



A novel energy-aware method for clustering and routing in IoT based on whale optimization algorithm & Harris Hawks optimization

Ehsan Heidari¹

Received: 10 June 2023 / Accepted: 22 December 2023 / Published online: 3 February 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2024

Abstract

Smart objects in the Internet of Things (IoT) communicate with one another, gather information, and respond to users requests. In these systems, wireless sensors are used to collect data and monitor the environment at the lowest level. In IoT applications, wireless sensor networks play a pivotal role. Since IoT devices often use batteries, efficiency is important to them such that IoT-related standards and research efforts focus more on energy saving or conservation. In this paper, we have used two meta-heuristics algorithm for clustering and routing in IoT. We cluster the network using a clustering method called WOA-clustering based on the meta-heuristic Whale Optimization Algorithm (WOA) and select the optimal cluster heads. We then use a routing method called HHO-Routing based on the Harris Hawks Optimization (HHO) algorithm, a novel meta-heuristic algorithm, to route the cluster heads to BS. The use of the above methods results in reduced power consumption for reaching the base station (BS). Also, to prove the optimal performance of the proposed methods, these methods were simulated and compared with five different methods in a similar context. It was observed that the consumed energy, the number of survival cycles for the death of the first node, and the data transmission rate were improved. The proposed method is simulated in cooja simulator, and for a more accurate evaluation, we compare it with UCCGRA, PSO-SD, PUDCRP, EECRA, EEMRP algorithms. We see that the proposed method performs better than other methods in terms of energy consumption and network lifespan.

Keywords Internet of things · Wireless sensor networks · Clustering · Routing · Energy consumption · Whale optimization algorithm · Harris Hawks optimization

Mathematics Subject Classification 68M18 · 68M10 · 68M12

✉ Ehsan Heidari
eh.heidari@iau.ac.ir

¹ Department of Computer Engineering, Doroud Branch, Islamic Azad University, Doroud, Iran

1 Introduction

IoT refers to a network in which every physical object is identified by a sensor, and it builds up a network with other objects. These objects can communicate with each other independently and exchange information. Recent advances in analog and digital electronic components manufacturing technology, as well as advances in the design of low-power wireless radio components, have made possible the production of small, inexpensive sensors capable of wireless communication [1–3]. According to the concept of IoT, different devices have the ability to communicate wirelessly for being tracked or controlled over the Internet or even via a simple smartphone application, which describes the term IoT. In the Internet of Things, the goal is to establish a connection between everything. In other words, the Internet is used to establish the conditions necessary for human communication with objects that requires one or more sensors [4–8]. In WSN, the information of the neighboring sensors is integrated and transmitted to an external BS. Thus, a new approach is developed for the reliable and accurate monitoring of high-risk or inaccessible environments, which is rapidly being utilized in many applications including the military, security [9], medical, and rescue industries [10–12], as well as the environmental and industrial applications (lamps, household appliances such as tea makers, dishwashers, or even automobiles [13, 14]).

Figure 1 shows the structure of an IoT consisting of two layers. The bottom layer is called the infrastructure layer, which consists of a large number of sensors and BS, also referred to as the IoT sensors layer. The task of this layer is only to transmit sensed data to BS. The second layer is called the application layer, which receives

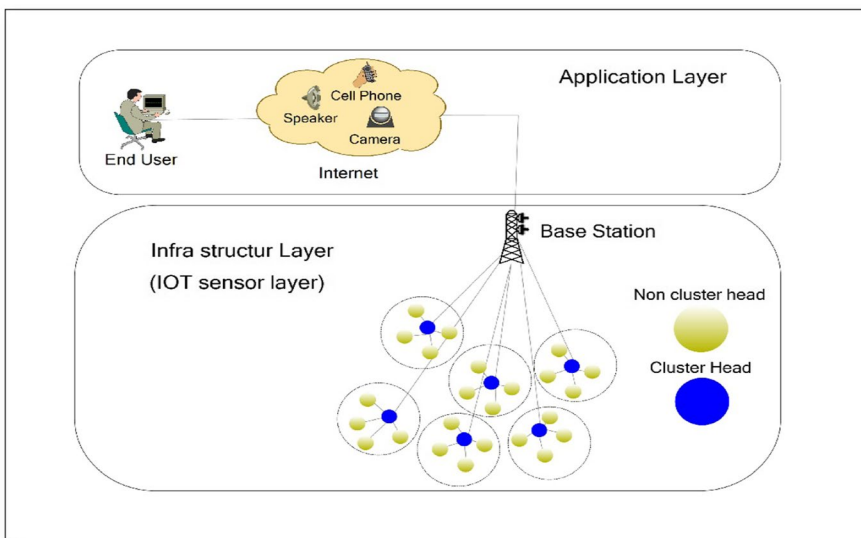


Fig. 1 The overall structure of the IoT

the lowest layer data obtained from BS via the Internet and serves applications. As you can see, wireless sensor networks play a pivotal role in IoT applications.

At the lowest level of the Internet of Things are sensors, and they use batteries. It is almost impossible to replace the batteries of the sensors. If the energy of these sensors is exhausted, they will practically fail, and the part of the network where these sensors are used will be inefficient. That's why reducing energy consumption is one of the most important research topics at IoT. Also, since data transmission consumes the most energy, the use of appropriate routing protocols to send data reduces energy consumption. Energy consumption is important to them, so IoT-related standards and research efforts have focused more on conserving or saving energy. Many nodes in Internet of Things are moving, and due to the dynamic nature of nodes in these networks, paths are continually changing, which will cause frequent changes in network topology. For sending data, there is a need to discover new paths. Other problems, such as low communication width due to the systemic nature of these networks and low power nodes, are other factors affecting the efficiency of routing algorithms in these networks. Therefore, these networks need to design their routing algorithms. That's why the issue of routing is so essential. Because the performance of wireless networks is highly dependent on network lifespan, the protocols provided in these types of networks should take into account the increase in network lifespan [15–23].

One of the solutions for reducing power in WSN is to apply a proper protocol for sending and receiving data [24]. Routing protocols are an example of energy-efficient protocols in WSN that are divided into three general categories: flat, location-based, and hierarchical routing protocols [25]. In flat routing protocols, all nodes work the same [26]. In the second method, the location of the node is used for routing [27]. In the third method, the network is divided into a number of clusters. In the clustering method, the nodes associated with the sensors are subdivided into specific groups called clusters. Each cluster has its own cluster head (CH), and other sensors are known as cluster members. Each of the sensors transmits the sensed information to their respective CH, and the CHs transmit the information gathered through a single path or multiple paths to the central node [28].

So far, much research has been done into data transmission routing in wireless sensor networks. Delivering data from the origin to the final destination is usually done in two ways. In the first method, the source sensor transmits its data to the destination (BS) with a few hops and through a number of sensor data interfaces (SDIs). In the second method, the sensor network is clustered, and a cluster-head sensor is specified in each cluster. Non-CH sensors first transmit their data to the CH sensor, and the CH transmit the data to the BS. Given that the position of the sensors in the network is not already known, the presence of an appropriate routing protocol for transmitting information is of great interest. The mentioned protocol must perform routing operations independently. The most important parameter in designing each routing protocol for WSN is the amount of power consumption and the network lifespan.

In this paper, we have used two meta-heuristics algorithm for clustering and routing in wireless sensor network. The meta-heuristics algorithms are used to globally search the domain to find the global or near global optimal results.

Advantages of metaheuristics include their simplicity, independency to the problem, flexibility, and gradient-free nature. Common sources of inspiration in the development of meta-heuristics are from physical phenomena, animal behavior, or evolutionary concepts. In addition, meta-heuristics are independent of the nature of the problem, meaning they do not need derivative information of the problem since they use a stochastic approach. This is in contrast to mathematical programming that typically requires detailed knowledge about the mathematical problem. This independency to the nature of the problem renders them a suitable tool for finding optimal solutions for a given optimization problem without being concerned about the nonlinearity types of the problem's search space and its constraints. Another bonus is their flexibility, which allows them to solve any kind of optimization problem without major changes in the algorithms' structure. They treat the problem as a black box with input and output states, and this feature empowers them as a potential candidate for a user-friendly optimizer. In addition, in contrast to the deterministic nature of mathematical methods, they mostly benefit from stochastic operators. Consequently, the probability of entrapment in local optima is reduced compared to conventional deterministic methods. This characteristic also renders them independent to the initial guess of solutions [29].

In this paper, a new two-phase routing method for the lower layer of the Internet of Things (WSN) is designed to reduce power consumption, increase network life, and improve service quality. In each phase, the service quality is guaranteed by new mechanisms and methodologies. In the first phase, for large-scale network management, the sensor nodes are primarily clustered by the meta-heuristic Whale Optimization Algorithm (WOA-C). The best CH is selected. To increase the service quality for clustering, high-energy, less distance to BS and more neighbors' parameters are used in the function. Each node then transmits its data to the CH. The CHs aggregate the received data and then transmit them to the BS using an optimized multi-phase routing method called the Harris Hawks Optimization algorithm (HHO-R). To do this, we use distance and energy factors in the fitness function. It can be seen that the new protocol proposed in this study improves network performance in terms of energy consumption and network lifespan. The main contribution of this work is to reduce the energy consumption using a novel routing method. It inherently considers the two above methods. First, a WOA is used for clustering and selecting the optimal cluster heads, and then a new HHO algorithm is applied to send data from the cluster heads to the BS. The results of the network performance analysis show that the proposed method can improve network performance in terms of energy consumption, network life, and the number of packets delivered.

In the rest of the paper and in Sect. 2, we describe the previous works that has been done in recent years. Chapter 3 describes the WOA and HHO optimization algorithms. Section 4 discusses the proposed method of applying the Whale Optimization Algorithm in the clustering phase and Harris Hawks Optimization algorithm in the data transmission phase. Section 5 presents the results of simulating the proposed method and comparing it with the related work. Finally, Sect. 6 gives concluding remarks and directions for future work.

2 Related work

In the past two decades, the use of the Internet of Things has attracted many scholars in various fields, which inevitably affects different aspects of human life [30, 31]. In the IoT, the goal is to establish a connection between everything. In other words, the Internet is used to establish the conditions necessary for human communication with objects that requires one or more sensors. Sensors need the power to communicate, and therefore batteries are used to power them. However, the most basic limitation of these sensors is their limited power source. For this reason, energy consumption must be reduced. Various methods have been proposed to extend the lifespan of WSNs and, as a result, increase the total network lifetime [31].

In [32], a clustering method is proposed to reduce energy consumption. In this method, the factors of “vote-based measurement” and “transmitting power of sensor nodes” have been used to perform clustering. Then a connected graph is created between the CHs, and the data transfer phase is performed.

One of the available methods is the use of energy-efficient routing protocols introduced by Javaid and colleagues [33]. They developed a method to improve imaging performance, environmental monitoring, and underwater processes, all controlled wirelessly, aimed at optimizing network lifetime, bandwidth limitation, and energy consumption. Their proposed protocol is based on an energy routing called BEAR that runs in 3 steps. In the first step, sensors share information about their position and residual energy. In second step, BEAR uses location information to select neighbor nodes and auxiliary and alternate sensor. The results show the good performance of this method [33]. In the same year, an article by Jian Shen and colleagues explored the improvement of energy consumption in the IoT using wireless sensor networks [34]. Their new protocol was called Energy-Efficient Centroid-based Routing Protocol (EECRP). Their proposed method was based on three main parts: (1) a new clustering method that provides the self-organizing capability for local nodes, (2) a new algorithm to adapt the clusters and CHs to the energy of all nodes, and (3) a way to improve energy consumption over long distances. Specifically, the residual energy of the nodes is used in the EECRP method to calculate the position of the central node. The simulation results show the improvement of this method against the LEACH method [34].

In the year 2019, Muhammad Faizan Ullah and colleagues presented an article in which a three-layer hybrid clustering mechanism is optimized such that in each run, control packets changes between nodes are constrained, and the bottom layer head is selected. The energy of the nodes is divided into levels based on which it is decided whether the cluster nodes need to enter a new phase of CH selection. The proposed mechanism helps limit the transmission of control packets between nodes while energy consumption between nodes is balanced. It is also stated in this paper that the top layer heads are selected by the BS, such that this action reduces backward transmission as much as possible [35].

Yichao Jin and colleagues have proposed content-based routing in IoT networks and its integration into the RPL (Routing Protocol for Low-Power and

Lossy Networks) [36]. Gathering large data, such as images and videos, from such networks often causes traffic. To solve it, this paper proposes the Content-Centric Routing (CCR) method [37]. This method reduces network traffic with the help of data aggregation and thus significantly reduces latency. In addition, after data aggregation, it is possible to eliminate the excess data transmission that reduces energy consumption, which is mainly devoted to wireless communications, thus limiting the battery.

Another efficient way to reduce energy consumption is the Energy-Efficient Content-Based Routing (EECBR) protocol for the Internet of Things. This method was proposed by Chelloug [38]. Internet convergence, WSNs, and RFID systems [39] have led to the concept of IoT.

Some researchers have used meta-heuristic algorithms, such as the Whale Optimization Algorithm (WOA), for clustering and optimal CH selection [40–42]. The WOA is a nature-inspired algorithm that can be easily implemented. This algorithm, with its productivity and exploration features, selects the appropriate CH.

In [43], an ant colony-based routing technique is introduced, in which the ant colony algorithm has been used to choose the best paths in order to send data in the Internet of Things. Then, the environment of the Internet of Things is divided into some areas. Afterward, an appropriate ant colony algorithm is used for each area. Finally, the algorithm performance is shown in terms of delay, bandwidth consumption, throughput, and energy consumption.

The Particle Swarm Optimization (PSO) algorithm is another meta-heuristic algorithm used in routing [44, 45]. In 2012, researchers proposed a PSO-based energy-aware method for CH selection [44]. Clustering nodes by collecting data at the CH has become one of the most important tools for extending network lifetime. In this paper, a PSO method is presented for generating energy-aware clusters by optimal CH selection. The PSO-based method is implemented inside the cluster instead of the BS. In the article [45], another PSO-based method called PUDCRP is introduced. In the PSO-based Uneven Dynamic Clustering Multi-Hop Routing Protocol (PUDCRP), the clusters distribution changes dynamically when some nodes are destroyed. The PSO algorithm is used to determine where the CH nodes are located. This is a PSO-based uneven dynamic clustering method that divides the network into circles of unequal sizes based on nodes distribution. A multipurpose function is introduced to select CH nodes from the loops.

In the article [46], an energy-efficient routing algorithm for IoT is presented with the help of WSN. Due to the non-uniform traffic distribution, an uneven cluster formation scheme is proposed for load balancing. Also, a distributed CH rotation method is proposed for balancing energy consumption in each cluster.

J. Huang [47] and colleagues proposed a Multi-Hop Energy-Efficient Routing protocol based on network clustering to minimize energy consumption. This algorithm optimizes the CH selection process by combining various factors such as nodes energy, nodes location, and network area level.

A large amount of research has been conducted on reducing energy consumption in wireless networks using various routing algorithms [48–57]. In most of the articles above, the main topic discussed is clustering and CH selection, and a repetitive single-step method is often used to send information from the CH to BS, which in

turn causes the premature death of the CHs that are further away from BS. However, in this paper, in addition to optimizing the clustering with the meta-heuristic Whale Optimization Algorithm (WOA), we propose a new multi-hop method for sending data from CHs to the BS. This will reduce power consumption and increase network lifetime. Since the power consumption is directly related to the distance between the nodes and that the more the distance, the higher the energy consumption, a multi-hop routing method called the Harris Hawks Optimization (HHO-R) algorithm is used to transmit data from the CHs that are further way from BS.

3 WOA and HHO meta-heuristic algorithms

In this paper, we used two meta-heuristic algorithms, namely Whale Optimization Algorithm (WOA) and Harris Hawks Optimization (HHO), for clustering and routing in IoT. Each of these two algorithms is explained later in this chapter. Metaheuristic algorithms create and utilize random variables. These techniques can be used to find optimal global results. Simplicity, independence from the problem, and flexibility are the advantages of this kind of algorithm [58]. They are usually inspired by physical phenomena, animal behavior, or evolutionary concepts. With the help of metaheuristic algorithms, optimization problems can be solved without major changes in the algorithm structure. They explain the problem like a black box with input and output values, which contrasts to the exact nature of mathematical algorithms. Using random variables, the likelihood of getting caught in the local optima is reduced [59].

3.1 Whale optimization algorithm (WOA) [29]

Since WOA is theoretically a metaheuristic algorithm, it can be used for optimization issues from a theoretical perspective. This is because it has both extraction and exploration capabilities. It helps the other agents take advantage of the best records in the domain. By adaptive variation of the search vector A , the WOA clustering algorithm can gently move among exploration and exploitation. More specifically, reducing A can lead to exploration for certain iterations ($|A| \geq 1$) while others are assigned to exploitation ($|A| < 1$). It is worthy to mention that the WOA only encompasses two adaptive internal parameters (A & C).

3.1.1 Encircling preys

Whales can detect the place of their preys and catch them. The optimal position in the search area is not recognized by analogy. The WOA method therefore presumes the best candidate solution is the target prey or a nearly optimum condition. This phenomenon is presented in Eqs. (1) and (2):

$$\vec{D} = |\vec{C} \cdot \vec{X}^*(t) - \vec{X}(t)| \quad (1)$$

$$\vec{X}(t-1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \quad (2)$$

where t shows the current iteration, \vec{A} and \vec{C} denote the coefficient vectors, and \vec{X} stands for the position vector. The position vector of the best solution is indicated by \vec{X}^* . It should be noted that \vec{X}^* must be updated if there is a better solution. The coefficient vectors \vec{A} and \vec{C} can be computed as follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \quad (3)$$

$$\vec{C} = 2 \cdot \vec{r} \quad (4)$$

where \vec{a} linearly decreases from 2 to 0 over the course of iterations (in both exploration and extraction phases), and \vec{r} is a random vector with values ranging from 0 to 1.

3.1.2 Bubble-net attacking method (extraction phase)

The bubble-net behavior of humpback whales can be mathematically modeled by:

The shrinking encircling mechanism: This action is achieved by reducing the value of \vec{a} in Eq. (3). It must be noted that the variation range of A decreases by a . More specifically, A is a random number in $[-a, +a]$, and \vec{a} ranges from 2 to 0 over the course of iterations. By selecting random numbers for A in $[-1, 1]$, the new position of a search agent can be determined everywhere among the original positions of the agent and the positions of the best current agent.

Note that the humpback whales swim along a spiral-shaped path and within a shrinking circle at the same time. A probability of 50% was assumed for the whale to select either the shrinking encircling mechanism or the spiral model in order to update the whale position over the course of optimization. The mathematical model is as follows:

$$\vec{X}(t+1) = \begin{cases} \vec{X}^*(t) - \vec{A} \cdot \vec{D} & \text{if } p < 0.5 \\ \vec{D}l \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) & \text{if } p \geq 0.5 \end{cases} \quad (5)$$

In this equation, P is a random number in $[0, 1]$, \vec{D} indicates the distance between the i th whale and the prey (the best solution so far), b is a constant to define the spiral logarithmic form, and l is a random number in $[1, 1]$. The pseudo code of the WOA algorithm is presented in Fig. 2 [29].

In this paper, we use the whale optimization algorithm for the clustering process.

3.2 Harris hawks optimization algorithm (HHO)

In their article, Heidari et al. [60] developed the Harris Hawks Optimization (HHO) algorithm based on the chasing style and cooperative behavior of Harris hawks. In an attempt to surprise their prey, some hawks swooped down on it from different angles. Also, hawks were able to select different chasing patterns depending on various patterns of prey flight and the scenes thereof. Every HHO algorithm consists


```

Initialize the whales population  $X_i$  ( $i = 1, 2, \dots, n$ )
Calculate the fitness of each search agent
 $X^*$ =the best search agent
while ( $t <$  maximum number of iterations)
  for each search agent
    Update  $a, A, C, l$ , and  $p$ 
    ifl ( $p < 0.5$ )
      if2 ( $|A| < 1$ )
        Update the position of the current search agent by the Eq. (1)
      else if2 ( $|A| \geq 1$ )
        Select a random search agent ()
        Update the position of the current search agent by the Eq. (2)
      end if2
    else if1 ( $p \geq 0.5$ )
      Update the position of the current search by the Eq. (5)
    end if1
  end for
  Check if any search agent goes beyond the search space and amend it
  Calculate the fitness of each search agent
  Update  $X^*$  if there is a better solution
   $t = t + 1$ 
end while
return  $X^*$ 

```

Fig. 2 Pseudo-code of the WOA algorithm

of three prominent phases, namely prey exploration, surprise pounce, and multiple types of surprise attack strategies. Each of the above phases is summarized below.

3.2.1 Phase 1: exploration phase

Mathematically modeled, this subsection is chiefly meant to wait, search, and detect the prey. The HHO algorithm is considered as an alternative or the best solution during every step. The position of the Harris hawks can be modeled based on the following equation:

$$X(t+1) = \begin{cases} X_{rand}(t) - r_1 |X_{rand}(t) - 2r_2 X(t)| & q \geq 0.5 \\ (X_{rabbit}(t) - X_m(t)) - r_3(LB + r_4(UB - LB)) & q < 0.5 \end{cases} \quad (6)$$

where X_t is the current position of the Harris hawks and X_{t+1} is its position in the next iteration, X_{rabbit} is the position of the rabbit, X_{rand} is the randomly selected hawk in the current population, r_i , $i = 1, 2, 3, 4$, q are random numbers between 0 and 1, U_b and L_b represent the upper and lower ranges of the variables and finally, X_m is the average position of the hawks, which can be calculated by the following Equation:

$$X_m(t) = \frac{1}{N} \sum_{i=1}^N X_i(t) \quad (7)$$

where the vector X_i is the location of each hawk, and N represents the total number of hawks.

3.2.2 Phase 2: transition from exploration to exploitation phase

The HHO algorithm can exchange the exploration and exploitation phases based on the escaping energy of the rabbits. Also, the following Equation can be used to calculate the energy of each rabbit:

$$E = 2E_0 \left(1 - \frac{t}{T}\right) \quad (8)$$

where E is the escaping energy of the rabbit, T is the maximum size of the iterations and $E_0 \in (-1, 1)$ is the initial energy during each step. The HHO algorithm is also able to judge the position of each rabbit depending on the change trend E_0 (the HHO algorithm starts the exploration phase to explore the location of each prey when $|E| \geq 1$. Otherwise, it exploits the neighborhood of the solutions during the exploitation phase).

3.2.3 Phase 3: exploitation phase

In the HHO algorithm, hawks lay a soft or hard siege to their prey to hunt it softly or hard from various angles based on the remaining energy of their prey. In this situation, the chance r related to a prey (successful escape occurs if $r < 0.5$) determines whether the prey escapes or not. Besides, the HHO algorithm lays a soft siege if $|E| \geq 0.5$. Otherwise, the hard siege will be laid. Taking into account the escaping phenomenon of the prey and chasing strategies of the Harris hawks, the HHO algorithm utilized four strategies to simulate the attacking stage, namely soft siege, hard siege, soft siege with progressive rapid dives, and hard siege with progressive rapid dives. The relation $|E| \geq 0.5$ indicates that the rabbit has enough energy to escape. However, values $|E|$ and r both determine whether the prey makes a successful escape or not.

(1) Soft siege ($r \geq 0.5$ and $|E| \geq 0.5$)

This measure can be written as follows:

$$X(t+1) = \Delta X(t) - E |J X_{rabbit}(t) - X(t)| \quad (9)$$

$$\Delta X(t) = X_{rabbit}(t) - X(t) \quad (10)$$

where $\Delta X(t)$ is the difference between the position vector of the rabbit and the current location in iteration t , $J=2(1-r5)$ is the random jump intensity of the rabbit

during the escaping process, and finally, r_5 is a random number in the range of 0 to 1.

(2) Hard siege ($r \geq 0.5$ and $|E| < 0.5$)

According to this strategy, the following formula can be used to update the present positions:

$$X(t + 1) = X_{rabbii}(t) - E|\Delta X(t)| \tag{11}$$

(3) Soft siege with progressive rapid dives ($r < 0.5$ and $|E| \geq 0.5$)

According to this strategy, the hawks specify their next step using the following Equation:

$$Y = X_{rabbii}(t) - E|JX_{rabbii}(t) - X(t)| \tag{12}$$

Taking into account the LF-based patterns, the hawks dive based on the following rules:

$$Z = X + S \times LF(D) \tag{13}$$

where D is the dimension of the problem and $S_{1 \times D}$ is a random vector. Besides, the levy flight, LF, can be calculated by the following relation:

$$LF(D) = 0.01 \times \frac{\mu \times \sigma}{|v|^{\frac{1}{\beta}}}, \sigma = \left(\frac{\Gamma(1 + \beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\left(\frac{\beta-1}{2}\right)}} \right)^{\frac{1}{\beta}}, \beta = 1.5 \tag{14}$$

where μ and v are random values between 0 and 1.

Therefore, updating the Hawks' positions is considered as the last strategy of this phase, which can be expressed by the following relation:

$$X(t + 1) = \begin{cases} Y & \text{if } F(Y) < F(X(t)) \\ Z & \text{if } F(Z) < F(X(t)) \end{cases} \tag{15}$$

(4) Hard siege with progressive rapid dives ($r < 0.5$ and $|E| < 0.5$)

In this step, the hawks are always in close proximity to the prey. The related behavior can be modeled as follows:

$$X(t + 1) = \begin{cases} Y & \text{if } F(Y) < F(X(t)) \\ Z & \text{if } F(Z) < F(X(t)) \end{cases} \tag{16}$$

where Y and Z can be calculated using the following two formulas:

$$Y = X_{rabbii}(t) - E|JX_{rabbii}(t) - X_m(t)| \tag{17}$$

Inputs: The population size N and maximum number of iterations T

Outputs: The location of rabbit and its fitness value

Initialize the random population X_i ($i = 1, 2, \dots, N$)

while (stopping condition is not met) **do**

Calculate the fitness values of hawks

Set X_{rabbit} as the location of rabbit (best location)

for (each hawk (X_i)) **do**

Update the initial energy E_0 and jump strength J

$E_0 = 2 \times \text{rand}() - 1$, $J = 2(1 - \text{rand}())$

Update the E using Eq. (8)

if ($|E| \geq 1$) **then** ▷ Exploration phase

Update the location vector using Eq. (6)

if ($|E| < 1$) **then** ▷ Exploitation phase

if ($r \geq 0.5$ and $|E| \geq 0.5$) **then** ▷ Soft besiege

Update the location vector using Eq. (9)

else if ($r \geq 0.5$ and $|E| < 0.5$) **then** ▷ Hard besiege

Update the location vector using Eq. (11)

else if ($r < 0.5$ and $|E| \geq 0.5$) **then** ▷ Soft besiege

with progressive rapid dives

Update the location vector using Eq. (15)

else if ($r < 0.5$ and $|E| < 0.5$) **then** ▷ Hard besiege

with progressive rapid dives

Update the location vector using Eq. (16)

Return X_{rabbit}

Fig. 3 Pseudo-code of the HHO algorithm

$$Z = Y + S \times LF(D) \quad (18)$$

where $X_m(t) = \frac{1}{N} \sum_{i=1}^N X_i(t)$

The pseudocode of the proposed HHO algorithm is reported in Fig. 3 [60].

In this paper, the HHO algorithm is used for the routing between CHs and the BS. These equations and update mechanisms will later be used in the routing.

4 Proposed method

The routing protocol in the Internet of Things framework presented in this paper is based on clustering using the WOA algorithm and choosing the optimal path using the HHO algorithm. In addition to reducing energy consumption, this protocol applies some criteria for quality of service. The protocol seeks to balance power consumption, which will increase network lifetime. It also includes criteria such as speed, network lifetime, and data transmission rates that are raised

regarding service quality. In this algorithm, clustering is used for network scalability. This process can not only reduce energy consumption but also balance the traffic load and improve the network scalability as the network dimensions grow. In addition, data aggregation can be done in the CHs to reduce data transmitted to BS and improve energy efficiency. Rather than direct routing, we use an energy-aware multi-hop routing method called HHO-Routing (HHO-R) based on the Harris Hawks Optimization (HHO) algorithm to transmit data from CHs to BS. The general steps of the proposed method are illustrated in Fig. 4. As shown in Fig. 4, the proposed method consists of two phases. In the first phase, the network is clustered by the WOA algorithm, and the optimal CHs are selected. Member nodes send their data to the CH. The CHs aggregate the collected data. They then send the data to the base station (BS) using a multi-hop method using the new HHO algorithm. The proposed method is described in four sections: the network model, the energy model, the WOA clustering model (the model of clustering and cluster-head selection based on the Whale Optimization Algorithm (WOA-C)), and finally, the HHO Data Transferring model (the model of routing from CH to BS based on the Harris Hawks Optimization (HHO-R)). In the remainder of this chapter, the above four models are explained.

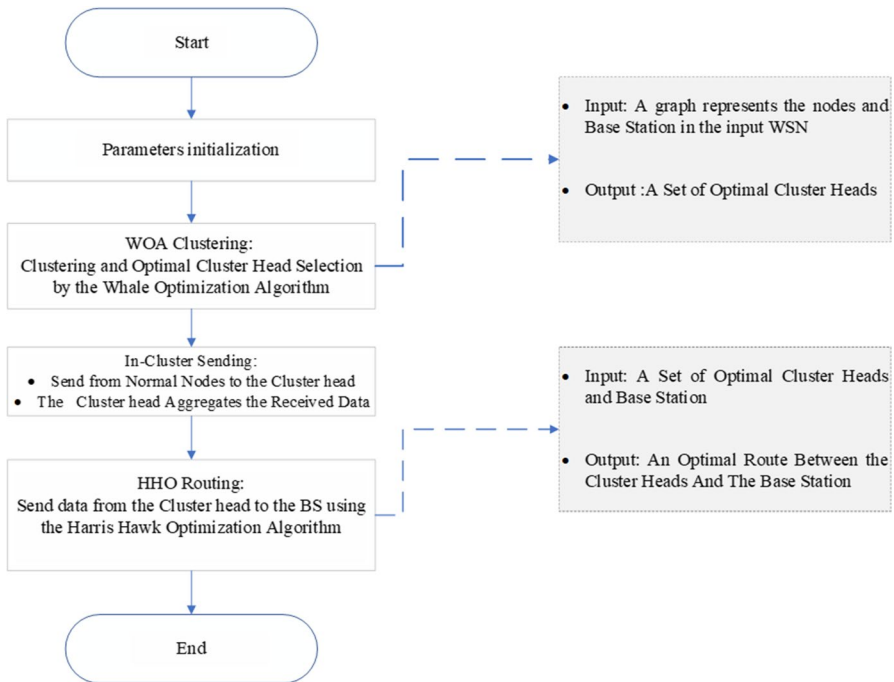


Fig. 4 flowchart of the proposed method

4.1 Network model

The network model consists of a number of N nodes or sensors with one transmitter and a receiver. The nodes are randomly distributed in an environment of 100×100 (m). All sensors are homogeneous and have the same initial energy. After the distribution of the sensors, their position is fixed, and also the sensors are informed of their position by GPS. BS location is also fixed.

4.2 Energy model

In this paper, we use the radio model described in [61] to calculate energy consumption. And it examines the energy consumption in data transmission (ETX), data reception (ERX) and data aggregation (EDA). The energy required to send or receive s bits is calculated from the following equations:

$$E_{amp} = \begin{cases} \epsilon F S d^2 & \text{if } d < d_0 \\ \epsilon T R d^4 & \text{if } d \geq d_0 \end{cases} \quad (19)$$

$$E_{TX}(s, d) = sE_{elec} + sE_{amp} \quad (20)$$

$$E_{RX}(s, d) = sE_{elec} \quad (21)$$

Here d represents the distance between two nodes and, d_0 , EFS , EDA , and E_{elec} are defined in Table 1.

4.3 WOA clustering model

To perform clustering, we use the WOA. The WOA is a meta-heuristic optimization algorithm mimicking the hunting behavior of humpback whales. The main difference between the WOA and the recently published papers is the simulated hunting behavior with random or the best search agent to chase the prey and the use of a spiral to simulate the bubble-net attacking mechanism of humpback whales. The efficiency of the WOA algorithm has been evaluated by solving many mathematical and structural optimization problems. Optimization results demonstrate that WOA is very competitive compared to the state-of-the-art optimization methods.

In a cluster-based routing for WSN, CHs consume more energy. This is due to the additional burden of receiving, collecting, and aggregating data from sensor nodes and transmitting the aggregated data to the BS. Therefore, proper selection of CHs plays an important role in saving energy consumption of sensor nodes to increase the network lifetime. In this paper, we use a method similar to the one proposed in the article [40] with a new and more efficient fitness function for clustering and selecting the CHs. The WOA uses a BS to perform the clustering operation, and as a result, the WOA determines the appropriate clusters for the

network. The BS sends the full network details to all sensor nodes through public broadcast. The message sent by the BS includes the number of CHs, the members associated with each CH, and the number of transfers for this configuration. All nodes receive packets sent by the BS, and so do the clusters. Therefore, the cluster formation stage is completed, followed by the data transfer phase by the second phase of the proposed method.

During implementation, all nodes transmit their data to the BS, whereby the BS compares the average energy of the nodes and states that the nodes whose energy value is higher than the average energy can be selected as CHs. To implement the whale optimization algorithm in clustering, after calculating the average energy at the BS, the nodes with medium-to-high energy are selected. Also, some CHs with minimum cost function will be selected using the whale optimization algorithm.

Now, n cluster-head nodes are considered as the search whales:

$$CH = CH_1, CH_2, \dots, CH_n \quad (22)$$

Considering that the nodes in the network are fixed, the desired whale position (CH) is denoted by CH_i in a two-dimensional space in order to mimic the position of whales in WOA. In other words:

$$Posi(t) = xi(t), yi(t) \quad (23)$$

So, the best position of the search whale is used to determine the best solution, which is used to select the most optimal CH. Also, for implementation, the position of the search whale on the screen is randomly selected first and then replaced with the nearest node. Next, the value of proportion with the corresponding cost function is calculated for all search whales, and the best search whale is selected. They are then updated based on the position of other search whales considering the superior search whale of the WOA parameters. The following are the steps of the proposed method and its pseudo-code to select the best CH based on the whale optimization algorithm.

The CH selection will be based on the fitness function. We also know that in the whale optimization algorithm, the fitness function plays a vital role in identifying and detecting the prey (target node). In this section, we first introduce the parameters of the fitness function and then introduce an efficient fitness function.

4.3.1 Fitness function parameters

Defining the appropriate fitness function plays an important role in the selection of the optimal cluster head by the WOA. To provide a good fitness function, we have used parameters to reduce energy consumption. For this purpose, in the proposed fitness function, we have used more residual energy factors, less distance to the base station, and fewer cluster heads. The function is used to minimize power consumption and extend network lifetime. So, we have to reduce the total transmission distance. Since CHs consume more energy than other nodes to transmit information,

we can reduce energy consumption by reducing the number of CHs. The fitness parameters are described in this section:

- The sum of all distances to BS (total distance): the sum of the distances from all sensor nodes to BS, which we call TD.
- The sum of the distances from normal nodes (non-CH nodes) to the CHs plus the sum of the distances from all CHs to BS which we call NCBD (the distance from non-CHs to CHs and from CHs to BS). To reduce energy consumption, the NCBD should not be too large.
- Residual energy: Since a CH consumes more energy, a node with more residual energy must be selected. The parameter, residual energy, is denoted by E .
- The total number of nodes and CHs: the total number of nodes and CHs are denoted by TN and TCH , respectively.

4.3.2 Fitness function

To become a CH, a node that has the most residual energy should be found. Also, the transmission distance is a factor that we need to minimize. Also, we can also consider reducing the number of CHs in our function. Like reducing the transmission distance, this can also affect energy efficiency because the CHs consume more energy than other nodes. So, our function is a function consisting of all of the above parameters that we define as an Eq. 24.

$$F = E + (TD - NCBD) + (TN - TCH) \quad (24)$$

where, as explained above, E is the residual energy of the node, TD is the sum of the distances from all nodes to BS, $NCBD$ is the sum of the distances from the usual nodes to the CHs plus the sum of the distances from all the CHs to BS. TN , and finally, TCH is the number of all nodes and CHs.

To calculate this function, we have to keep the TN and TD values constant, but the values E , TCH , and $NCBD$ must be changing.

Factors such as higher residual energy, shorter transmission distance ($NCBD$), and fewer CHs increase the fitness value of each individual. Our proposed function attempts to find a suitable solution by increasing the value of the function.

In fact, TD is the distance from all nodes to BS and $NCBD$ is the sum of the distances from all nodes to the CHs and from CHs to BS. The larger the $TD-NCBD$ difference value, the greater the reduction in the communication distance of nodes to BS, which we performed by clustering. The result is a reduction in energy consumption for communication between the nodes and BS. To normalize the equation, and in order for the distance value not to be too large compared to other parameters, we can divide $TD-NCBD$ by TD , which is always constant; that is, we can use $(TD-NCBD)/TD$ instead of $TD-NCBD$. Now that the equation is normal, we can multiply it by a constant number, which makes the distance parameter more important for the equation.

The last part of the relationship concerns the importance of the number of CHs. As mentioned earlier, cluster-head nodes consume more energy than other nodes. So, we need to reduce the number of CHs to an absolute minimum. To do so, we deduct the number of clusters per generation from the total number of nodes. The more the difference, the less the number of CHs. To normalize the relationship, we divide this part by the constant TN; that is, we write the third part of the equation as (TN-TCH)/TN. As mentioned above, if we want to increase the effectiveness of the third part, the number of CHs, in computing the relationship, we can multiply it by a constant number.

With these definitions, the fitness function can be defined as follows:

$$F(CH_i) = 10 * E + \frac{(TD - NCBD)}{TD} + \frac{10 * (TN - TCH)}{TN} \tag{25}$$

The best solution is to obtain the highest value of the fitness function and have enough residual energy. In this case, we can say that there are enough neighboring nodes to convert to a cluster-head node. The above definitions are summarized in Fig. 5.

Once the most optimal arrangement of the CHs and the nodes related to each cluster has been determined, the BS transmits the data containing identification information to each node on the network to determine their location.

Algorithm 1	WOA-Based Clustering for Each Round
1:	Input: A graph represents the nodes and Base Station in the input WSN
2:	Output: A Set of Optimal Cluster heads
3:	for each round do
4:	Initialize Whale Population (search agents)
5:	Initialize Random Position for each whale
6:	Evaluate Fitness Function Using Eq.25 at Position p
7:	if Fitness(p) > Fitness(X*) then
8:	X*=p
9:	set X* as best solution
10:	end if
11:	while r<rmax do
12:	while i<imax do
13:	Update the values of a, A, C, l, p
14:	if p<0.5 then
15:	Use Shrinking encircling Mechanism Using Eq.2
16:	else if p≥0.5 then
17:	Use Spiral Updating Mechanism Using Eq.5
18:	end if
19:	Update the positions of all search agents
20:	end while
21:	Calculate the fitness function for all whales
22:	Update X* with with the position of best solution
23:	end while
24:	Send the set of Cluster Heads to Base station
25:	end for

Fig. 5 The pseudo-code of the WOA clustering algorithm

4.4 The data transfer model using Harris Hawks optimization (HHO-R)

After the clusters were formed and the best CHs were selected by the whale optimization algorithm, it was now time to send the sensed data to the BS. At this point, the cluster member nodes first transmit their data to the CH. The CH then aggregates the data and transmits it to BS. The above two data transmission stages, namely transmitting data from normal nodes to the CH and from the CH to BS, are explained below.

4.4.1 Data transmission from normal nodes to CHs

Within clusters, each node is associated only with CH. The clusters are designed such that each node can communicate directly with its CH. This type of clustering is referred to as the single-step clustering (Fig. 6). In doing so, we use the Carrier Sense Multiple Access (CSMA) method described in the article [62].

4.4.2 Data transmission from CHs to BS

In inter-cluster communications, data is exchanged between CHs and BSs. The starting point for transmitting data to outside the cluster is CH. This will reduce the impact and latency and will also save energy. Transmitting packets from CH to BS can be done in two ways. The CH can communicate with BS directly (single step) after it has collected the data of the member nodes and performed the necessary processing on them. In this method, depending on the distance from CH to BS, a great deal of energy may be required for communicating with BS. In fact, the CHs close to BS interact with it with low energy consumption and the main load of energy is on the CHs that are further away from BS. These CHs may run out of energy quickly, and as a result, the energy balance in the network is lost. Another drawback of this method is that the network capability to expand is limited by the transmission radius of sensors over the IoT network. This means that the network can only be developed to the extent that the farthest CH can directly communicate with BS. However, in general, it can be said that this method can be useful when the transmission distance is short (Fig. 7).

Another method is the multi-hop transmission of the packets to BS. The networks based on IoT sensors consist of a large number of nodes distributed across the network. The distance between the nodes is not too long. In such a structure, multi-hop communications lead to more efficient energy consumption. In this case, high-power transmission is not required. Also, multi-hop communications overcome some of the effects of signal propagation over long distances and result in an increase in the network lifetime and network expansion capabilities. We also do not have distance constraints in this method. Of course, in the case of the multi-hop transmission, if a repetitive path is used, it will lead to higher energy consumption in the middle CHs and CHs close to BS. The reason is that these CHs should constantly receive data from other CHs and transmit them to BS or CHs close to BS. This is illustrated in Fig. 8.

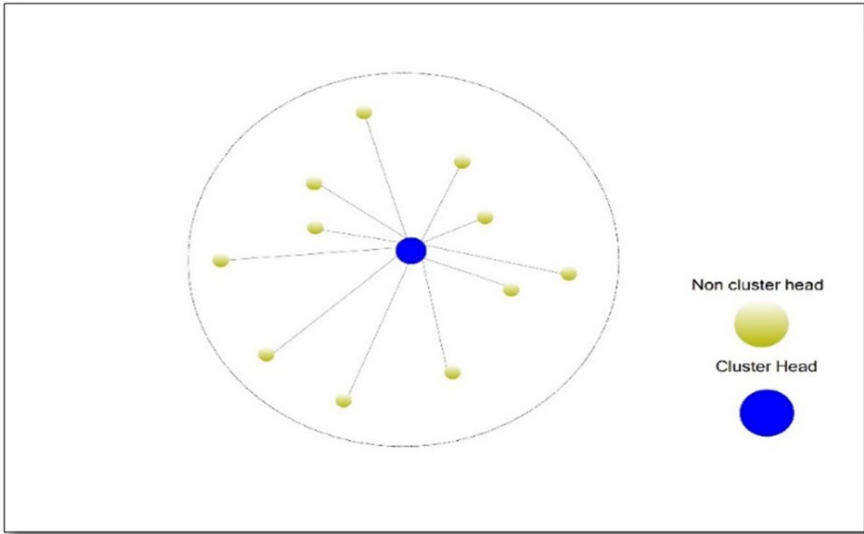


Fig. 6 Data transmission from normal nodes to CHs

For this reason, in this paper, we propose an optimal energy-aware multi-hop routing method using the Harris Hawks Optimization algorithm (HHO-R).

Harris hawks can reveal a variety of chasing patterns based on the dynamic nature of scenarios and the escaping patterns of the prey. This work mathematically mimics such dynamic patterns and behaviors to develop an optimization algorithm.

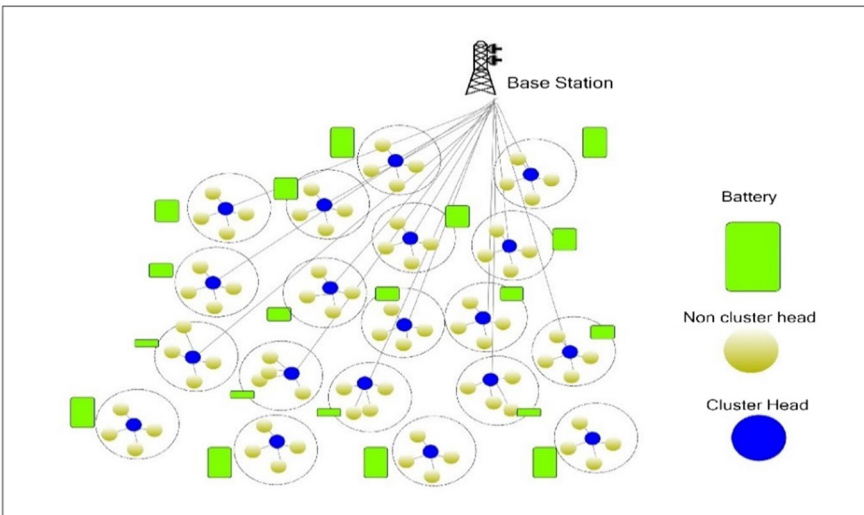


Fig. 7 Direct and single-step data transmission from CHs to BS

The effectiveness of the HHO optimizer is checked, through a comparison with other nature-inspired techniques, on 29 benchmark problems and several real-world engineering problems. The statistical results and comparisons show that the HHO algorithm provides very promising and occasionally competitive results compared to well-established metaheuristic techniques. [60]

All nodes are clustered by the previous phase. After each clustering phase, the location and specifications of the CHs are sent as input for the next phase to select the optimal path. In the second phase, the routing operation is performed by these CHs with the help of the Harris Hawks Optimization in twenty rounds. Then, the clustering phase is performed again by the WOA, and the new CHs are delivered to the second phase. In each of these twenty rounds, the cluster member nodes first send their data as a single hop to the CH. To prevent collisions in this post, we have used TDMA time-division multiplexing. The CH then aggregates the received data and sends the aggregated data to the BS by multi-hop sending using the Harris Hawks Optimization algorithm.

In order to select the optimal path between the CH and the BS by the HHO-R algorithm, the energy level factors of the CH, the distance to the next CH, and the distance to BS are considered as function parameters. The proposed HHO-R algorithm works better for optimization than other methods (Fig. 9).

This algorithm is also centrally controlled and is activated by BS. The proposed algorithm operates on a cyclic basis. Each cycle starts with the setup phase. At the beginning of each setup phase, all the CHs transmit the information related to residual energy and their location to BS. The BS then executes a Harris Hawks Optimization (HHO) algorithm to select the best path that minimizes the function. Below, we describe the various steps of the HHO algorithm for choosing the optimal path.

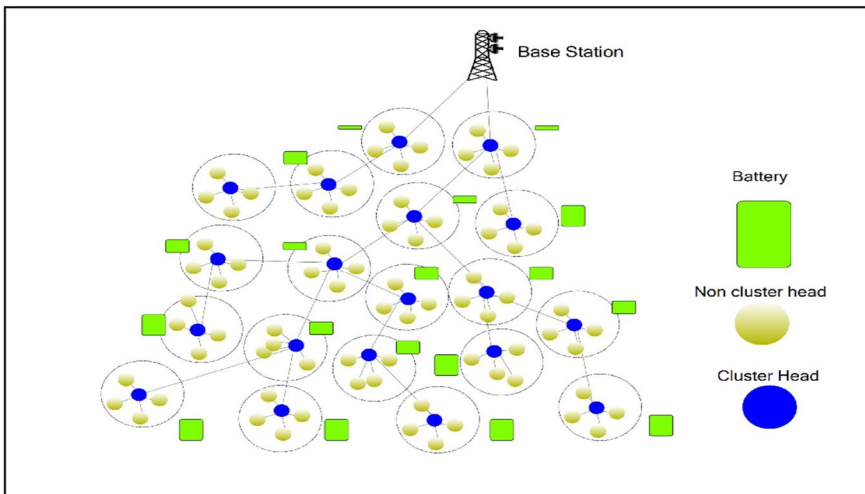


Fig. 8 Multi-hop routing

In the exploration or identification phase, by considering the q value between zero and one, we randomly select one of the two following strategies with equal chance to select the next CH:

$$X_{CH}(t + 1) = \begin{cases} X_{rand}(t) \cdot r1 \mid X_{rand}(t) \cdot 2r2X(t) & q \cdot 0.5 \\ (X_{BS}(t) \cdot X_m(t)) \cdot r3(LB + r4(UB \cdot LB)) & q < 0.5 \end{cases} \quad (26)$$

where $X(t+1)$ is the position of CH in the next iteration, $X_{BS}(t)$ is the position of BS, $X(t)$ is the present position of CH, $r1, r2, r3, r4$ and q are random numbers between zero and one, which are updated in each iteration, UB and LB are the upper and lower range of variables, $X_{rand}(t)$ is the random selection of CH, and X_m is the mean of the current CH positions obtained from relation 27.

$$X_m(t) = \frac{1}{N} \sum_{i=1}^N X_{CH_i}(t) \quad (27)$$

Here $X_{CH_i}(t)$ is the position of each CH in the iteration t denoting the total number of CHs. The CHs energy decreases after each period of data transmission to BS. This is illustrated by relation 28.

$$E_{CH} = 2E_{0CH} \left(1 - \frac{t}{T} \right) \quad (28)$$

where E is the wasted energy of CH, T is the maximum number of rounds, and E_0 is the initial energy of CH. Given the E value obtained, we decided to explore or exploit it. If E is greater than one, exploration is done according to Eq. 8, and if it is less than one, exploitation is done according to Eqs. 11–20. The different modes of exploitation are described below.

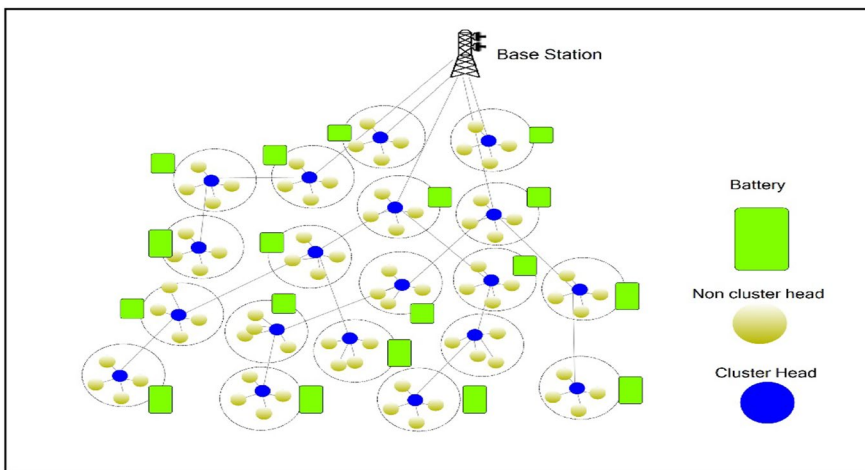


Fig. 9 Optimal multi-hop routing from CH to BS by HHO algorithm

If E is smaller than one, we enter the exploitation phase. Another factor needed to simulate the algorithm is the chance of prey escaping, which is denoted by r . r is a number between 0 and 1, which is selected at random. Finally, in this section, we have two variables, E and r , whose values are between zero and one. A combination of these two variables yields four different states as follows:

- $E > 0.5, r \geq 0.5$ Soft besiege: Using Eqs. 9, 10
- $E < 0.5, r \geq 0.5$ Hard besiege: Using Eq. 11
- $E \geq 0.5, r < 0.5$ Soft besiege with progressive rapid dives: Using Eqs. 12, 13, 14, 15
- $E < 0.5, r < 0.5$ Hard besiege with progressive rapid dives: Using Eqs. 16, 17, 18

After performing the above steps and applying the relationships related to exploration and exploitation, the best path is selected by choosing the optimal CHs. In the Harris Hawks Optimization (HHO) algorithm, the fitness function also plays an important role in prey identification. We use the following parameters in the fitness function:

- *The distance* The distance between CH and BS is denoted by $D_{CH, BS}$.
- *The CH residual energy* Due to the higher energy consumption of the CH, a CH with more residual energy should be selected. The residual energy parameter is denoted by E_{CH} . The fitness function used to select the next CH is as follows:

$$F = \alpha E_{CH_i} + \beta \frac{1}{D_{CH_i, BS}} \quad (29)$$

Algorithm 2	Routing Based HHO For Each Round
1:	Input: A Set of Optimal Cluster Heads and BS
2:	Output: An Optimal Route Between the Cluster Heads and The Base Station
3:	for each round do
4:	Initialize the random Population X_i ($i=1, 2, \dots, N$)
5:	Initialize hawks' population
6:	while (X_{ch} is not equal to BS) do
7:	Evaluate Fitness Function of Cluster Heads Using Eq.31
8:	Set $X_{Base\ Station}$ as the location of Base Station (best location)
9:	for (each cluster head (X_{CH_i})) do
10:	Update the initial energy E_0 and jump strength J
11:	Update the ECH using Eq. 30
12:	if ($ E \geq 1$) then ▷ Exploration phase
13:	Update the location vector using Eq. (28)
14:	if ($ E < 1$) then ▷ Exploitation phase
15:	if ($r \geq 0.5$ and $ E \geq 0.5$) then ▷ Soft besiege
16:	Update the location vector using Eq. (11)
17:	else if ($r \geq 0.5$ and $ E < 0.5$) then ▷ Hard besiege
18:	Update the location vector using Eq. (13)
19:	else if ($r < 0.5$ and $ E \geq 0.5$) then ▷ Soft besiege with progressive rapid dives
20:	Update the location vector using Eq. (17)
21:	else if ($r < 0.5$ and $ E < 0.5$) then ▷ Hard besiege with progressive rapid dives
22:	Update the location vector using Eq. (18)
23:	end while
24:	end for
25:	Send the Data Across the best route to Base station
26:	end for

Fig. 10 The pseudo-code of optimal path selection using the Harris Hawks Optimization (HHO-R)

Alpha and Beta are randomly selected from zero to 1. The CH energy, E , is the residual energy of the i -th CH and $D_{CH_i, BS}$ is the distance between CH and BS. The best solution is obtained when a function with the highest value is achieved. After the CH has transmitted its data to BS, the previous steps will be repeated for transmitting the next data. The mechanism of transmitting data from the CH to BS by the Harris Hawks Optimization (HHO) algorithm is illustrated in Fig. 10.

4.5 Computational complexity analysis

The computational complexity of an optimization algorithm is presented by a function relating the running time of the algorithm to the input size of the problem. For this purpose, Big-O notation is used here as common terminology [70]. In this paper, we used two WOA and an HHO for clustering and routing. In the following, we will discuss the time complexity of these two algorithms.

- WOA complexity

The time complexity of WOA is divided into three aspects: whale population initialization, fitness calculation, and update of population individual position. The time complexity of initialization is $O(N)$ that N shows the number of whale populations, the time complexity of calculating the population fitness is $O(N)$, the time complexity of updating the whale individual position is $O(T \times N) + O(T \times N \times D)$, where T is the maximum number of iterations and D is the dimension of specific problems. Therefore, computational complexity of WOA is $O(N \times (T + TD + 1))$.

Therefore, the overall computational complexity is defined as:

$$O(WOA) = O(N \times (T + TD + 1)) \tag{30}$$

- HHO complexity

Complexity is dependent upon the initialization, fitness evaluation, and updating of hawks. Note that with N hawks, the computational complexity of the initialization process is $O(N)$. The computational complexity of the updating mechanism is $O(T \times N) + O(T \times N \times D)$, which is composed of searching for the best location and updating the location vector of all hawks, where T is the maximum number of iterations and D is the dimension of specific problems. Therefore, computational complexity of HHO is $O(N \times (T + TD + 1))$.

Therefore, the overall computational complexity is defined as:

$$O(HHO) = O(N \times (T + TD + 1)) \tag{31}$$

Finally, based on the following equation,

$$O[(\alpha f + \beta) + (\mu g + \rho)] = O(f) + O(g) \tag{32}$$

So that the α, β, μ and ρ coefficients are constant values, we conclude that the complexity of the algorithm is generally equal to:

$$\begin{aligned}
 O(\text{algorithm}) &= 100(O(\text{WOA}) + 20 \times O(\text{HHO})) \\
 &= 100(O(N \times (T + TD + 1) + 20 \times O(N \times (T + TD + 1))) \\
 &= 2100 \times O(N \times (T + TD + 1)) = O(N \times (T + TD + 1))
 \end{aligned}
 \tag{33}$$

In the mentioned algorithm, for one execution of WOA function, HHO function is executed twenty times and this operation is repeated one hundred times. As shown, the proposed method has reasonable time complexity. Therefore, HHO routing can be considered an efficient algorithm.

5 Simulation and performance evaluation

In this section, a summary of the results obtained from the proposed method, the results of the simulation, and the comparison of the proposed method with previous methods are given. In the simulation, 100 rounds of clustering were repeated, and 20 rounds of routing were performed for each clustering. This is illustrated in Fig. 11.

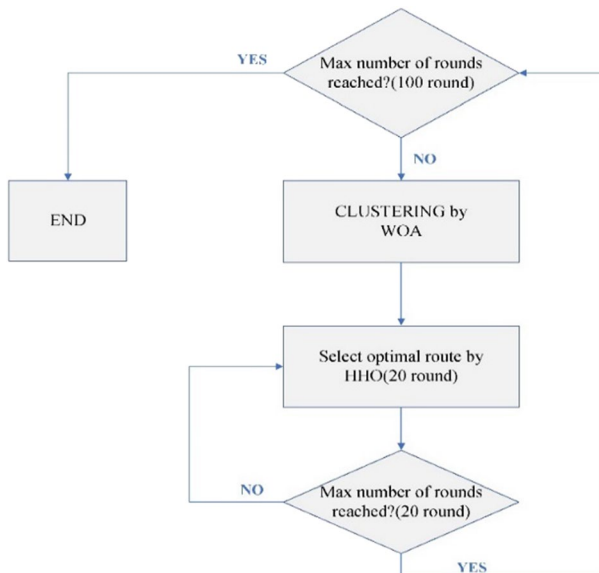


Fig. 11 Number of repetitions of the simulation

Table 1 Simulation parameters

Parameters	Values
Network environment (m)	100*100
Number of sensor nodes	200
BS location	(50, 150)
Structure of arrangement of sensor nodes in the network	Uniform and random playback
Nodes energy	J 0.5
Energy dissipated per bit (Eelec)	50 nJ/b
Energy for data aggregation (EDA)	5 nJ/bit
Amplifier energy for transmitter (ϵ_{TR})	pJ/bit/m ⁴ 0.0013
Amplifier energy for free space (ϵ_{FS})	pJ/bit/m ² 10
Number of simulation rounds	100 rounds of clustering—For each clustering, 20 rounds of data are transmitted (2000 rounds in total)
d0	78m

5.1 Simulation parameters

To simulate the proposed method, a cooja simulator has been used in the Contiki operating system [63, 64]. The structure of the wireless sensor network is based on the parameters in Table 1, much like the parameters in the article [43].

5.2 Simulation and comparison results

For a fair comparison between the different methods, the proposed method has been compared with three different algorithms. To compare and measure the performance of the proposed method, we use the following metrics:

- *The number of Data Delivery* The amount of data sent to the BS is one of the important parameters for routing algorithms.
- *Residual energy* The value of this metric can indicate the energy efficiency of the algorithm. This value decreases with increasing repetition of the algorithm. More residual energy indicates better algorithm performance.
- *The number of surviving nodes* If energy consumption is optimal, more nodes will survive. The lower number of missing nodes also indicates an improvement in algorithm performance.
- *Network Delay* Another parameter for comparing algorithms is the delay time. Delay time is usually defined as "the time between sending a packet and reaching the destination".

As can be seen in Table 1, the experiments were performed with 200 nodes to compare with other methods but to further explain the proposed method and also to show its scalability, we also performed experiments with 50, 100, and 200 nodes.

The number of data delivery rates for 50, 100, and 200 nodes were investigated, and the results indicated a high delivery rate in the proposed network. Given the increase in the number of nodes and the extraction of the results, we find that as the number of nodes in the sensor network increases, the delivery rate in the network will also increase. Figure 10 shows the rate of increase in the number of delivered data.

Figure 12 illustrates two points: Firstly, as the simulation round increases, the packet transmission rate increases as well, and secondly, as the number of nodes increases, the delivery rate increases as well. Increasing the delivery rate in the network is such that the proposed method involves a higher delivery rate than that of the comparative methods of other articles. Comparative methods include UCCGRA [32], PSO-SD [44], PUDCRP [45], EECRA [46], EEMRP [47].

In Fig. 13, we can see that the proposed method performs better and shows improvement over the other methods.

The amount of residual energy in the network was specified for 50, 100, and 200 sensor nodes in a 100×100 m network at 100, 500, 1000, and 2000 iterations. As the number of nodes in the network increases, the amount of residual energy in the network decreases due to the increase in the number of transmitted messages. An increase in the number of nodes in the network leads to an increase in the overhead resulting from the transmitted messages in the network, which reduces the amount of residual energy in the network. Figure 14 shows the amount of residual energy under different simulation conditions.

According to the results of Fig. 15, we find that the proposed method, compared with the methods of other similar articles, has a higher residual energy level. This comparison is shown in Fig. 15.

In this Figure, we find that the proposed method yields better results and shows improvement over the other methods.

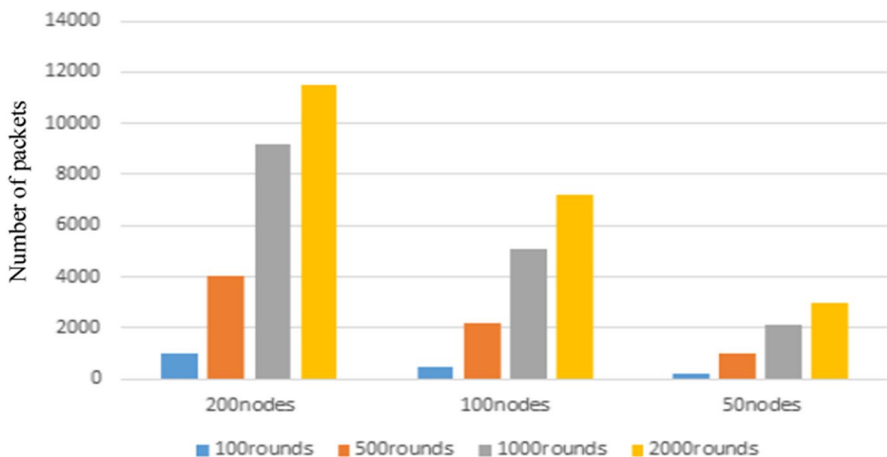


Fig. 12 Number of Data delivery (in bytes) under different simulation conditions

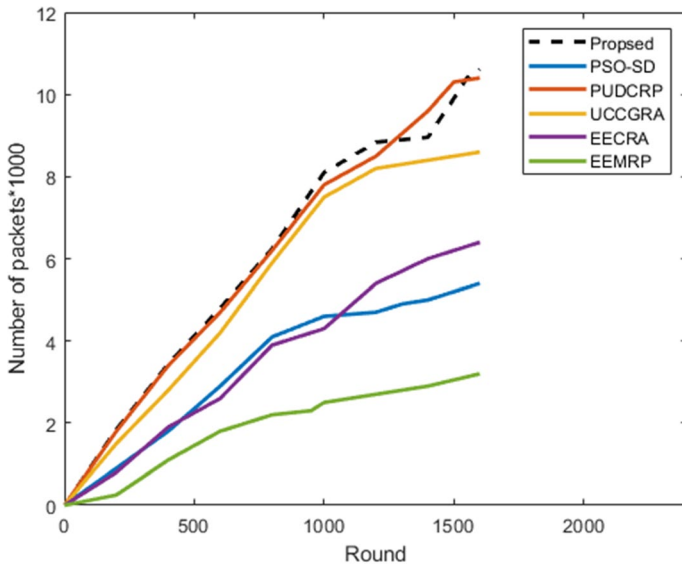


Fig. 13 Data delivery rate with 200 nodes

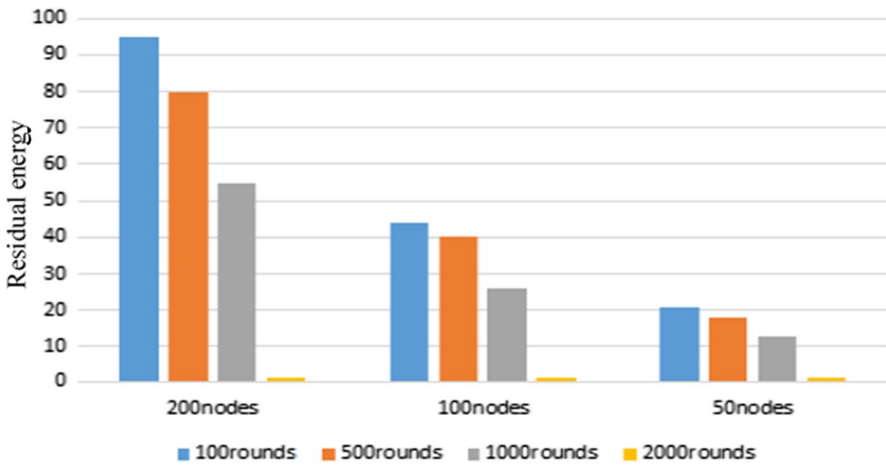


Fig. 14 The amount of residual energy under different simulation conditions

The number of dead nodes is calculated for 50, 100, and 200 nodes in different simulations. Table 2 shows that as the number of nodes increases, the number of dead nodes increases due to the rise in the number of messages. Since the first dead node is one of the most important parameters of any network, the number of dead nodes in different simulation conditions is presented in the form of a table (Table 2).

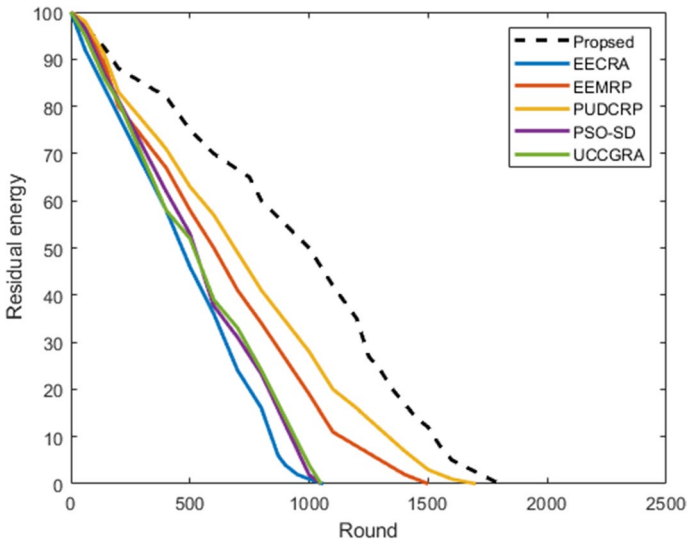


Fig. 15 A comparison of residual energy of the proposed method with those of other papers

Table 2 Number of dead nodes in different simulation conditions

Number of nodes	Number of dead nodes	First dead node
200	200 nodes in 2000 rounds 17 nodes in 1000 rounds 0 nodes in 500 rounds	In the iteration round 870
100	100 nodes in 2000 rounds 13 nodes in 1000 rounds 0 nodes in 500 rounds	In the iteration round 663
50	50 nodes in 2000 rounds 11 nodes in 1000 rounds 0 nodes in 500 rounds	In the iteration round 605

Table 3 The first dead node in different methods

First dead node	Method	Row
In round 300	UCCGRA	1
In round 375	EEMRP	2
In round 305	PSO-SD	3
In round 730	PUDCRP	4
In round 670	EECRA	5
In round 870	Proposed	6

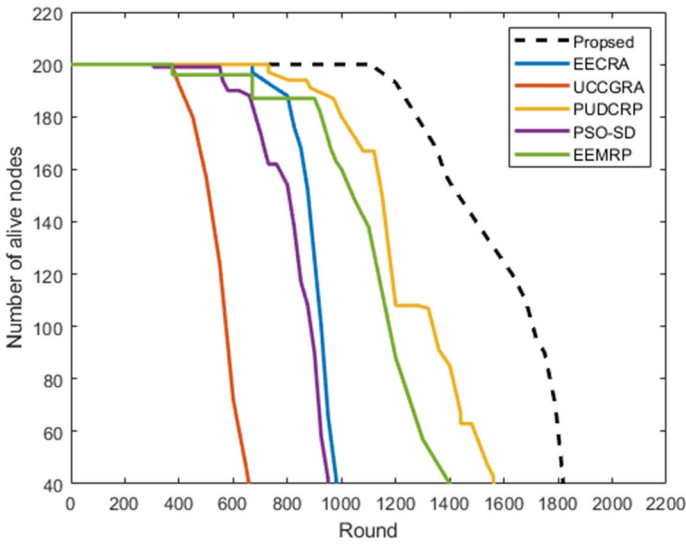


Fig. 16 A comparison of the number of surviving nodes in different methods

In Table 2, we observe that the first dead node for the proposed method is obtained at a high number of rounds (605 for 50 nodes, 663 for 100 nodes, and 870 for 200 nodes), and the smaller number of dead nodes indicates the robustness of the method in data transmission and optimal routing. To compare the proposed method in terms of dead nodes, a comparison is made between the comparative methods, and the results are presented in Table 3.

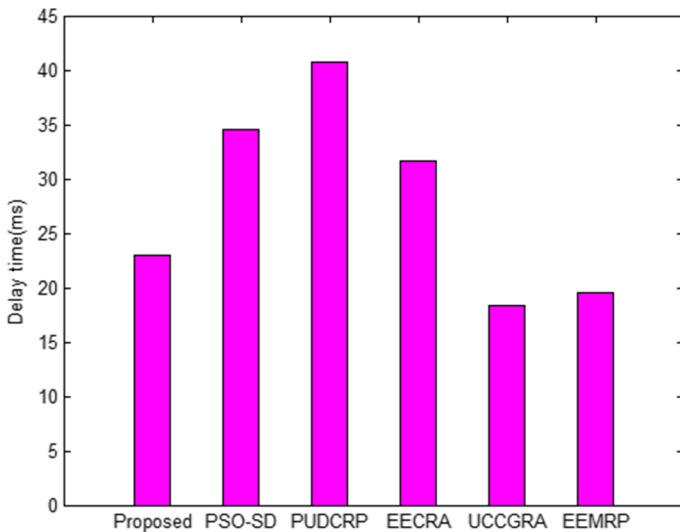


Fig. 17 Comparison of the network delay in different methods

In fact, the proposed method shows less number of dead nodes than the comparative methods. Also, for better comparison, a comparison is made between the living nodes of different methods, which is shown in Fig. 16. By looking carefully at this figure, we find that the proposed method has performed much better than the other methods.

Another parameter for comparing algorithms is the delay time. Delay time is usually defined as "the time between sending a packet and reaching the destination". In this paper, we calculate the delay time in one round as "the time of receiving all packet by the BS in one round". The delay time of different methods is shown in Fig. 17. As shown in the figure, the proposed method has a good performance in terms of delay time.

6 Conclusions and future work

In this study, clustering and routing in the proposed network were investigated using the Whale Optimization Algorithm (WOA) to increase data transmission rate, reduce dead nodes, and control the energy consumption of sensor nodes used in IoT. In the proposed method, the Whale Optimization Algorithm (WOA) and the Harris Hawks Optimization (HHO) algorithm were used in the clustering phase and data transmission phase, respectively. Also, maximum optimization was obtained by the proposed method. It was observed that with this method, the energy consumption, the number of survival iterations for the death of the first node, and the data transmission rate improved. Future work can also use other meta-heuristic methods for clustering and routing to optimize IoT energy consumption. For future studies, other metaheuristic optimization algorithms can be used in clustering and routing WSN. It can also be used in three-dimensional sensor networks for further research by adopting the parameters of the proposed method for 3D. One possible application would be considering the 3D coordinates of the nodes and calculate all the distances in three dimensions.

Author contributions The author wrote all sections of the main manuscripts.

Data availability Data sharing not applicable.

Declarations

Conflict of interest The author declares no competing interest.

References

1. Zhang Q, Zh C, Yang LT, Chen Z, Zhao L, Li P (2017) An incremental CFS algorithm for clustering large data in industrial internet of things. *IEEE Trans Ind Inf* 13:1193–1201. <https://doi.org/10.1109/TII.2017.2684807>

2. Yao X, Farha F, Li R, Psychoula I, Chen L, Ning H (2020) Security and privacy issues of physical objects in the IoT: challenges and opportunities. *Digit Commun Netw*. <https://doi.org/10.1016/j.dcan.2020.09.001>
3. Qing L, Cong L (2016) Efficient cluster routing design under the environment of internet of things based on location. In: 2016 international conference on intelligent transportation, big data & smart city (ICITBS). IEEE, pp 318–323. doi: <https://doi.org/10.1109/ICITBS.2016.63>.
4. Conti M, Kaliyar P, Lal C (2017) REMI: a reliable and secure multicast routing protocol for IoT networks. In: Proceedings of the 12th international conference on availability, reliability and security. ACM, p 8. doi: <https://doi.org/10.1145/3098954.3106070>.
5. Pan MS, Yang SW (2017) A lightweight and distributed geographic multicast routing protocol for IoT applications. *Comput Netw* 112:95–107. <https://doi.org/10.1016/j.comnet.2016.11.006>
6. Mahdavinejad MS, Rezvan M, Berekatain M, Adibi P, Barnaghi P, Sheth AP (2018) Machine learning for internet of things data analysis: a survey. *Digit Commun Netw* 4(3):161–175. <https://doi.org/10.1016/j.dcan.2017.10.002>
7. Al-Turjman F (2019) Cognitive routing protocol for disaster-inspired internet of things. *Futur Gener Comput Syst* 92:1103–1115. <https://doi.org/10.1016/j.future.2017.03.014>
8. Sobral J, Rodrigues J, Rabelo R, Saleem K, Furtado V (2019) LOADng-IoT: an enhanced routing protocol for internet of things applications over low power networks. *Sensors* 19(1):150. <https://doi.org/10.3390/s19010150>
9. Lim J, Ko YB, Kim D, Kim DA (2018) Stepwise approach for energy efficient trust evaluation in military IoT networks. In: 2018 international conference on information and communication technology convergence (ICTC). IEEE, pp 689–692. doi: <https://doi.org/10.1109/ICTC.2018.8539353>.
10. Agre J, Clare L (2000) An integrated architecture for cooperative sensing networks. *Computer* 33(5):106–108. <https://doi.org/10.1109/2.841788>
11. Ernst T, Guillemaud R, Mailley P, Polizzi JP, Koenig A, Boisseau S, Gerbelot-Barillon R (2018) Sensors and related devices for IoT, medicine and smart-living. In: 2018 IEEE Symposium on VLSI Technology. IEEE, pp 35–36. doi: <https://doi.org/10.1109/VLSIT.2018.8510692>.
12. Ben Arbia D, Alam M, Kadri A, Ben Hamida E, Attia R (2017) Enhanced IoT-based end-to-end emergency and disaster relief system. *J Sens Actuator Netw* 6:19. <https://doi.org/10.3390/jsan6030019>
13. Turkmen A, Peng AS, Miller M, Dassow B, Bauer D, Mills L, Batzler R (2019) Development of a home energy monitoring system: a capstone project experience. In: 2019 IEEE power and energy conference at Illinois (PECI). IEEE, pp 1–6. doi: <https://doi.org/10.1109/PECI.2019.8698778>.
14. Chilipirea C, Ursache A, Popa DO, Pop F (2016) Energy efficiency and robustness for IoT: building a smart home security system. In: 2016 IEEE 12th international conference on intelligent computer communication and processing (ICCP). IEEE, pp 43–48. doi: <https://doi.org/10.1109/ICCP.2016.7737120>.
15. Hameed MK, Idrees AK (2021) Cuckoo scheduling algorithm for lifetime optimization in sensor networks of IoT. In: Inventive systems and control: proceedings of ICISC 2021. Springer Singapore, pp 171–187
16. Hameed MK, Idrees AK (2021) Distributed DBSCAN protocol for energy saving in IoT networks. In: International conference on communication, computing and electronics systems: proceedings of ICCCES 2020. Springer Singapore, pp 11–24
17. Idrees AK, Al-Mamory SO, Couturier R (2020) Energy-efficient particle swarm optimization for lifetime coverage prolongation in wireless sensor networks. In: International conference on new trends in information and communications technology applications. Springer international publishing, Cham, pp 200–218
18. Idrees AK, Al-Yaseen WL (2021) Distributed genetic algorithm for lifetime coverage optimisation in wireless sensor networks. *Int J Adv Intell Paradig* 18(1):3–24
19. Idrees AK, Al-Mamory SO, Idrees SK, Couturier R (2021) ETOP: energy-efficient transmission optimization protocol in sensor networks of IoT. In: International conference on new trends in information and communications technology applications. Springer international publishing, Cham, pp 176–186
20. Idrees AK, Couturier R (2022) Energy-saving distributed monitoring-based firefly algorithm in wireless sensors networks. *J Supercomput* 78(2):2072–2097
21. Hameed MK, Idrees AK (2022) Sensor device scheduling-based cuckoo algorithm for enhancing lifetime of cluster-based wireless sensor networks. *Int J Comput Appl Technol* 68(1):58–69

22. Idrees AK, Jawad LW (2023) Energy-efficient data processing protocol in edge-based IoT networks. *Ann Telecommun.* <https://doi.org/10.5121/jgraphoc.2011.310218>
23. Idrees AK, Witwit AJ (2021) Energy-efficient load-balanced RPL routing protocol for internet of things networks. *Int J Internet Technol Secur Trans* 11(3):286–306
24. Heidari E, Movaghar A, Motameni H, Homayun E (2021) Routing in internet of things using cellular automata. In: *International conference on innovative computing and communications*. Springer, Singapore, pp 875–884
25. Akkaya K, Younis M (2005) A survey on routing protocols for wireless sensor networks. *Ad Hoc Netw* 3:325–349. <https://doi.org/10.1016/j.adhoc.2003.09.010>
26. Heinzelman W, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: *Proceedings of the 33rd international conference on system science (HICSS'00)*. Hawaii, USA, pp 1–10. doi: <https://doi.org/10.1109/HICSS.2000.926982>.
27. Tandon A, Srivastava P (2019) Location based secure energy efficient cross layer routing protocols for IoT enabling technologies. *Int J Innov Technol Explor Eng (IJITEE)* 8:368–374
28. Heidari E, Movaghar A (2011) An efficient method based on genetic algorithms to solve sensor network optimization problem. *GRAPH-HOC* 3:18–33. <https://doi.org/10.5121/jgraphoc.2011.310218>
29. Mirjalili S, Lewis A (2016) The whale optimization algorithm. *Adv Eng Softw* 95:51–67. <https://doi.org/10.1016/j.advengsoft.2016.01.008>
30. Kharkongor C, Chithralekha T, Varghese R (2017) Trust and energy-efficient routing for internet of things—energy evaluation model. In: *Proceedings of the 5th international conference on frontiers in intelligent computing: theory and applications*. Springer, Singapore, pp 585–597. doi: https://doi.org/10.1007/978-981-10-3153-3_58.
31. Heidari E, Movaghar A, Motameni H, Barzegar B (2022) A novel approach for clustering and routing in WSN using genetic algorithm and equilibrium optimizer. *Int J Commun Syst* 35(10):e5148. <https://doi.org/10.1002/dac.5148>
32. Xia H, Zhang RH, Yu J, Pan ZK (2016) Energy-efficient routing algorithm based on unequal clustering and connected graph in wireless sensor networks. *Int J Wirel Inf Netw* 23(2):141–150. <https://doi.org/10.1007/s10776-016-0304-5>
33. Javaid N, Cheema S, Akbar M, Alrajeh N, Alabed MS, Guizani N (2017) Balanced energy consumption based adaptive routing for IoT enabling underwater WSNs. *IEEE Access* 5:10040–10051. <https://doi.org/10.1109/ACCESS.2017.2706741>
34. Shen J, Wang A, Wang C, Hung PC, Lai CF (2017) An efficient centroid-based routing protocol for energy management in WSN-assisted IoT. *IEEE Access* 5:18469–18479. <https://doi.org/10.1109/ACCESS.2017.2749606>
35. Faizan Ullah M, Imtiaz J, Maqbool KQ (2019) Enhanced three layer hybrid clustering mechanism for energy efficient routing in IoT. *Sensors* 19:829. <https://doi.org/10.3390/s19040829>
36. Jin Y, Gormus S, Kulkarni P, Sooriyabandara M (2016) Content centric routing in IoT networks and its integration in RPL. *Comput Commun* 89:87–104. <https://doi.org/10.1016/j.comcom.2016.03.005>
37. Jin Y, Kulkarni P, Gormus S, Sooriyabandara M (2012) Content centric and load-balancing aware dynamic data aggregation in multihop wireless networks. In: *2012 IEEE 8th international conference on wireless and mobile computing, networking and communications (WiMob)*. IEEE, pp 179–186. doi: <https://doi.org/10.1109/WiMOB.2012.6379072>.
38. Chelloug SA (2015) Energy-efficient content-based routing in internet of things. *J Compt Commun* 3(12):9. <https://doi.org/10.4236/jcc.2015.312002>
39. Jia X, Feng Q, Fan T, Lei Q (2012) RFID technology and its applications in internet of things (IoT). In: *2nd international conference on consumer electronics, communications and networks (CECNet)*. Yichang, pp 1282–1285. doi: <https://doi.org/10.1109/CECNet.2012.6201508>.
40. Jadhav AR, Shankar T (2017) Whale optimization based energy-efficient cluster head selection algorithm for wireless sensor networks. *arXiv preprint arXiv:1711.09389*
41. Heidari E, Movaghar A, Motameni H, Barzegar B (2022) Reducing energy consumption in IoT by a routing whale optimization algorithm. *Malays J Comput Sci* 35(2):142–157
42. Al-Janabi TA, Al-Rawashidy HS (2017) Efficient whale optimisation algorithm-based SDN clustering for IoT focused on node density. In: *2017 16th annual mediterranean Ad Hoc networking workshop (Med-Hoc-Net)*. IEEE, pp 1–6. Doi: <https://doi.org/10.1109/MedHocNet.2017.8001651>.
43. Said O (2017) Analysis, design and simulation of internet of things routing algorithm based on ant colony optimization. *Int J Commun Syst* 30(8):e3174. <https://doi.org/10.1002/dac.3174>
44. Singh B, Lobiyal DK (2012) A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks. *HCIS* 2:13. <https://doi.org/10.1186/2192-1962-2-13>

45. Ruan D, Huang J (2019) A PSO-based uneven dynamic clustering multi-hop routing protocol for wireless sensor networks. *Sensors* 19:1835. <https://doi.org/10.3390/s19081835>
46. Wang Z, Qin X, Liu B (2018) An energy-efficient clustering routing algorithm for WSN-assisted IoT. In: 2018 IEEE wireless communications and networking conference (WCNC). IEEE, pp 1–6. doi: <https://doi.org/10.1109/WCNC.2018.8377171>.
47. Huang J, Hong Y, Zhao Z, Yuan Y (2017) An energy-efficient multi-hop routing protocol based on grid clustering for wireless sensor networks. *Clust Comput* 20:3071–3083. <https://doi.org/10.1007/s10586-017-0993-2>
48. Abdulzahra AMK, Al-Qurabat AKM (2022) A clustering approach based on fuzzy C-means in wireless sensor networks for IoT applications. *Karbala Int J Mod Sci* 8(4):579–595
49. Abdulzahra AMK, Al-Qurabat AKM, Abdulzahra SA (2023) Optimizing energy consumption in WSN-based IoT using unequal clustering and sleep scheduling methods. *Internet Things* 22:100765
50. Al-Qurabat AKM, Mohammed ZA, Hussein ZJ (2021) Data traffic management based on compression and MDL techniques for smart agriculture in IoT. *Wirel Pers Commun* 120(3):2227–2258
51. Al-Qurabat AKM (2022) A lightweight Huffman-based differential encoding lossless compression technique in IoT for smart agriculture. *Int J Com Dig Sys*. <https://doi.org/10.12785/ijcds/110109>
52. Saedi IDI, Al-Qurabat AKM (2022) Perceptually important points-based data aggregation method for wireless sensor networks. *Baghdad Sci J* 19(4):0875–0875
53. Al-Qurabat AKM, Abdulzahra SA (2020) An overview of periodic wireless sensor networks to the internet of things. In: IOP conference series: materials science and engineering, 928(3). doi: <https://doi.org/10.1088/1757-899X/928/3/032055>.
54. Al-Qurabat AKM, Salman HM, Finjan AAR (2022) Important extrema points extraction-based data aggregation approach for elongating the WSN lifetime. *Int J Comput Appl Technol* 68(4):357–368
55. Saedi IDI, Al-Qurabat AKM (2022) An energy-saving data aggregation method for wireless sensor networks based on the extraction of extrema points. In: AIP conference proceedings. AIP Publishing, vol 2398, no 1
56. Nedham WB, Al-Qurabat AKM (2022) An improved energy efficient clustering protocol for wireless sensor networks. In: 2022 international conference for natural and applied sciences (ICNAS). IEEE, pp. 23–28
57. Abdulzahra SA, Al-Qurabat AKM, Idrees AK (2021) Compression-based data reduction technique for IoT sensor networks. *Baghdad Sci J* 18(1):0184–0184
58. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. *Adv Eng Softw* 69:46–61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>
59. Faramarzi A, Heidarnejad M, Stephens B, Mirjalili S (2020) Equilibrium optimizer: a novel optimization algorithm. *Knowl-Based Syst* 191:105190. <https://doi.org/10.1016/j.knosys.2019.105190>
60. Heidari AA, Mirjalili S, Faris H, Aljarah I, Mafarja M, Chen H (2019) Harris hawks optimization: algorithm and applications. *Futur Gener Comput Syst* 97:849–872. <https://doi.org/10.1016/j.future.2019.02.028>
61. Heinzelman WB, Chandrakasan AP, Balakrishnan H (2002) An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans Wirel Commun* 1(4):660–670. <https://doi.org/10.1109/TWC.2002.804190>
62. Dener M, Bay ÖF (2012) Medium access control protocols for wireless sensor networks: literature survey. *Gazi Univ J Sci* 25:455–464
63. Osterlind F, Dunkels A, Eriksson J, Finne N, Voigt T (2006) Cross-level sensor network simulation with cooja. In: Proceedings 2006 31st IEEE conference on local computer networks. IEEE, pp 641–648. doi: <https://doi.org/10.1109/LCN.2006.322172>.
64. Al-Nidawi Y, Salman N, Kemp AH (2014) Mesh-under cluster-based routing protocol for IEEE 802.15. 4 sensor network. In: European wireless 2014; 20th European wireless conference. VDE, pp 1–7

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.