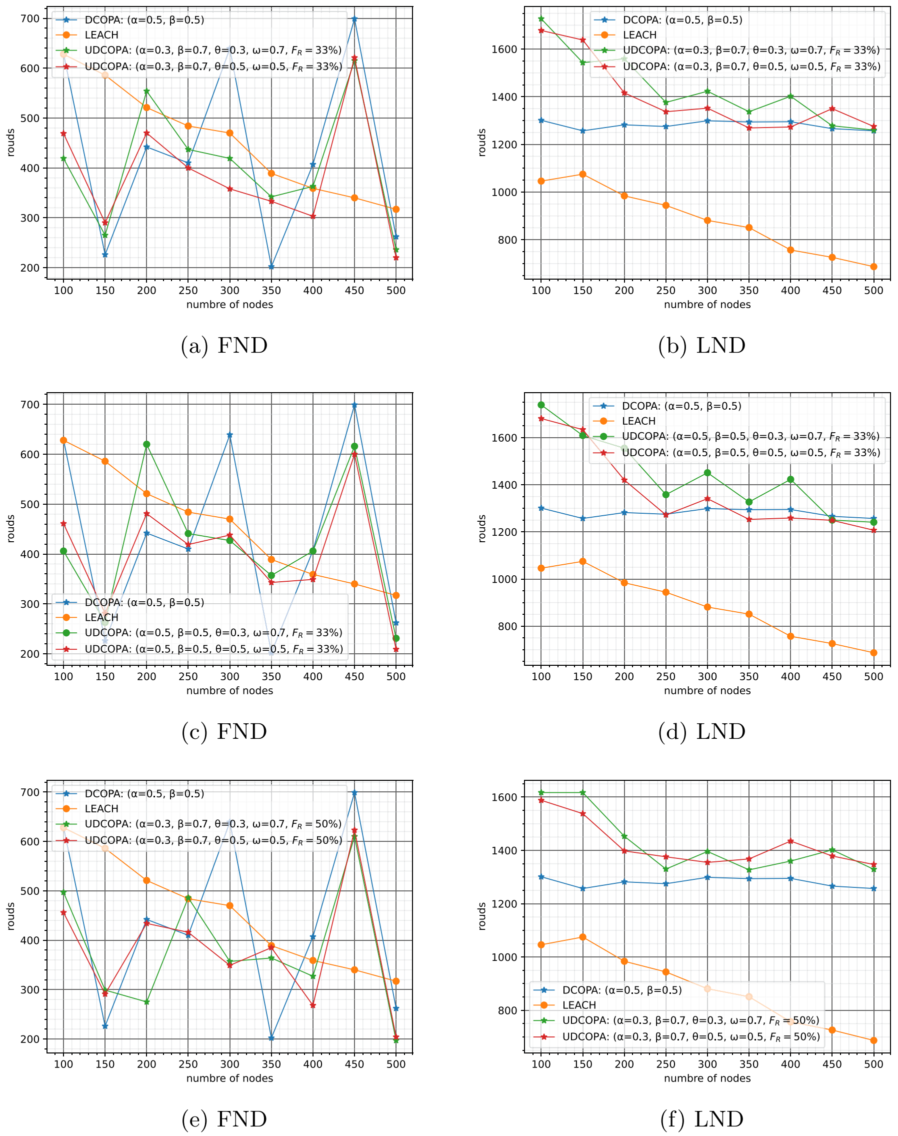


Fig. 8 Analysis of the effect of scalability on network lifetime parameters

unchanged. Indeed, density is not considered in the calculation of K_{opt}, as indicated in Formula 5.

An analysis of the impact of scalability on network lifetime parameters for the different protocols is presented in Fig. 8. We observe, through the FND metric illustrated in Fig. 8a, c, e, that the LEACH protocol is characterized by a progressive decrease in FND as the number of nodes in the network increases. This is because, in some rounds, we can have a very limited number of CHs with a large number of member nodes (in some cases, we could even have the whole network constitute a single cluster), which directly affects node mortality.

We observe fluctuations in the FND metric, with significant deviations for DCOPA and UDCOPA. This is due to the nature of the generated network, which has isolated nodes not solicited by other CHs. These nodes, not being normal nodes belonging to a given cluster, see the timer responsible for their proclamation as CHs constantly reach zero, leading them to declare themselves as CHs several times in succession. This leads to their premature exhaustion and, ultimately, the complete depletion of their energy, resulting in the rapid attainment of FND for such a network. We observed this type of behavior when simulating DCOPA [6], but by calculating the confidence interval for several networks, we obtained really good results.

With regard to the LND metric illustrated in Fig. 8b, d, f, the LEACH protocol shows very negative results toward the scalability effect, due to all the drawbacks described above. DCOPA, on the other hand, shows very stable performance despite network extensions of up to 500 nodes, due to its assignment of the CH role to nodes according to their local criteria, directly influencing energy consumption. UDCOPA, with its different configurations, reflects good results that exceed the DCOPA protocol, with a slight decrease once the number of nodes starts to increase and then shows stability. The ARC plays a central and pivotal role in UDCOPA's architecture.

9 Conclusion

UDCOPA is a distributed unequal clustering algorithm for data communications in WSN-based IoT networks that improve on the DCOPA protocol, in particular, by defining the radius within which this CH will search for the nodes that make up its cluster. Our goal is to improve the adaptability of the RC, the context-sensitivity of the criteria for a CH and the dynamic adjustment of the RC for balancing energy consumption, balancing the extent of clusters and the number of nodes allowed per cluster, and balancing the distribution of CHs and clusters in the coverage area. We realized that the same clustering algorithm is applied to CHs once they have been elected in the DCOPA protocol, without taking into consideration their disparities and differences in terms of R_{Enrg} and DistBS. To remedy this disadvantage, we have introduced a new approach, called UDCOPA, which allows each CH to establish its own RC, taking into account its R_{Enrg} and DistBS criteria. This multicriteria modeling is implemented on a weighted sum basis to ensure that clustering is optimally adapted to the specific criteria of each CH. Simulation results showed that our UDCOPA protocol outperformed both the DCOPA and LEACH protocols in terms of energy management, load balancing, scalability and network lifetime. Our contribution has significantly enhanced clustering in WSN-based IoT networks by proposing an unequal and multicriteria approach, UDCOPA, for defining the RC of CHs that will be able to make decisions for modifying their RC for a more adapted and contextual clustering. The promising results of our simulations demonstrate the effectiveness and significant improvement of our approach over comparison protocols. The clusters formed by the UDCOPA protocol with an ARC are of unequal size. Smaller clusters have a smaller ARC, and fewer member nodes are closest to the BS. On the other hand, clusters further away from the BS have a larger ARC and a larger number of member nodes. This characteristic can lead to imbalances, as the number of CHs in areas closer to the BS can increase, allowing a node to be elected as CH several times, resulting in excessive energy dissipation for some nodes close to the BS. For the most remote CHs, they will have a large number of member nodes and increased communication, which also leads to excessive energy consumption. For future work, there are plans to incorporate additional criteria aimed at enhancing ARC calculation, refining Kopt calculation and integrating density as a factor that could influence ARC.

Author Contributions F. Mir was contributed conception and design of study, methodology, software, formal analysis, analysis and interpretation of data, validation, writing—original draft and writing—review and editing. F. Meziane was involved in supervision, methodology, validation and writing—review and editing.

Data availability No datasets were generated or analyzed during the current study.

Declaration

Conflict of interest The authors declare no conflict of interest.

References

1. Adeel A et al (2019) A survey on the role of wireless sensor networks and iot in disaster management. In: Durrani T, Wang W, Forbes S (eds) Geological disaster monitoring based on sensor networks. Springer, Singapore, pp 57–66
2. Asin A, Gascon D (2012) 50 sensor applications for a smarter world. Libelium Comunicaciones Distribuidas, Tech Rep, pp 589–594
3. Jha V, Sharma R (2022) An energy efficient weighted clustering algorithm in heterogeneous wireless sensor networks. J Supercomput 78(12):14266–14293
4. Singh H, Singh D (2021) Hierarchical clustering and routing protocol to ensure scalability and reliability in large-scale wireless sensor networks. J Supercomput 77(9):10165–10183
5. Banerjee A, De SK, Majumder K, Dash D, Chattopadhyay S (2022) Construction of energy minimized WSN using GA-SAMP-MWPSO and k-mean clustering algorithm with LDCF deployment strategy. J Supercomput 78(8):11015–11050
6. Mir F, Meziane F (2023) DCOPA: a distributed clustering based on objects performances aggregation for hierarchical communications in IoT applications. Cluster Comput 26:1077–1098
7. Mir F, Meziane F (2024) Novel adaptive DCOPA using dynamic weighting for vector of performances indicators optimization of IoT networks. Expert Syst Appl 247:123212
8. Belton V, Stewart T (2002) Multiple criteria decision analysis: an integrated approach. Springer, Berlin
9. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. IEEE, p 10
10. Mir F, Meziane F (2023) Unequal clustering protocol in IoT networks based on multiple criteria processing. In: ICISS ’23. ACM, pp 132–136
11. Mir F, Bounceur A, Meziane F (2014) Regression analysis for energy and lifetime prediction in large wireless sensor networks. In: 2014 International Conference on Advanced Networking Distributed Systems and Applications. IEEE, pp 1–6

12. Farman H, Javed H, Ahmad J, Jan B, Zeeshan M (2016) Grid-based hybrid network deployment approach for energy efficient wireless sensor networks. J Sens 2016(1):2326917
13. Farman H et al (2018) Multi-criteria based zone head selection in internet of things based wireless sensor networks. Future Gener Comput Syst 87:364–371
14. Liao Y, Qi H, Li W (2012) Load-balanced clustering algorithm with distributed self-organization for wireless sensor networks. IEEE Sens J 13(5):1498–1506
15. Shivhare A, Singh VK, Kumar M (2020) Anticomplementary triangles for efficient coverage in sensor network-based IoT. IEEE Syst J 14(4):4854–4863
16. Shivhare A, Singh VK, Kumar M (2023) Event detection using the user context in sensor based IoT. Wireless Netw 29(6):2577–2589
17. Maurya MK et al (2022) Cluster based smart random walk for data aggregation in wireless sensor network. In: 2022 IEEE 15th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC). IEEE, pp 98–104
18. Yu J, Qi Y, Wang G, Guo Q, Gu X (2011) An energy-aware distributed unequal clustering protocol for wireless sensor networks. Int J Distrib Sens Netw 7(1):202145
19. Ye M, Li C, Chen G, Wu J (2005) EECS: an energy efficient clustering scheme in wireless sensor networks. In: 24th IEEE International Performance, Computing, and Communications Conference. IEEE, pp 535–540
20. Jain A (2022) Unequal clustering protocols for wireless sensor networks-taxonomy, comparison and simulation. Wireless Pers Commun 124(1):517–571
21. Jyun-Yuan C, Shanq-Jang R, Ray-Guang C, Teng-Tai H (2006) PADCP: power-aware dynamic clustering protocol for wireless sensor networks
22. Jeong J, Culler D, Oh J-H (2007) Empirical analysis of transmission power control algorithms for wireless sensor networks. In: 2007 Fourth International Conference on Networked Sensing Systems. IEEE, pp 27–34
23. Cheng H, Yang S, Cao J (2013) Dynamic genetic algorithms for the dynamic load balanced clustering problem in mobile ad hoc networks. Expert Syst Appl 40(4):1381–1392
24. Li C, Ye M, Chen G, Wu J (2005) An energy-efficient unequal clustering mechanism for wireless sensor networks. In: IEEE International Conference on Mobile Adhoc and Sensor Systems Conference. IEEE, p 8
25. Tong W, Jiyi W, He X, Jinghua Z, Munyabugingo C (2013) A cross unequal clustering routing algorithm for sensor network. Meas Sci Rev 13(4):200–205
26. Soro S, Heinzelman WB (2005) Prolonging the lifetime of wireless sensor networks via unequal clustering. In: 19th IEEE International Parallel and Distributed Processing Symposium. IEEE, p 8
27. Jain A, Reddy B (2014) Sink as cluster head: An energy efficient clustering method for wireless sensor networks. In: 2014 International Conference on Data Mining and Intelligent Computing (ICDMIC). IEEE, pp 1–6
28. Jiang C-J, Shi W-R, Tang X-L et al (2010) Energy-balanced unequal clustering protocol for wireless sensor networks. J China Univ Posts Telecommun 17(4):94–99
29. Bagci H, Yazici A (2010) An energy aware fuzzy unequal clustering algorithm for wireless sensor networks. In: International Conference on Fuzzy Systems. IEEE, pp 1–8
30. Liu P, Huang T-l, Zhou X-y, Wu G-x (2010) An improved energy efficient unequal clustering algorithm of wireless sensor network. In: 2010 International Conference on Intelligent Computing and Integrated Systems. IEEE, pp 930–933
31. Mao S, Zhao C, Zhou Z, Ye Y (2013) An improved fuzzy unequal clustering algorithm for wireless sensor network. Mobile Netw Appl 18:206–214
32. Kaur T, Kumar D (2018) Particle swarm optimization-based unequal and fault tolerant clustering protocol for wireless sensor networks. IEEE Sens J 18(11):4614–4622
33. Guiloufi ABF, Nasri N, Kachouri A (2016) An energy-efficient unequal clustering algorithm using 'Sierpinski triangle' for WSNs. Wireless Pers Commun 88:449–465
34. Arjunan S, Pothula S (2019) A survey on unequal clustering protocols in wireless sensor networks. J King Saud Univ Comput Inf Sci 31(3):304–317
35. Kim J-H, Hussain CS, Yang W-C, Kim D-S, Park M-S (2008) Produce: a probability-driven unequal clustering mechanism for wireless sensor networks. IEEE, pp 928–933
36. Yu J, Qi Y, Wang G (2011) An energy-driven unequal clustering protocol for heterogeneous wireless sensor networks. J Control Theory Appl 9:133–139

37. Lee S, Choe H, Park B, Song Y, Kim C-K (2011) LUCA: an energy-efficient unequal clustering algorithm using location information for wireless sensor networks. Wireless Pers Commun 56:715–731
38. Chen G, Li C, Ye M, Wu J (2009) An unequal cluster-based routing protocol in wireless sensor networks. Wireless Netw 15:193–207
39. Lai WK, Fan CS, Lin LY (2012) Arranging cluster sizes and transmission ranges for wireless sensor networks. Inf Sci 183(1):117–131
40. Mazumdar N, Om H (2015) Coverage-aware unequal clustering algorithm for wireless sensor networks. Procedia Comput Sci 57:660–669
41. Gajjar S, Talati A, Sarkar M, Dasgupta K (2015) FUCP: fuzzy based unequal clustering protocol for wireless sensor networks. In: 2015 39th National Systems Conference (NSC). IEEE, pp 1–6
42. Logambigai R, Kannan A (2016) Fuzzy logic based unequal clustering for wireless sensor networks. Wireless Netw 22:945–957
43. Baranidharan B, Santhi B (2016) DUCF: distributed load balancing unequal clustering in wireless sensor networks using fuzzy approach. Appl Soft Comput 40:495–506
44. Xunli F, Feiefi D (2015) Shuffled frog leaping algorithm based unequal clustering strategy for wireless sensor networks. Appl Math Inf Sci 9(3):1415–1426
45. Abo-Zahhad M, Ahmed SM, Sabor N, Sasaki S (2014) A new energy-efficient adaptive clustering protocol based on genetic algorithm for improving the lifetime and the stable period of wireless sensor networks. Int J Energy Inf Commun 5(3):47–72
46. Pravin RA et al (2024) Stochastic cluster head selection model for energy balancing in IoT enabled heterogeneous WSN. Meas Sens 35:101282
47. Sennan S, Ramasubbareddy S, Dhanaraj RK, Nayyar A, Balusamy B (2024) Energy-efficient cluster head selection in wireless sensor networks-based internet of things (IoT) using fuzzy-based Harris Hawks optimization. Telecommun Syst. https://doi.org/10.1007/s11235-024-01176-9
48. Chandrasekaran SK, Rajasekaran VA (2024) Energy-efficient cluster head using modified fuzzy logic with WOA and path selection using enhanced CSO in IoT-enabled smart agriculture systems. J Supercomput 80(8):11149–11190
49. Gupta S, Snigdh I (2022) Leveraging data aggregation algorithm in Lora networks. J Supercomput 78(15):16861–16875
50. Heinzelman WB, Chandrakasan AP, Balakrishnan H (2002) An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wireless Commun 1(4):660–670
51. Junping H, Yuhui J, Liang D (2008) A time-based cluster-head selection algorithm for LEACH. In: 2008 IEEE Symposium on Computers and Communications. IEEE, pp 1172–1176
52. Batra PK, Kant K (2016) LEACH-MAC: a new cluster head selection algorithm for wireless sensor networks. Wireless Netw 22:49–60

Authors and Affiliations

Foudil Mir[1] · Farid Meziane[2]

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

Farid Meziane
F.Meziane@derby.ac.uk

[1] Medical Computing Laboratory (LIMED), Faculty of Exact Sciences, University of Bejaia, 06000 Bejaia, Algeria

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