

Article

# Energy-Efficient Clustering Routing Protocol for Wireless Sensor Networks Based on Yellow Saddle Goatfish Algorithm

Alma Rodríguez <sup>1,2</sup>, Carolina Del-Valle-Soto <sup>1,\*</sup> and Ramiro Velázquez <sup>3</sup>

<sup>1</sup> Facultad de Ingeniería, Universidad Panamericana, Álvaro del Portillo 49, Zapopan, Jalisco 45010, Mexico; anrodriguez@up.edu.mx

<sup>2</sup> Departamento de Electrónica, Universidad de Guadalajara, CUCEI. Av. Revolución 1500, Guadalajara, Jalisco C.P 44430, Mexico

<sup>3</sup> Facultad de Ingeniería, Universidad Panamericana. Josemaría Escrivá de Balaguer 101, Aguascalientes, Aguascalientes 20290, Mexico; rvelazquez@up.edu.mx

\* Correspondence: cvalle@up.edu.mx

Received: 23 July 2020; Accepted: 3 September 2020; Published: 4 September 2020



**Abstract:** The usage of wireless sensor devices in many applications, such as in the Internet of Things and monitoring in dangerous geographical spaces, has increased in recent years. However, sensor nodes have limited power, and battery replacement is not viable in most cases. Thus, energy savings in Wireless Sensor Networks (WSNs) is the primary concern in the design of efficient communication protocols. Therefore, a novel energy-efficient clustering routing protocol for WSNs based on Yellow Saddle Goatfish Algorithm (YSGA) is proposed. The protocol is intended to intensify the network lifetime by reducing energy consumption. The network considers a base station and a set of cluster heads in its cluster structure. The number of cluster heads and the selection of optimal cluster heads is determined by the YSGA algorithm, while sensor nodes are assigned to its nearest cluster head. The cluster structure of the network is reconfigured by YSGA to ensure an optimal distribution of cluster heads and reduce the transmission distance. Experiments show competitive results and demonstrate that the proposed routing protocol minimizes the energy consumption, improves the lifetime, and prolongs the stability period of the network in comparison with the stated of the art clustering routing protocols.

**Keywords:** wireless sensor networks; yellow saddle goatfish algorithm; clustering routing protocol; metaheuristic algorithms; energy efficiency; Bio-inspired algorithms

## 1. Introduction

In recent years, there has been significant exploitation in wireless communications because they allow great flexibility for data transmission. Wireless communications provide a combination of connectivity and mobility using air as a medium [1]. Among the various existing wireless technologies, the most popular are wireless sensor networks. They require an efficient communication establishment, from the energy and computational point of view, which is guaranteed by developing optimal algorithms for communication protocols.

Wireless networks of distributed sensors are composed of a collection of electro-mechanical microdevices, scattered within a defined area. These sensors can reveal multiple information about territory. Exchanging and transmitting this data over a channel for wireless sensor networks [2]. Wireless sensors depend on the use of batteries, which generates the main problem because they contain limited energy autonomy. Thus, a base station is required to send information and process it.

Hierarchical or cluster-based routing is a well-known technique with particular advantages related to scalability and efficiency in communications. The concept of hierarchical routing has also been applied to achieve energy efficiency in WSNs. In a hierarchical architecture, high energy nodes can be used to process and send information [3]. In contrast, low energy nodes can only be used for sensor work in regions close to the target. Therefore, creating clusters and assigning particular tasks to Cluster Heads (CHs) can significantly contribute to the scalability of the system, its lifetime, and its energy efficiency. Scalability is a critical factor in WSNs, which has not been well resolved in most of the protocols due to the initial assumptions. For instance, cluster-based protocols commonly assume a single sink and cluster heads within the coverage of the sink [4]. Under those assumptions, WSNs will lack scalability, making communications highly expensive in terms of the required energy. Therefore, increasing the number of nodes or the diameter of the network will cause an overload that grows exponentially and a bottleneck consisting of only one sink that can strangle the network. Hierarchical routing is an efficient way to decrease energy consumption within the cluster and perform information aggregation. It also allows merging tasks to reduce the transmitted packet number to the sink [5]. A correct protocol must have mechanisms for saving energy in the nodes since recharging their batteries is usually tough, not possible, or even dangerous. All aspects of the nodes, from the hardware to their performed operations, will help to distribute the energy load [6]. Thanks to hierarchical routing having different roles to help us to differentiate potential phases in protocol operations, the existence of two operation modes in the sensors can foresee a higher energy expenditure. However, different transmission modes, such as cluster head and normal sensor, can be beneficial if implemented in every layer [7].

Routing protocols for wireless sensor networks (WSNs) represent an engineering area that has attracted the attention of several research communities due to its rapid development and evolution. Among all approaches, the clustering routing protocol corresponds to the most effective method to manage the energy consumption of all sensor nodes. It allows efficiently increasing the working cycle of sensor networks [8,9]. The Low Energy Adaptive Clustering Hierarchy Protocol (LEACH) [10] represents the most popular clustering routing protocol for homogeneous WSNs. Under the operation of LEACH, cluster heads (CHs) are randomly determined while all the sensor nodes are constantly interchanged as CHs. The energy expenditure is uniformly assigned to each node in order to enlarge the network performance. Several protocols that consider the LEACH scheme as a basis have been introduced in the literature with different performance levels [11–15]. Another interesting routing protocol is the Vice-CH-enabled [16] algorithm, which reduces the overhead provoked by clustering through the consideration of a two-step procedure. In its operation, the CHs are determined to take into account only the nodes that currently survive. The numerical results of Vice-CH-enabled demonstrate that its performance is higher than those based on the LEACH protocol [17].

The LEACH and Vice-CH-enabled protocols are devised to operate with homogeneous WSNs. However, they maintain a low performance when they are applied to HWSNs. Based on the concepts of LEACH and Vice-CH-enabled schemes, other routing protocols [18,19] were conceived to fulfill the necessities involved by HWSNs. One of the most popular protocols for HWSNs is the Stable Election Protocol (SEP) [20]. In its operation, SEP considers two different node types: advanced and normal elements. In general, advanced nodes present a higher probability than normal elements to become CHs. Therefore, the SEP protocol applies different threshold schemes depending on the node type. The idea behind these techniques is to use the energy of advanced nodes efficiently and avoid the consumption of normal nodes in order to extend the working cycle of the network. Several other protocols have been proposed based on the SEP paradigm. Some examples include the Modified Stable Election Protocol (M-SEP) [21], which evaluates the energy of each node in terms of the complete network to select the nodes that will become CHs. Under such conditions, nodes with higher energy levels present a better opportunity to be CHs than those that maintain low energy values. With the M-SEP method, the life cycle of the network and transmission rate of information is considerably enhanced. Another interesting SEP approach is the P-SEP scheme [22]. The P-SEP evaluates the uncertainty of the node selection through the estimation in each round of an energy threshold. By using

this energy threshold, it is avoided that a node with a lower energy level than the threshold value turns into a CH. With some concepts similar to P-SEP, the clustering protocol Distributed Energy-Efficient Clustering Algorithm (DEEC) [23] was introduced for HFWSNs. DEEC also uses a threshold limit for selecting the sensor nodes as CHs. Its main characteristic is that the period of time in which a sensor node behaves as a CH depends on the sensor energy level. The idea behind this decision is that high energy nodes produce stable transmission cycles in the network.

On the other hand, metaheuristic methods are optimization schemes able to solve complex systems.

These techniques are inspired by our scientific understanding of biological or social processes that from an abstraction level, can be considered to be search strategies. Some examples of popular metaheuristic methods include the Particle Swarm Optimization (PSO) [24], Genetic Algorithms [25], the Artificial Bee Colony (ABC) algorithm [26], the Differential Evolution (DE) method [27], the Gravitational Search Algorithm (GSA) [28] and the Flower Pollination Algorithm (FPA) [29]. Metaheuristic schemes do not need convexity, continuity differentiability, or certain initial conditions, which corresponds to an important advantage in comparison to other techniques. Despite their interesting results, these search strategies maintain different difficulties when they are applied to high multi-modal optimization problems.

The Yellow Saddle Goatfish Algorithm (YSGA) [30] represents a metaheuristic scheme that emulates the hunting behavior of the Yellow Saddle Goatfish. The algorithm uses different operators that allow a better balance of exploration-exploitation, reducing the typical flaws of other metaheuristic methods such as the premature convergence and the tendency to be trapped in local optima. Such characteristics have motivated their use in different complex engineering optimization problems such as system control [31] and energies [32]. In contrast to other metaheuristic algorithms, YSGA implements Lévy flights and logarithmic spirals to update the position of search individuals. Furthermore, YSGA considers groups of particles to guide the search through different directions while covering more land space.

As an alternative to classical approaches, the problem of designing clustering protocols was also faced with metaheuristic techniques [33–35]. In the literature, metaheuristic methods demonstrated obtaining a better performance than those based on traditional computational techniques in terms of accuracy and robustness. From the metaheuristic perspective, the process of a clustering routing protocol is translated into an optimization problem. Therefore, an objective function is defined to evaluate the quality of a solution. Guided by this objective function, the algorithm modifies the configurations of sensor nodes until obtaining the solution that presents a longer work cycle for the network [36,37]. Several clustering protocols have been proposed based on metaheuristic principles. Some examples include the Energy Centers using Particle Swarm Optimization (EC-PSO) [38]. Using this approach, by using a geometric scheme, the CHs are firstly selected. Then, the PSO algorithm is considered to identify the central sensor of the network, which will represent the CH. The EC-PSO method also includes a mechanism to avoid the consideration of nodes with low energy levels. Another interesting example is Genetic-Algorithm-Based Energy-Efficient Clustering (GAEEC) [39]. In this clustering protocol, the Genetic algorithms are used to identify the CH sensors through a clustering metric. The GA considers as the objective function, a model that combines the energy of the nodes, and the transmission cost. Recently, a routing protocol based on the Gray Wolf Optimizer has been also proposed [17]. In this scheme, different fitness functions are defined to evaluate the node characteristics of each sensor. Such values are considered to be weights that are dynamically modified depending on the distance among the nodes of the network. Therefore, the idea is to identify the configuration of CH nodes that reduce the aggregate value of the weights. Although metaheuristic methods present interesting results, they have a critical problem, such as their low premature convergence. This fact generates that such methods frequently obtain sub-optimal CH configurations, which reduces their working cycles. The difference between the proposed YSGA clustering protocol and other protocols that also implement metaheuristic algorithms is the strategy to select CHs. In our method, we have encoded candidate solutions to automatically choose the number of CHs and avoid a fixed percentage

of CHs or predefined probabilities to select CHs. Our strategy ensures the optimal number of CHs is chosen in every round by the implementation of the YSGA. Furthermore, the proposed method includes a different cost function to guide the search toward the clustering network configuration that better fit the requirements of energy savings and load balancing.

In this work, a new clustering routing protocol for Wireless Sensor Networks is introduced. The scheme is developed based on the Yellow Saddle Goatfish Algorithm (YSGA). Under the proposed approach, the number of cluster heads and the selection of optimal cluster heads is determined by the YSGA algorithm, while sensor nodes are assigned to its nearest cluster head. Therefore, the cluster structure of the network is updated by the YSGA to ensure an optimal distribution of cluster heads and reduce the transmission distance. Experiments show competitive results and demonstrate that the proposed routing protocol minimizes the energy consumption, improves the lifetime, and prolongs the stability period of the network in comparison with the state-of-the-art clustering routing protocols.

The next sections of the paper are structured as follows. In Section 2, the main characteristics of the Yellow Saddle Goatfish Algorithm (YSGA) are reviewed. Section 3 describes the proposed clustering protocol. Section 4 presents the numerical results of the proposed protocol compared with other well-known methods. Finally, the conclusions are discussed in Section 5.

## 2. Review of the Yellow Saddle Goatfish Algorithm

The YSGA is a recent metaheuristic algorithm inspired by the yellow saddle goatfish behavior. This optimization method considers a population of individuals divided into different groups, where each subpopulation is created using the k-means algorithm. Individuals in each group can play two different roles: chaser and blocker. Additionally, exchange roles and change zone operators are included in the search strategy of the YSGA.

### 2.1. Chaser Behavior

In every subpopulation, the individual with the best fitness value is the chaser  $\Phi_l$ . This particle leads the group through the search, and its behavior is modeled by a Lévy flight random process. Thus, the location of the chaser is defined as:

$$\Phi_l^{t+1} = \Phi_l^t + \alpha \left( \frac{u}{|v|^{1/\beta}} \right) (\Phi_l^t - \Phi_{best}^t) \quad (1)$$

$$0 < \beta \leq 2$$

where  $\alpha$  is the step size, which value is 1. The values of  $u$  and  $v$  are estimated from the normal distribution as:

$$u \sim N(0, \sigma_u^2)$$

$$v \sim N(0, \sigma_v^2)$$

Considering  $\Gamma$  as the Gamma function,  $\sigma_u$  and  $\sigma_v$  are defined as:

$$\sigma_u = \left\{ \frac{\Gamma(1+\beta) \sin \frac{\pi\beta}{2}}{\Gamma(\frac{1+\beta}{2}) \beta 2^{(\beta-1)/2}} \right\}^{1/\beta}, \quad \sigma_v = 1 \quad (2)$$

The Lévy index  $\beta$  controls the tail of the probability distribution. The value of  $\beta$  is calculated as:

$$\beta = 1.99 + \frac{0.001t}{t_{max}/10} \quad (3)$$

where  $t_{max}$  is the maximum number of iterations, while  $t$  is the current number of iterations. The best chaser among groups is the global best particle  $\Phi_{best}$ . Thus, the position of the global best is updated by the following equation:

$$\Phi_{best}^{t+1} = \Phi_{best}^t + \alpha \left( \frac{u}{|v|^{1/\beta}} \right) \quad (4)$$

## 2.2. Blocker Behavior

Every group has one chaser individual. Therefore, the remaining particles in each group are considered blockers  $\varphi_g$ . The position of blockers are updated according to a logarithmic spiral defined as:

$$\varphi_g^{t+1} = D_g \cdot e^{b\rho} \cdot \cos 2\pi\rho + \Phi_l \quad (5)$$

Here,  $\rho$  is a random number in the interval of  $[a, 1]$ , where  $a$  is linearly decreased from  $-1$  to  $-2$  over iterations. The value of the parameter  $b$  is  $1$ . The distance  $D_g$  between the blocker and the corresponding chaser is calculated as:

$$D_g = |r \cdot \Phi_l - \varphi_g^t| \quad (6)$$

The random number  $r$  is among  $[-1, 1]$ .

## 2.3. Exchange of Roles and Change of Zone

The exchange of roles allows blocker individuals to become a chaser particle. This is a simple mechanism that updates the chaser if a blocker is better in terms of the fitness value. On the other hand, the change of zone is a strategy to escape from local optima. If a better solution is not found in a determined time, then the change of zone operation is executed according to the following equation:

$$p_g^{t+1} = \frac{\Phi_{best} + p_g^t}{2} \quad (7)$$

This equation updates the position of every particle  $p_g^t$  in the population without considering its role.

Although YSGA implements Lévy flights in its search strategy, as do other algorithms such as Cuckoo Search (CS), there are significant differences among them. As an example, CS uses a global population of host nests for the search. On the other hand, YSGA uses subpopulations of search agents with blockers and a chaser in each subpopulation. Besides, the mainly updated mechanism in CS is Lévy flights, while the updated mechanisms in YSGA are Lévy flights (for exploration) and a logarithmic spiral path for exploitation. Additionally, the CS algorithm includes a selection operator, which discards the worst solutions under a certain probability to replace them with new ones. In contrast, YSGA implements other operators, such as the exchange of roles and the change of zone. These operators help to improve the solutions, so they do not need to be removed from the population by a selection mechanism. Furthermore, the exchange of roles promotes diversity in every subpopulation, while the change of zone avoids stagnation in local optima.

## 3. Proposed Clustering Routing Protocol

In this section, the clustering routing protocol based on the Yellow Saddle Goatfish algorithm is introduced. The operation of the proposed method is organized in two main phases: the configuration phase or set-up phase and the operation phase or steady-state phase.

### 3.1. Configuration Phase

Initially, network assumptions and its topological structure must be defined since different network configurations can be established depending on the requirements of the application. The network

assumptions comprise the characteristics of the sensor nodes. On the other hand, in the network structure, the establishment of WSNs includes the deployment of nodes in the sensing area to create a topological structure where the data will be collected. Under such considerations, the proposed protocol considers the following network assumptions and structure. This configuration has been adopted to be consistent with other related works.

Network assumptions:

1. The network has one base station  $BS$ , a set of cluster heads  $CH$ , and a set of sensor nodes  $n$
2. The power of the base station is externally supplied, while the energy of sensor nodes is limited
3. A sensor node will be considered dead when it is out of power
4. All sensor nodes are homogeneous

Network structure:

- a. Initially, all nodes are randomly deployed in the sensing area
- b. Nodes location will not change during the whole life of the network
- c. The base station is placed at the center of the sensing area
- d. The number of clusters is not fixed
- e. Every normal node (also called leaf node) is added to its nearest cluster head

Once the network topology is determined, the entire network connection establishment process begins, where the set-up phase is executed. In the set-up phase, initial CHs are chosen to build the initial cluster network configuration. In the selection of the optimal CHs, the proposed strategy considers the following energy consumption model.

### Energy Consumption Model

In WSNs, the activities that consume most of the energy are data transmission and reception. The energy consumption for transmitting or receiving data depends on the distance  $d$  and the size of the data packet. Under such considerations, the required energy to transmit a  $l$ -bit data packet is defined in Equation (8).

$$E_{TX}(l, d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2, & d \leq d_{th} \\ lE_{elec} + l\varepsilon_{mp}d^4, & d > d_{th} \end{cases} \quad (8)$$

where  $E_{TX}$  is the energy consumption for transmitting data,  $l$  is the data length,  $E_{elec}$  is the energy dissipation for transmitting or receiving 1 bit of data,  $\varepsilon_{fs}$  is the coefficient of energy dissipation in the free space model,  $\varepsilon_{mp}$  is the coefficient of energy dissipation in the multi-path attenuation model, and  $d_{th}$  is the transmission distance threshold defined in Equation (9).

$$d_{th} = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \quad (9)$$

The required energy to receive a  $l$ -bit data packet is calculated in Equation (10).

$$E_{RX}(l) = lE_{elec} \quad (10)$$

Since a normal sensor node  $n_i$  only transmit data to the cluster head, the following equation can calculate its energy consumption:

$$E(n_i) = E_{TX}(l, d) \quad (11)$$

However, the energy consumption of a cluster head must include the energy consumption of receiving packets from cluster member nodes, aggregating data, and transmitting aggregated data to the base station. Therefore, the energy consumption of a cluster head  $CH_j$  is calculated as Equation (12).

$$E(CH_j) = E_{RX}(l)N_j + (N_j + 1)lE_{DA} + E_{TX}(l, d) \quad (12)$$

where  $N_j$  is the number of member nodes in cluster  $j$ , and  $E_{DA}$  is the energy for 1 bit of data aggregation. Under the above considerations, the residual energy of a leaf sensor node  $n_i$  can be estimated by Equation (13).

$$E_r^t(n_i) = \begin{cases} E_r^{t-1}(n_i) - E_{TX}(l, d), & E_r^{t-1}(n_i) > 0 \\ 0, & E_r^{t-1}(n_i) \leq 0 \end{cases} \quad (13)$$

While the residual energy of a cluster head  $CH_j$  is described in Equation (14).

$$E_r^t(CH_j) = \begin{cases} E_r^{t-1}(CH_j) - E(CH_j), & E_r^{t-1}(CH_j) > 0 \\ 0, & E_r^{t-1}(CH_j) \leq 0 \end{cases} \quad (14)$$

During the set-up phase, the YSGA is implemented to select initial CHs and build the initial cluster network configuration considering the presented energy consumption model. Once the YSGA has chosen the cluster heads, all the remaining nodes will join the nearest CH. However, if the distance of a leaf node to its closest CH is higher than the distance to the base station, then the leaf node will not be part of a cluster. Instead, it will transmit the information directly to the base station.

The process of selecting CHs to build an optimal network configuration by the YSGA is described in the operation phase.

### 3.2. Operation Phase

In the operation phase, the information is transferred to the BS once the selection of optimal CHs has established the network connections. Data transmission occurs every round, where one round is completed every time the information is transferred from normal nodes to CHs, and then to the BS. After each round, the network is reconfigured by executing the YSGA algorithm to select the new optimal CHs and rebuild the network links. The process for choosing CHs and the optimal cluster network configuration is described in the next subsections.

#### 3.2.1. Selection of Optimal Cluster Heads

In the proposed protocol, the YSGA algorithm determines the number of CHs and which nodes will be selected as CHs. In the selection, two main criteria are considered: the distance from CH to the base station and the residual energy of CH.

Regarding the first criteria, establishing the most efficient connection will promote energy savings. Therefore, the selection of CHs considers reducing the length of the links. Additionally, it is assumed that links work at the same frequency, and crosslinks are eliminated because they generate interference during data transmission.

Concerning the second criterion, the proposed method selects sensor nodes as CHs if its residual energy is higher enough. Cluster heads spend more energy than normal nodes because, according to the energy consumption model, CHs consume energy when receiving, aggregating, and transmitting data. Hence, sensor nodes with the highest residual energy must be selected as CHs for load balancing.

In contrast to other clustering routing protocols, the number of CHs is not fixed. Instead, the number of CHs is dynamically changed in order to build the best network configuration in every round. Furthermore, the selection of CHs is not random or probabilistic, such as in LEACH or DEEC protocols. In contrast, in our method, the YSGA algorithm automatically finds the optimal number of CHs and selects the best sensor nodes to become CHs in every round. This strategy allows the implementation of the proposed protocol for a wide range of WSN applications without concern about determining the number of clusters.

Once the optimal cluster heads are selected, each sensor node is incorporated into the nearest cluster head. Nevertheless, if the distance from the sensor node to the BS is shorter than the length to the CH, then the sensor node is not clustered, so the information of this node is transmitted directly to the base station. After completed the clustering process, a potential network configuration is prepared

for transmission, and it can be evaluated to determine if it is optimal. The procedure for selecting the best cluster network configuration will be explained in the following section.

### 3.2.2. Selection of Optimal Cluster Network Configuration

In the proposed protocol, the YSGA algorithm finds the optimal cluster network configuration  $C$  by selecting the optimal set of cluster heads. Thus, in every iteration, cluster heads are determined to generate a new network structure. After that the network is evaluated in terms of its fitness value. In this work, we implement a cost function to evaluate the effectiveness of the network configuration. The objective function includes four terms: the total intra-cluster distance, the total distance from cluster heads to the base station, the energy consumption, and the residual energy of cluster heads. The cost function is defined in Equation (15).

$$f(C) = w_1 f_1(C) + w_2 f_2(C) + w_3 f_3(C) + w_4 f_4(C) \quad (15)$$

where  $w_1, w_2, w_3$ , and  $w_4$  are weighted constant factors that control the contribution of every element in the formulation.

The first term is the total intra-cluster distance, which is calculated as:

$$f_1(C) = \frac{d_{max}(\sum_{j=1}^k d(n_j, CH_j)) - \sum_{j=1}^k d(n_j, CH_j)}{d_{max}(\sum_{j=1}^k d(n_j, CH_j)) - d_{min}(\sum_{j=1}^k d(n_j, CH_j))} \quad (16)$$

Considering the subset of sensor nodes  $n_j$  assigned to cluster  $j$  and  $CH_j$  as the cluster head, the maximum and minimum total intra-cluster distance from every sensor node in  $n_j$  to its cluster head  $CH_j$  is  $d_{max}$  and  $d_{min}$ , respectively, while the total intra-cluster distance  $d$ .

The second term is the total distance from cluster heads to the base station, formulated as:

$$f_2(C) = \frac{d_{max}(\sum_{j=1}^k d(CH_j, BS)) - \sum_{j=1}^k d(CH_j, BS)}{d_{max}(\sum_{j=1}^k d(CH_j, BS)) - d_{min}(\sum_{j=1}^k d(CH_j, BS))} \quad (17)$$

Here, the maximum and minimum total distance from every cluster head  $CH_j$  to the base station  $BS$  is  $d_{max}$  and  $d_{min}$ , respectively, while the total distance from cluster heads to the base station is  $d$ .

The third term is the energy consumption, calculated as:

$$f_3(C) = \frac{E_{max}(\sum_{j=1}^k \sum_{i=1}^{N_j} E(n_{ij}) + E(CH_j)) - E(\sum_{j=1}^k \sum_{i=1}^{N_j} E(n_{ij}) + E(CH_j))}{E_{max}(\sum_{j=1}^k \sum_{i=1}^{N_j} E(n_{ij}) + E(CH_j)) - E_{min}(\sum_{j=1}^k \sum_{i=1}^{N_j} E(n_{ij}) + E(CH_j))} \quad (18)$$

where  $E(n_{ij})$  is the energy consumption of the sensor node  $i$  that belongs to cluster  $j$ . The number of member nodes in cluster  $j$  is  $N_j$ . The maximum and minimum total energy consumption of the network is  $E_{max}$  and  $E_{min}$ , respectively.

The fourth term is the sum of the residual energy ratio between sensor node members and its cluster head:

$$f_4(C) = \frac{E_{r_{max}}\left(\sum_{j=1}^k \frac{\sum_{i=1}^{N_j} E_r(n_{ij})}{E_r(CH_j)}\right) - E_r\left(\sum_{j=1}^k \frac{\sum_{i=1}^{N_j} E_r(n_{ij})}{E_r(CH_j)}\right)}{E_{r_{max}}\left(\sum_{j=1}^k \frac{\sum_{i=1}^{N_j} E_r(n_{ij})}{E_r(CH_j)}\right) - E_{r_{min}}\left(\sum_{j=1}^k \frac{\sum_{i=1}^{N_j} E_r(n_{ij})}{E_r(CH_j)}\right)} \quad (19)$$

where the residual energy of the cluster head is  $E_r(CH_j)$ , while the residual energy of sensor node members is  $E_r(n_{ij})$ . The maximum and minimum total residual energy of the network is  $E_{max}$  and  $E_{min}$ , respectively.

A smaller value of  $f(C)$  indicates that the selected cluster heads have generated a better cluster network configuration. In that sense, the YSGA algorithm will choose a set of cluster heads to build different cluster networks in every iteration in order to find the optimal network configuration. The proposed algorithm is summarized in Algorithm 1.

---

**Algorithm 1.** Proposed routing protocol.

---

Input: number of alive nodes

Output: a set of  $k$  clusters

while  $t \leq t_{max}$

```

|   Select cluster heads according to the search strategy of YSGA
|   if the distance from a sensor node to the nearest cluster head is shorter than to the base station
|   |   Assign the sensor node to the nearest cluster head
|   else
|   |   The sensor node will transmit directly to the base station
|   end if else
|   Calculate the objective function of the current cluster network configuration  $f(C^t)$ 
|   if  $f(C^t) < f(C^{opt})$ 
|   |   Update the optimal cluster network configuration as:
|   |    $f(C^{opt}) \leftarrow f(C^t)$ 
|   end if
end while

```

---

#### 4. Simulation and Experimental Results

Estimating the proposed protocol's performance, we take into account the network lifetime, network stability period, total residual energy, residual energy deviation, and the number of received data packets by the sink node. These metrics are described as follows:

- Network lifetime: It is the time between the start of the network running and the death of the first node. Besides, it is known as the network stability period.
- Network instability period: It is the time between the death of the first node and the death of all nodes.
- Residual energy: This metric is calculated by considering the total residual energy of all sensor nodes. Usually, this value is presented as a percentage indicator of the total remaining energy of the network, which is represented in Equation (20).

$$p(t) = \frac{\sum_{i=1}^N E_r(n_i, t)}{\sum_{i=1}^N E_o(n_i, t)} \quad (20)$$

where  $p(t)$  is the percentage of the total residual energy in the round  $t$ . The total residual energy in round  $t$  is  $\sum_{i=1}^N E_r(n_i, t)$  and the sum of the initial energy  $E_o$  of all the sensor nodes in the network is  $\sum_{i=1}^N E_o(n_i, t)$ .

- Residual energy deviation: The variation between the node with the most remaining energy and the node with the least remaining energy. The residual energy deviation in the round  $t$  is calculated as follows in Equation (21).

$$D_R(t) = E_{r_{max}}(\mathbf{n}, t) - E_{r_{min}}(\mathbf{n}, t) \quad (21)$$

where  $E_{r_{\max}}(\mathbf{n}, t)$  and  $E_{r_{\min}}(\mathbf{n}, t)$  are the maximum and minimum residual energy among the set of nodes  $\mathbf{n}$  in the network considering round  $t$ . However, to facilitate the analysis of the results, the residual energy deviation is reported in percentage units.

- Throughput: The number of data packets received by the BS.

Our method has been compared against other popular techniques such as Low Energy Adaptive Clustering Hierarchy Protocol (LEACH) [14], Stable Election Protocol (SEP) [21], and Distributed Energy-Efficient Clustering Algorithm (DEEC) [23].

For consistency in the experiments, parameter settings of Table 1 have been considered for all the routing protocols in comparison. Regarding the tuning parameters of the YSGA algorithm, Table 2 shows the adopted values in which the authors have reported the best performance of the YSGA. In Addition to the settings of Table 1, Table 3 shows the values of additional parameters considered in every method in comparison.

**Table 1.** Simulation parameters.

Parameter	Value
Sensing area	100 m <sup>2</sup>
Number of sensor nodes $N$	100
Packet size $l$	4000 bits
Energy for 1 bit of data aggregation $E_{DA}$	5 nJ/bit
The energy dissipation for transmitting or receiving 1 bit of data $E_{elec}$	50 nJ/bit
The coefficient of energy dissipation in the free space $\varepsilon_{fs}$ (transmission coefficient amplifier)	10 pJ/bit/m <sup>2</sup>
The coefficient of energy dissipation in the multi-path attenuation model $\varepsilon_{mp}$ (transmission coefficient amplifier)	0.0013 pJ/bit/m <sup>4</sup>
Initial energy of sensor nodes $E_0$	0.07 J

**Table 2.** Setting parameters of the proposed method.

Parameter	Value
Population size	20
Maximum number of iterations	50
$w_1, w_2, w_3, w_4$	1
$\alpha$	1
$b$	1

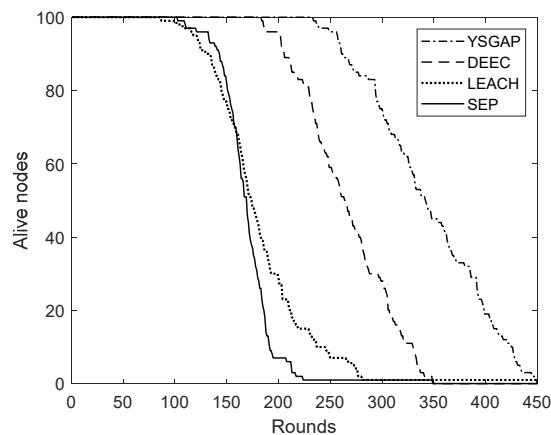
**Table 3.** Setting parameters of compared methods.

	Parameter	Value
DEED	Probability of CHs $p_{opt}$	0.1
LEACH	Percentage of CHs $p$	0.05
SEP	Probability of CHs $p$	0.1

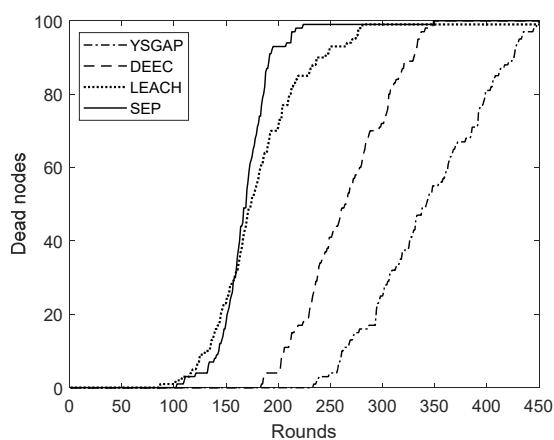
The simulation was implemented in MATLAB R2019a(The MathWorks, Inc. Natick, Massachusetts, USA), in a computer with a processor Intel(R) Core (TM)i7-8550ucpu@1.80GHz1.99GHz.

#### 4.1. Network Lifetime

Figures 1 and 2 show the obtained results of the simulations in terms of the network lifetime. The history of alive nodes is reported in Figure 1, while the number of dead sensor nodes is illustrated in Figure 2. From these figures, it is evident that the proposed clustering routing protocol performs better than the protocols in comparison. Besides, nodes start to die after 230 rounds in the proposed protocol, while nodes begin to die after 180 rounds in DEEC.



**Figure 1.** Network lifetime: alive nodes history.

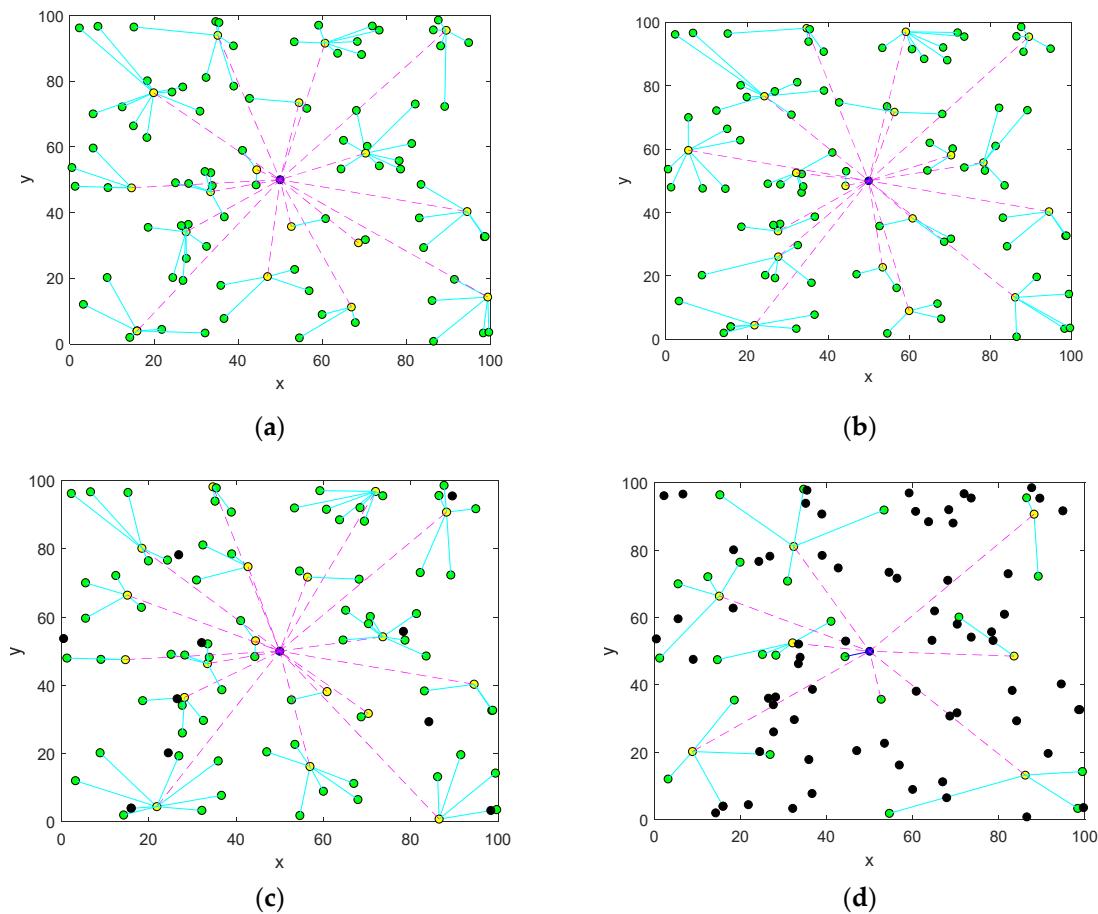


**Figure 2.** Network lifetime: dead nodes history.

The proposed method and DEEC might seem similar in their resulting patterns since both schemes are based on similar performing principles. Under such conditions, both techniques could generate similar rates of energy consumption in every round. However, the proposed strategy includes additional factors that influence the selection of CHs, which leads to an optimal cluster configuration that extends the lifetime of the network. In contrast, the DEEC method obtains suboptimal cluster topologies that produce higher consumption. Moreover, selected CHs do not always have the maximum remaining energy, which leads to inefficient energy expenditure. Furthermore, cluster formation also stops quickly in DEEC, which forces normal nodes to communicate directly to the base station, causing faster exhaustion of nodes, and consequently, reducing the network lifetime.

Considering the achieved results, our method manages to prolong the lifetime of the network more rounds than the other techniques. These results were obtained because our strategy is able to automatically change the number of clusters in every round to balance the energy load. Besides, the selection of sensor nodes to become CHs is also dynamic. Therefore, the roles of sensor nodes can change in each round to obtain the network configuration that reduces the energy consumption and increases the lifetime of the nodes.

Additionally, the clustering evolution of the network is presented in Figure 3, where the clustering formation is visualized in different rounds. This figure shows how clusters are changing as different cluster heads are selected over time. Furthermore, it reveals the distribution of dead nodes and how they are ignored in the clustering process.

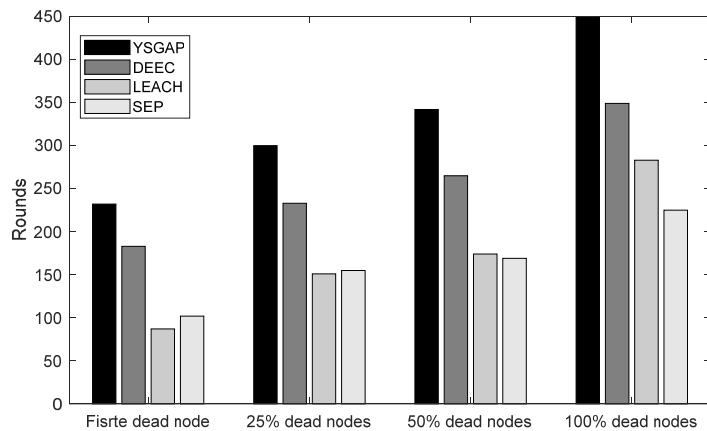


**Figure 3.** Clustering evolution of the network: (a) network configuration in round 5, (b) in round 100, (c) in round 270, and (d) in round 400.

In the simulation, sensor nodes are represented as circles, while communication links are symbolized as connection lines. From Figure 3, the base station can be observed at the center of the sensing area, represented as a blue circle. On the other hand, alive sensor nodes can be identified as green circles, while cluster heads are yellow circles, and dead nodes are simulated as black circles. The links that communicate intracluster nodes can be observed as cyan lines, while communication links from cluster heads to the base station are represented as magenta dashed lines. The connection between independent nodes and the base station is represented as blue lines. Figure 3a–d illustrate the optimal cluster network configuration in rounds 5, 100, 270, and 400, respectively.

#### 4.2. Network Instability Period

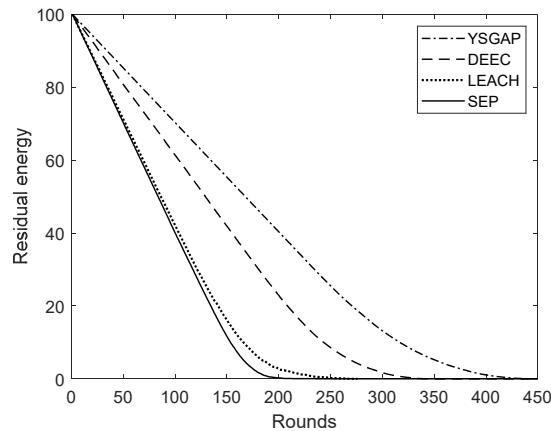
For a closer inspection, Figure 4 shows in which round has died the first sensor node. Additionally, the round in which the half of the nodes have died, and the death of all nodes is also reported in this figure. According to the obtained results, it is observed that the proposed protocol achieves the highest values. In our method, the first dead node occurs after 230 rounds, the 50% of the nodes died after 340 rounds, and the 100% of the sensor nodes died almost in 450 rounds. The protocols in comparison show inferior indices regarding the first, half, and last dead in the network.



**Figure 4.** Network lifetime: percentage of dead nodes.

#### 4.3. Residual Energy

The total residual energy of the network considers the remaining energy of all nodes in each round. This metric is reported in Figure 5, where it is evident that the proposed method manages to save more energy than the other protocols. The total residual energy is reported considering 100% of the residual energy in round 0, and 0% of the residual energy in the last round. As it was expected, our approach finds the network configuration that spends the lowest energy in every round. Therefore, the residual energy of the network is higher than the obtained with the other methods. From the figure, we can conclude that the proposed protocol achieves the lowest energy expenditure and lengthens the life of the nodes.

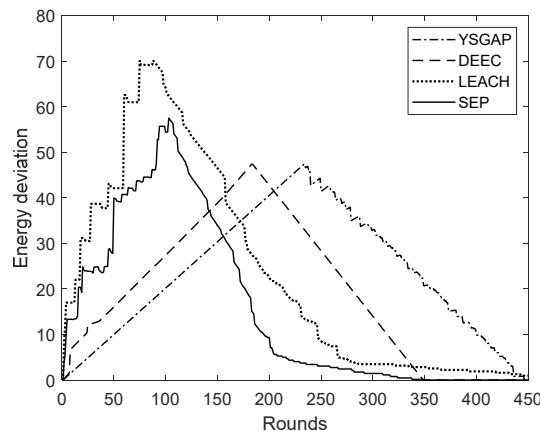


**Figure 5.** Total residual energy.

#### 4.4. Energy Deviation

The energy deviation considers the difference between the sensor node with the highest residual energy and the sensor node with the lowest residual energy. This is an indicator of load balancing. If the energy deviation is high, then the energy expenditure is not balanced among sensor nodes. The unbalanced energy expenditure causes premature node death. Consequently, the network lifetime is reduced. For a better load balancing, it is expected to reduce the highest peak.

Figure 6 shows the energy deviation of the reported protocols. From the figure, it can be observed that the highest energy deviation is reached by LEACH, while the lowest is achieved by DEEC and YSGAP. However, the YSGAP protocol reaches the highest peak after 230 rounds, while DEEC after 180 rounds. The later the highest peak is reached, the better. Therefore, we can conclude that YSGAP outperforms its opponents in terms of the energy deviation.

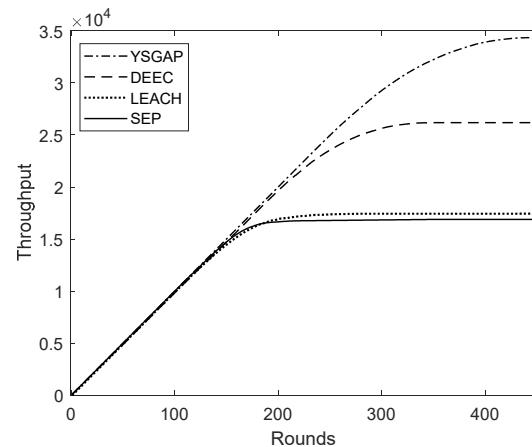


**Figure 6.** Energy deviation.

#### 4.5. Throughput

The number of the sent packet to the base station is known as throughput. The more sent packet, the better. This means that there was more information collected by the sensor nodes.

Figure 7 shows the total number of sent packets to the base station in every round. From the figure, it is observed that the YSGAP sends more data packets to the base station than the other protocols. This was as expected since the proposed method prolongs the network lifetime, and consequently, the collected data is superior.



**Figure 7.** Throughput: total number of packets sent to the BS.

## 5. Conclusions

In this work, a novel clustering routing protocol called YSGAP was proposed. The method finds the optimal network configuration to reduce the energy consumption in every round and prolong the network lifetime. In its operation, the YSGAP protocol automatically determines the number of clusters heads and selects the sensor nodes that will play the role of cluster heads in every round. Our method was compared with some of the most popular clustering routing protocols in the literature, such as DEEC, LEACH, and SEP. Experimental results proved that the proposed strategy outperforms the methods in comparison. The evaluation of the proposed routing protocol was carried out, demonstrating that our approach manages to considerably increase the lifetime of the network, while also providing robustness in communications, fault tolerance, and time-bounded response questions.

As feature work, the proposed protocol can be enhanced by including more terms in the cost function, such as considering uniform clusters sizes for a better load balancing. Additionally, the

encoded candidate solutions might be arranged differently so that the dimensionality of the optimization problem can be considerably reduced. This change can increase the scalability of the sensor network. Furthermore, the off-line execution of the protocol can consider the main drawback of the proposed method. Therefore, a significant improvement can be its development in parallel to speed up the process of finding the optimal cluster configuration. This improvement will allow the protocol to run online in real time applications.

**Author Contributions:** Relatively to the present manuscript, A.R. developed the energy algorithm, prepared, and executed the scenario and test-bed, interpreted and analyzed the results, and designed the methodology. C.D.-V.-S. was involved in the algorithm design, she supervised the research methodology and the approach of this work, she performed the formal analysis and interpreted the results. R.V. reviewed, interpreted, and drafted the simulation results, he reviewed the methodology and the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Lin, X.; Shroff, N.B.; Srikant, R. A tutorial on cross-layer optimization in wireless networks. *IEEE J. Sel. Areas Commun.* **2006**, *24*, 1452–1463.
2. Gheisari, M.; Alzubi, J.; Zhang, X.; Kose, U.; Saucedo, J.A.M. A new algorithm for optimization of quality of service in peer to peer wireless mesh networks. *Wirel. Netw.* **2019**, *6*, 1–9. [[CrossRef](#)]
3. Nam, D. Comparison Studies of Hierarchical Cluster-Based Routing Protocols in Wireless Sensor Networks. *CATA* **2020**, *69*, 334–344.
4. Rajagopal, S.; Vani, R.; Kavitha, J.C.; Saravanan, R. Lifetime Improvement of Wireless Sensor Networks Using Tree-Based Routing Protocol. In *EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing*; Springer: Cham, Switzerland, 2020; pp. 51–61.
5. Aadri, A.; Idrissi, N. An energy efficient hierarchical routing scheme for wireless sensor networks. *Comput. Sci. Inf. Technol.* **2017**, 137–148. [[CrossRef](#)]
6. Sabor, N.; Sasaki, S.; Abo-Zahhad, M.; Ahmed, S.M. A comprehensive survey on hierarchical-based routing protocols for mobile wireless sensor networks: Review, taxonomy, and future directions. *Wirel. Commun. Mob. Comput.* **2017**, *2017*, 23. [[CrossRef](#)]
7. Bandyopadhyay, S.; Coyle, E.J. An energy efficient hierarchical clustering algorithm for wireless sensor networks. In Proceedings of the IEEE INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428), San Francisco, CA, USA, 30 March–3 April 2003; IEEE: Piscataway, NJ, USA, 2003; Volume 3, pp. 1713–1723.
8. Liu, X. A survey on clustering routing protocols in wireless sensor networks. *Sensors* **2012**, *12*, 11113–11153. [[CrossRef](#)]
9. Mazumdar, N.; Roy, S.; Nayak, S. A survey on clustering approaches for wireless sensor networks. In Proceedings of the 2018 2nd International Conference on Data Science and Business Analytics (ICDSBA), Changsha, China, 21–23 September 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 236–240.
10. Heinzelman, W.R.; Chandrakasan, A.; Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, HI, USA, 7 January 2000; IEEE: Piscataway, NJ, USA, 2000; p. 10.
11. Tong, M.; Tang, M. LEACH-B: An improved LEACH protocol for wireless sensor network. In Proceedings of the 2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM), Chengdu, China, 23–25 September 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 1–4.
12. Geetha, V.A.; Kallapur, P.V.; Tellajeera, S. Clustering in wireless sensor networks: Performance comparison of leach & leach-c protocols using ns2. *Procedia Technol.* **2012**, *4*, 163–170.
13. Marappan, P.; Rodrigues, P. An energy efficient routing protocol for correlated data using CL-LEACH in WSN. *Wirel. Netw.* **2016**, *22*, 1415–1423. [[CrossRef](#)]

14. Razaque, A.; Mudigulam, S.; Gavini, K.; Amsaad, F.; Abdulgader, M.; Krishna, G.S. H-LEACH: Hybrid-low energy adaptive clustering hierarchy for wireless sensor networks. In *Proceedings of the 2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT), Farmingdale, NY, USA, 29 April 2016*; IEEE: Piscataway, NJ, USA, 2016; pp. 1–4.
15. Jerbi, W.; Guermazi, A.; Trabelsi, H. O-LEACH of routing protocol for wireless sensor networks. In *Proceedings of the 2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGIV), Beni Mellal, Morocco, 29 March–1 April 2016*; IEEE: Piscataway, NJ, USA, 2016; pp. 399–404.
16. Pachlor, R.; Shrimankar, D. VCH-ECCR: A centralized routing protocol for wireless sensor networks. *J. Sens.* **2017**, *2017*, 1–10. [\[CrossRef\]](#)
17. Zhao, X.; Ren, S.; Quan, H.; Gao, Q. Routing Protocol for Heterogeneous Wireless Sensor Networks Based on a Modified Grey Wolf Optimizer. *Sensors* **2020**, *20*, 820. [\[CrossRef\]](#)
18. Osamy, W.; Salim, A.; Khedr, A.M. An information entropy-based-clustering algorithm for heterogeneous wireless sensor networks. *Wirel. Netw.* **2020**, *26*, 1869–1886. [\[CrossRef\]](#)
19. Dutt, S.; Agrawal, S.; Vig, R. Cluster-head restricted energy efficient protocol (CREEP) for routing in heterogeneous wireless sensor networks. *Wirel. Pers. Commun.* **2018**, *100*, 1477–1497. [\[CrossRef\]](#)
20. Smaragdakis, G.; Matta, I.; Bestavros, A. *SEP: A Stable Election Protocol for Clustered Heterogeneous Wireless Sensor Networks*; Boston University Computer Science Department: Boston, MA, USA, 2004.
21. Singh, D.; Panda, C.K. Performance analysis of modified stable election protocol in heterogeneous wsn. In *Proceedings of the 2015 International Conference on Electrical, Electronics, Signals, Phuket, Thailand, 20–21 December 2015*; Communication and Optimization (EESCO): Visakhapatnam, India, 2015; pp. 24–25.
22. Naranjo, P.G.V.; Shojafar, M.; Mostafaei, H.; Pooranian, Z.; Baccarelli, E. P-SEP: A prolong stable election routing algorithm for energy-limited heterogeneous fog-supported wireless sensor networks. *J. Supercomput.* **2017**, *73*, 733–755. [\[CrossRef\]](#)
23. Qing, L.; Zhu, Q.; Wang, M. Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. *Comput. Commun.* **2006**, *29*, 2230–2237. [\[CrossRef\]](#)
24. Kennedy, J.; Eberhart, R. Particle swarm optimization. In *Proceedings of the ICNN’95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995*; IEEE: Piscataway, NJ, USA, 1995; Volume 4, pp. 1942–1948.
25. Goldberg, D.E.; Holland, J.H. Genetic algorithms and machine learning. *Mach. Learn.* **1988**, *3*, 95–99. [\[CrossRef\]](#)
26. Karaboga, D. *An Idea Based on Honey Bee Swarm for Numerical Optimization*; Technical Report-tr06; Engineering Faculty, Computer Engineering Department, Erciyes University: Kayseri, Turkey, 2005; Volume 200, pp. 1–10.
27. Storn, R.; Price, K. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **1997**, *11*, 341–359. [\[CrossRef\]](#)
28. Rashedi, E.; Nezamabadi-Pour, H.; Saryazdi, S. GSA: A gravitational search algorithm. *Inf. Sci.* **2009**, *179*, 2232–2248. [\[CrossRef\]](#)
29. Yang, X.S. Flower Pollination Algorithm for Global Optimization. In *International Conference on Unconventional Computing and Natural Computation*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 240–249.
30. Zaldivar, D.; Morales, B.; Rodríguez, A.; Valdivia, G.A.; Cuevas, E.; Pérez-Cisneros, M. A novel bio-inspired optimization model based on yellow saddle goatfish behavior. *Biosystems* **2018**, *174*, 1–21. [\[CrossRef\]](#)
31. Micev, M.; Čalasan, M.; Oliva, D. Fractional Order PID Controller Design for an AVR System Using Chaotic Yellow Saddle Goatfish Algorithm. *Mathematics* **2020**, *8*, 1182. [\[CrossRef\]](#)
32. Latif, A.; Das, D.C.; Barik, A.K.; Ranjan, S. Illustration of demand response supported co-ordinated system performance evaluation of YSGA optimized dual stage PIFOD-(1+ PI) controller employed with wind-tidal-biodiesel-based independent two-area interconnected microgrid system. *IET Renew. Power Gener.* **2020**, *14*, 1074–1086. [\[CrossRef\]](#)
33. Zhao, X.; Cui, Y.; Guo, Z.; Hao, Z. An Energy-Efficient Coverage Enhancement Strategy for Wireless Sensor Networks Based on a Dynamic Partition Algorithm for Cellular Grids and an Improved Vampire Bat Optimizer. *Sensors* **2020**, *20*, 619. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Zhao, X.Q.; Cui, Y.P.; Gao, C.Y.; Guo, Z.; Gao, Q. Energy-efficient coverage enhancement strategy for 3-D wireless sensor networks based on a vampire bat optimizer. *IEEE Internet Things J.* **2019**, *7*, 325–338. [\[CrossRef\]](#)

35. Ghosh, N.; Banerjee, I.; Sherratt, R.S. On-demand fuzzy clustering and ant-colony optimisation-based mobile data collection in wireless sensor network. *Wirel. Netw.* **2019**, *25*, 1829–1845. [[CrossRef](#)]
36. Kong, L.; Pan, J.S.; Snášel, V.; Tsai, P.W.; Sung, T.W. An energy-aware routing protocol for wireless sensor network based on genetic algorithm. *Telecommun. Syst.* **2018**, *67*, 451–463. [[CrossRef](#)]
37. Potthuri, S.; Shankar, T.; Rajesh, A. Lifetime improvement in wireless sensor networks using hybrid differential evolution and simulated annealing (DESA). *Ain Shams Eng. J.* **2018**, *9*, 655–663. [[CrossRef](#)]
38. Wang, J.; Gao, Y.; Liu, W.; Sangaiah, A.K.; Kim, H.J. An improved routing schema with special clustering using PSO algorithm for heterogeneous wireless sensor network. *Sensors* **2019**, *19*, 671. [[CrossRef](#)]
39. Singh, S.P.; Sharma, S.C. Genetic-algorithm-based energy-efficient clustering (GAEEC) for homogenous wireless sensor networks. *IETE J. Res.* **2018**, *64*, 648–659. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).