



## An Advanced Clustering Protocol Based on Modified Differential Search Algorithm for Data Gathering in Wireless Sensor Networks

Kun Nursyaiful Priyo Pamungkas<sup>1</sup>

Waskitho Wibisono<sup>1\*</sup>

Supeno Djanali<sup>1</sup>

<sup>1</sup>*Department of Informatics, Institut Teknologi Sepuluh Nopember, Indonesia*

\* Corresponding author's Email: [supeno@its.ac.id](mailto:supeno@its.ac.id)

---

**Abstract:** Wireless sensor networks (WSN) have been widely applied to monitor objects in harsh and hazardous environments. The limited energy resource in WSN is a major problem. Energy efficiency is required for WSN to operate for a long time. The cluster in WSN is a strategy that can help to extend the life of the network. However, the optimal cluster formation and cluster head (CH) selection is a non-deterministic polynomial-time hard problem. Besides, many clustering-based protocols build clusters after defining the cluster heads among all sensor nodes on the network. This process triggers inefficient energy consumption. Furthermore, the unbalanced load between CH and cluster members (CMs) also affects the network lifetime. This paper proposes an advanced clustering protocol based on a modified differential search algorithm to save energy consumption and extend network life. To form clusters, this protocol determines the optimal number of clusters. In this proposed protocol, CH selection takes place at the cluster level. The process of cluster formation and CH selection is proposed by utilizing a modified differential search algorithm. Furthermore, the rotation system and the determination of relay nodes are proposed to control the load distribution. The proposed protocol's performance is compared with various data gathering protocols for WSN, namely LEACH, PEGASIS, LEACH-GA, and HFAPSO. The experimental results show that the proposed protocol can increase network lifetime in terms of First Node Dies (FND), Half of Nodes Die (HND), and Last Node Dies (LND). On average, the MDSAC protocol increases FND by 1596 rounds, 2456 rounds, 1211 rounds, and 1060 rounds compared to LEACH, PEGASIS, LEACH-GA, and HFAPSO. The MDSAC protocol also increases the average HND by 5764 rounds, 4918 rounds, 5221 rounds, and 4308 rounds compared to other comparison protocols. Even the MDSAC protocol managed to improve the average LND of 7784 rounds, 8739 rounds, 9001 rounds, and 7161 rounds, compared to the four comparison protocols. These results highlight the contribution of the proposed approach in maintaining network lifetime in WSN as one of the important challenges in WSN research and development.

**Keywords:** Clustering, Data gathering, Differential search algorithm, Wireless sensor networks.

---

### 1. Introduction

Wireless sensor networks consisted of numerous sensor nodes with limited resources deployed to gather data from various environments. Because of these purposes, the sensor node is a vital element in WSN. Sensor nodes work continuously to retrieve observational data where sensor nodes are distributed. In general, the sensor nodes are spread over a large area and a harsh environment. In this condition, the maintenance and repair of the sensor node are not accessible. The sensor node also has limited resources, especially energy, because it only uses the

battery as an energy source [1–4]. Thus, it is difficult to replace or to recharge the sensor node's battery when the sensor node energy runs out [5, 6].

Energy is the main challenge in many WSNs research [7]. Energy efficiency is a critical effort in WSN so that sensor nodes can still provide data related to areas or objects of observation. Based on the literature in [8–10], the sensor node's most massive energy consumption is during data transmission. The factor of the transmission distance between the nodes to the sink affects energy consumption. On the other hand, the sensor node also has a limited range. A proper network topology is required for efficient data transmission to be achieved.

The clustering method is a strategy to form a network structure that can support energy efficiency [11, 12]. Many researchers pay great attention to the cluster approach [9]. The cluster strategy's basic idea is to reduce data redundancy because redundant data transmission causes excessive energy consumption. Based on this basic idea, not all sensor nodes need to transmit data to the sink. The formation of a cluster-based topology divides the sensor nodes into two categories: cluster members and cluster heads. Sensor nodes included in the category of cluster members must retrieve data from the observation area and send data to the cluster head. Each cluster has one sensor node as the cluster head. The cluster head is in charge of receiving data from cluster members, processing the received data, and sending data to the sink. The cluster strategy offers advantages besides energy savings [13], namely: (1). prevent collisions; (2). bandwidth savings; (3). increased scalability; and (4). reduce overhead.

Low Energy Adaptive Clustering Hierarchy (LEACH) is a prominent cluster-based protocol in WSN [14, 15]. In the LEACH protocol, cluster formation is preceded by selecting CH. CH selection is made in a distributed manner based on probability and random values. The LEACH protocol offers advantages [16], among others (1). reduce sensor nodes to sink communication costs; (2). avoiding data redundancies; and (3). provides the opportunity for all sensor nodes to become CH at least once. However, the weakness of the LEACH protocol lies in the CH selection scheme, which ignores parameters such as energy and position. Thus, CH can deplete energy quickly and die prematurely. Until now, many researchers have paid great attention to the LEACH protocol [16, 17].

Power-Efficient Gathering in Sensor Information Systems (PEGASIS) is another well-known protocol in WSN [18]. PEGASIS offers a different cluster concept from LEACH. PEGASIS only builds one cluster in the form of a chain. In each round, one sensor node is randomly selected as the leader in the chain to communicate to the sink [3]. Each sensor node in the chain can transmit and receive data from neighboring nodes. Furthermore, each sensor node performs data aggregation and sends data to the next neighboring node. Thus, the leader receives the aggregated data from the chain members. However, this mechanism can trigger a long delay when the number of sensor nodes increases, or the network area size increases. Like LEACH, researchers are still interested in developing a variant of the PEGASIS protocol [8].

In general, clustering consists of three processes: the CH selection process, merging sensor nodes to

CH, and sending data by cluster members to CH and CH to sink. Data transmission from the CH to the sink can be single-hop or multi-hop. The focal point of the cluster strategy is the CH selection process. Suppose this selection process results in a suboptimal set of nodes as CH. In that case, energy efficiency will be difficult to achieve, and the nodes cannot operate for a long time. A cluster-based protocol should be able to select the correct node as a CH. However, the selection of this CH is an NP-hard problem. Thus, an intelligence approach is needed to overcome this problem. Apart from selecting the CH, another process that needs attention is the process of finding and merging sensor nodes to the closest CH. The process of broadcasting by CH and receiving messages by sensor nodes requires energy. Thus, if this process is not appropriately considered, this process has the potential to increase energy consumption. This energy consumption increases as the size of the network area grows.

Another problem that can arise in this cluster strategy is the CH rotation. Sensor nodes cannot operate as CH continuously because CH has a more significant load than sensor nodes as cluster members. Nodes operating as CH in each round continuously result in uneven load distribution and rapid CH death. However, a balanced CH rotation requires an optimal number of clusters and a relatively stable cluster. Therefore, a new protocol is needed to overcome these problems.

In nature, some living organisms have shown the ability to sustain life even under limited conditions. They work in a herd or colony system in a distributed and intelligent manner to achieve their global goals. Collective behaviour such as birds, bees, and ants have specific patterns that are adopted by intelligent algorithms, including Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Artificial Bee Colony (ABC). In the context of WSN, these intelligent algorithms can help solve various problems [17-19]. Several intelligent algorithms have become the basis of protocols in WSN, such as PSO in [20], ACO in [21], and ABC in [22]. Each of these intelligent algorithms has certain advantages that can affect protocol performance. The Differential Search algorithm (DSA) is a relatively new bio-inspired algorithm. This algorithm is inspired by the migration patterns of the herd of living organisms. Based on the research in [23], it has shown good performance compared to some other algorithms. However, the DSA is sensitive to random number generator so that the performance of the DSA can decrease in some instances.

This paper proposes an advanced clustering protocol based on a modified differential search

algorithm to minimize energy consumption and extend network life. This proposed protocol estimates the optimal number of clusters to build clusters. Instead of involving all sensor nodes on the network, this proposed protocol selects CH in each cluster. We formulate a fitness function to select the best CH. This fitness function ensures that the selected CH does not run out of energy when running all tasks and that the CH location is not close to the network area boundary. The cluster formation and selection of this CH utilizes a modified differential search algorithm. Besides, the proposed protocol regulates CH rotation to distribute the load. A dual-hop communication approach is also adopted by selecting relay nodes to share the load with CHs positioned further away from the sink. In these strategies, the proposed protocol can minimize the consumption of energy and prolong network lifetime.

This paper is divided into five parts. Section 2 describes studies related to the proposed protocol. Section 3 contains the explanation of the radio, energy, and network models used in this study. In Section 4, we describe the mechanism for the proposed protocol. Section 5 explains the simulation environment, simulation parameters, and discusses the simulation experimental results. Finally, Section 6 concludes the results of the research.

## 2. Related work

Hemavathi and Nandakumar proposed a clustering protocol based on PSO optimization [24]. This protocol focuses on simplifying the exploration process for the PSO algorithm for selecting cluster heads. Both researchers argue that efforts to minimize iteration on the PSO are needed to reduce overhead and avoid data traffic congestion. Before the cluster was formed, this protocol selected several sensor nodes as CH. The residual energy level of each sensor node determines the probability of being CH. After the CH is selected, the clusters are formed on the network. Although this protocol has demonstrated its superiority over LEACH, the period of stability is still relatively short.

Srikanth and Prasad presented a clustering protocol based on fuzzy k-means and a centralized midpoint algorithm called Energy Efficient Multi-hop Routing Clustering Protocol (EEMRCP) [25]. This protocol selects CH based on two factors, namely residual energy, and intra-cluster distance. Whereas the CH to sink distance is used to determine whether the CH sends data to the sink directly or multi-hop. Multi-hop communication can assist CH that is far from the sink. However, this strategy becomes inefficient when the density is very low, and

the network area is vast.

Yadav, Kumar, and Vijendra proposed enhancing the LEACH protocol by utilizing the Particle Swarm Optimization algorithm [26]. The new protocol resulting from the integration of LEACH and PSO is named LEACH-PSO. The LEACH-PSO protocol makes use of the PSO algorithm to determine the optimal number of clusters and CH. The energy consumption of the nodes becomes the basis for selecting the optimal CH. LEACH-PSO performs better than LEACH. However, the LEACH-PSO protocol can produce a fluctuating number of clusters in each round. As a result, the load distribution can be uneven.

In 2019, Karthick and Palanisamy had developed a clustering protocol based on the Krill Herd (KH) algorithm [27]. This protocol focuses on optimal cluster head selection. The selection of cluster heads utilizes the KH algorithm by taking into account the intra-cluster distance and the ratio between all sensor nodes' energy with the remaining energy of the CH candidates. This selection scheme allows the sensor node, which has the most residual energy and the closest distance to other nodes, to continuously be selected as CH. After the CH selection is complete, the clusters are formed. This protocol proves its superiority in energy efficiency and increased network lifetime compared to LEACH and GA based protocols. However, there are times when the KH algorithm is unable to maintain a balance between local search and global search. So, the KH algorithm can be trapped in the optimum locale. Furthermore, the KH algorithm does not converge on optimal results.

Ouldzira, Lagraini, Mouhsen, Chhiba, and Tabyaoui proposed an extension of the LEACH protocol under MG-LEACH [28]. This protocol has three phases: the set building phase, the setup phase, and the steady-state phase. As with LEACH, CH selection and cluster formation take place in the setup phase. The initial energy and residual energy of the sensor nodes are parameters that affect the threshold value. Considering the energy factor in MG-LEACH is an effort to ensure that the selected CH has sufficient energy to complete the task. However, MG-LEACH did not take the location into account so that a sensor node that was far from the sink could be selected as CH.

Radhika and Sivakumar developed a new protocol that combines the LEACH protocol with micro genetic algorithms [7]. This new protocol is named  $\mu$ GA-LEACH. The  $\mu$ GA-LEACH protocol has three phases: the preparation phase, the setup phase, and the steady-state phase. The CH selection process takes place in the preparation phase using the

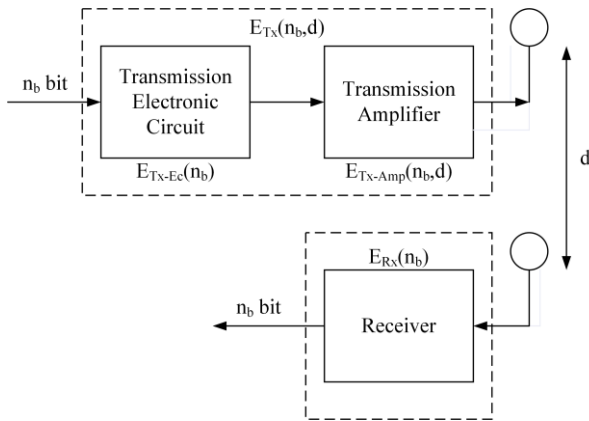


Figure. 1 Radio model

$\mu$ GA algorithm. The average distance of the sensor node to the CH candidate, the average distance of the CH candidate to the sink, and the ratio of all sensor nodes' residual energy to the remaining energy of the CH candidate are the basis for consideration in determining CH. The  $\mu$ GA-LEACH protocol has shown better performance than LEACH, LEACH-C, LEACH-GA, and GADA-LEACH. However, this protocol does not estimate the residual energy of the nodes when they become CH. Thus, CH may die before completing his task.

In 2020, Pitchaimanickam and Murugaboopathi proposed developing a LEACH-C protocol based on hybrid firefly and particle swarm optimization (HFAPSO) [29]. This protocol was proposed to overcome the disadvantages of CH selection in LEACH-C. To produce optimal CH, the HFAPSO protocol combines the advantages of FA and PSO. There are two factors used by HFAPSO to assess the feasibility of a sensor node as CH, namely the intra-cluster communication cost factor and the residual energy factor. The HFAPSO protocol has shown good performance in reducing energy consumption and increasing network lifetime compared to LEACH-C protocol and clustering protocol based on FA. However, the HFAPSO protocol does not focus on the CH load rotation system. Thus, a node can be selected as CH repeatedly.

### 3. System model

#### 3.1 Energy and radio model

In this study, the radio model used refers to the model proposed by [14] because many studies use it [30]. This radio model consists of three parts: the transmitter, the amplifier, and the receiver. Fig. 1 illustrates the radio model used in this study.

In this radio transmitter and receiver, electronic components need the energy to process 1-bit data,

Table 1. The symbol list in system model

Symbol	Meaning
$E_{Tx}(n_b, d)$	Required energy to transmit $n_b$ bits of data over a distance $d$ .
$E_{ec}$	Consumed energy in the radio electronic circuit.
$E_{fs}$	Free space consumed energy in the radio amplifier.
$E_{mpf}$	Multi-path fading consumed energy in the radio amplifier.
$n_b$	Length of data in bit.
$d$	Distance between transmitter and receiver.
$d_t$	Threshold distance.
$E_{Rx}$	Required energy to receive data.
$E_{da}$	Required energy to aggregate data.
$E_{DA}$	Total required energy to aggregate $n_b$ bits of data.

namely  $E_{ec} = 50$  nJ. Whereas in the amplifier section, the energy consumed depends on the propagation model. This study uses two propagation models, namely, free space and multi-path fading. The distance between the transmitter and receiver influences the choice of this propagation model. If the distance between transmitter and receiver, i.e.,  $d$  is less than the  $d_t$  distance threshold, then the free space propagation model is selected. It affects energy consumed by the amplifier is  $E_{fs} = 10$  pJ/bit/m<sup>2</sup>. In contrast, the multi-path fading propagation model is chosen, and the energy consumption in the amplifier section is  $E_{mpf} = 0.0013$  pJ/bit/m<sup>4</sup>. Thus, when the sensor node transmits  $n_b$ -bit in a distance of  $d$ , the sensor node will consume an amount of energy:

$$E_{Tx}(n_b, d) = E_{Tx-Ec}(n_b) + E_{Tx-Amp}(n_b, d) \quad (1)$$

$$E_{Tx}(n_b, d) = \begin{cases} n_b(E_{ec} + E_{fs}d^2), & d < d_t \\ n_b(E_{ec} + E_{mpf}d^4), & d \geq d_t. \end{cases} \quad (2)$$

The following formula can assign the  $d_t$  threshold value:

$$d_t = \sqrt{\frac{E_{fs}}{E_{mpf}}}. \quad (3)$$

Meanwhile, the sensor node that receives data consumes energy by:

$$E_{Rx}(n_b) = n_b E_{ec}. \quad (4)$$

If the sensor node receiving data is CH, then CH will spend energy to aggregate data every 1-bit is  $E_{da}$ .

Thus, when CH receives  $n_b$ -bit, the data aggregation process consumes as much energy as:

$$E_{DA}(n_b) = n_b E_{da}. \quad (5)$$

Some of the symbols used in this energy and radio model are listed in Table 1.

### 3.2 Network model

We consider a cluster-based WSN model combining single-hop and dual-hop. All sensor nodes are randomly distributed in the network area. After all sensor nodes have been deployed, the sensor nodes' positions do not change. The energy source of the sensor nodes is a battery of the same capacity. We assume that the sensor nodes have the exact technical specifications so that each sensor node has the same capabilities in terms of sensing observation objects, processing data, and communication. Data from sensing sensor nodes in the same cluster have a high correlation so that CH can use data aggregation techniques to combine all data from sensor nodes into single data. Each sensor node can communicate with other sensor nodes and sinks directly. Communication between two sensor nodes requires the same energy because communication is symmetrical.

Furthermore, each sensor node has the capabilities to know their location as well as the sink's location. The sink has knowledge about all sensor nodes in the network area. We also assume that the location of the sink is stationary. Besides, the sink has rich energy resources because it uses external energy sources.

## 4. The proposed protocol

This study proposes a protocol with the name Modified DSA for Clustering, which is abbreviated as MDSAC. The MDSAC protocol combines a

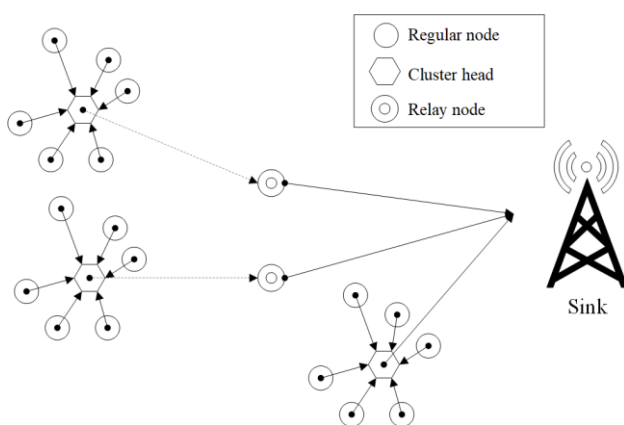


Figure. 2 The topology of the MDSAC protocol

centralized and distributed approach. A centralized approach is needed when the MDSAC protocol runs the cluster formation process on the network. In this process, the sink requests important data from sensor nodes, such as location coordinates and energy levels. Then, the sink does the computation after all the data is collected. Meanwhile, the distribution approach is taken CH selection because it performs this process at the cluster level. The MDSAC protocol is proposed and developed to achieve two important goals, namely: (1). durable network; and (2). high scalability.

The MDSAC protocol divides sensor nodes into three categories, namely: regular nodes, relay nodes (RNs), and CHs. We use the term round denotes operating time in WSN. Each round in the MDSAC protocol consists of two phases, namely the setup phase and the data-gathering phase. In the setup phase, the sink constructs the WSN topology by forming a virtual backbone if needed and cluster. Fig. 2 illustrates the MDSAC protocol topology. In detail, the setup phase consists of several stages, as follows:

1. Broadcast NTC\_Request message: sink broadcasts the message to all sensor nodes. The message sent by the sink contains a request to the sensor node to send ID, position, status, and energy information. Each sensor node also replies to the NTC\_Reply message to the sink containing the ID, position, status, and energy.
2. Network topology construction: after the sink receives messages from all sensor nodes, the sink determines whether creating a virtual backbone is required. If needed, the sink selects multiple sensor nodes as relay nodes. On the other hand, the sink calculates the optimal number of clusters directly. This calculation refers to Eq. (21). Based on the results of these calculations, sinks form clusters in the network. Section 5.2 describes cluster formation in more detail.
3. Transmit NC\_Msg: after the sink completes the formation of the network topology, sink sends an NC\_Msg message to each sensor containing the sensor node ID along with the sensor node cluster-ID and VCHs position. Each sensor node that receives the message will update its configuration.
4. CH selection: each cluster selects CH sequentially. The CH selection mechanism that takes place in each cluster is described in section 5. The sensor node selected as CH sends a CH\_ADV message containing the CH ID to other sensor nodes in the same cluster.

The data-gathering phase is the phase of collecting and transmitting data. In this phase, the



sensor nodes collect data from the observation area, and then the sensor nodes send it to the CH based on the TDMA schedule. Fig. 3 presents the TDMA schedule that applies to MDSAC protocol. This scheduled data transmission can avoid data collisions. Thus, the data received by CH is correct and intact. Next, the data received from all cluster members aggregated by CH. The aggregation process is needed to summarize all received data into a single data. Thus, data traffic to the sink is not congested. Although this process also requires energy. After CH complete aggregation, CH transmits data from the results of this aggregation process to the sink.

#### 4.1 Relay node's selection as the virtual backbone

In the MDSAC protocol, the existence of relay nodes is conditional. The relay node has a function only to reduce the CH load, which is positioned away from the sink. Thus, not all CHs send data to the sink via relay nodes. This concept differs from other protocols that use relay nodes to forward data from each CH to the sink. This continuous deployment of relay nodes is good in terms of sharing the CH load. However, the relay nodes expend energy to receive data from CH.

We define the set of relay nodes as  $NR = \{nr_1, nr_2, \dots, nr_n\}$ . This set of relay nodes forms a virtual backbone for several CHs. This virtual backbone is formed when there is a sensor node with a distance of  $d_{tos} \geq d_t$ . Suppose the sensor node distance to the sink is  $d_{tos}$ , the energy available at the sensor node is  $E_{av-ns}$ , and the average energy availability across all sensor nodes that are still operating in the network  $\overline{E_{av-ns}}$ , then the MDSAC protocol chooses sensor nodes as relay nodes on condition as follows:

$$if\ d_{tos} \geq 50 \ \& \ d_{tos} \leq 55 \ \& \ \frac{E_{av-ns}}{\overline{E_{av-ns}}} \geq 1 \Rightarrow \text{Role} = NR \quad (6)$$

The distance requirement for sensor nodes to sink as the basis for selecting the relay nodes refers to research [31]. Furthermore, the sensor nodes selected as NR send announcement messages to other sensor nodes.

#### 4.2 Cluster topology construction

The MDSAC protocol integrates sensor nodes into clusters before CH selection. This strategy is different from other clustering protocols. Before the cluster is established on the network, the sink assigns an optimal number of clusters after the sink receives an NTC\_Reply reply message from the sensor nodes.

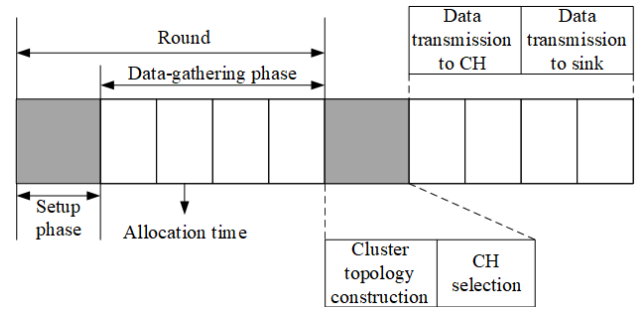


Figure. 3 TDMA schedule in MDSAC protocol

The optimal number of clusters needs to be sought because it also affects energy efficiency in the network.

Assume that  $H$  sensor nodes are still alive, namely  $NS_L = ns_1, ns_2, ns_3, \dots, ns_h$  spread over the area  $S \times S$  observation area. The sink is placed right in the center of the observation area with a transmission range distance of  $R_s$ . Another assumption is that if there are  $n_c$  clusters formed on the network, then each cluster will have an average number of sensor nodes, which is  $\frac{H}{n_c}$ , with the number of cluster members being  $\frac{H}{n_c} - 1$  and one CH. Each cluster also occupies an area whose average area is:

$$\overline{L_C} = \frac{S^2}{n_c} \quad (7)$$

If we assume that the shape of the cluster is the circle, then we can determine the radius of the cluster, i.e.:

$$R_C = \frac{S}{\sqrt{\pi n_c}} \quad (8)$$

Because in general, the sensor nodes are randomly distributed to the observation area, the density of a cluster is  $\rho(x, y)$ . Thus, the authors in [32] argue that if we assume  $x_{ch} = 0$  and  $y_{ch} = 0$ , then the expected quadratic transmission distance from the cluster member to CH  $E[d_{toCH}^2]$  can be formulated as follows:

$$\begin{aligned} E[d_{toCH}^2] &= \iint ((x - x_{ch})^2 + (y - y_{ch})^2) \rho(x, y) dx dy \\ E[d_{toCH}^2] &= \iint (x^2 + y^2) \rho(x, y) dx dy \\ E[d_{toCH}^2] &= \iint r_c^2 \rho(r_c, \vartheta) dr d\vartheta. \end{aligned} \quad (9)$$

On Eq. (17), the density level  $\rho(r_c, \vartheta)$  is assumed to be constant [32]. Regarding Eq. (8), Eq. (9) can still be simplified to:

$$E[d_{toCH}^2] = \rho \int_{\vartheta=0}^{2\pi} \int_{r_c=0}^{\frac{s}{\sqrt{\pi n_c}}} r_c^3 dr d\vartheta$$

$$E[d_{toCH}^2] = \frac{s^2}{2\pi n_c}. \quad (10)$$

After the expected distance from the cluster member to the CH can be calculated, and the average distance of the sensor node to the sink  $d_{toS}$  is known. Indeed, we can also calculate the total energy consumed in the network at each round.

$$E_{Total} = n_c E_{cluster} \quad (11)$$

$$E_{Total} = \begin{cases} n_b \left( HE_{ec} + HE_{da} + n_c E_{fs} d_{toS}^2 \right) + HE_{ec} + HE_{fs} \frac{s^2}{2\pi n_c}, & d < d_t \\ n_b \left( HE_{ec} + HE_{da} + n_c E_{mpf} d_{toS}^4 \right) + HE_{ec} + HE_{fs} \frac{s^2}{2\pi n_c}, & d \geq d_t. \end{cases} \quad (12)$$

In [27], communication between the CH and sink uses a multi-path fading propagation mode to calculate total energy in the cluster for one round. In this study, the MDSAC protocol considers whether the mean distance to the sink is less than the threshold distance or vice versa. This assumption indeed results in determining the optimum number of clusters. We can obtain the optimum number of clusters by finding the  $E_{Total}$  derivative with regards to  $n_c$ .

$$n_{copt} = \begin{cases} \sqrt{\frac{H}{2\pi} \frac{s}{d_{toS}}}, & d_{toS} < d_t \\ \sqrt{\frac{H}{2\pi} \sqrt{\frac{E_{fs}}{E_{mpf}}} \frac{s}{d_{toS}^2}}, & d_{toS} \geq d_t. \end{cases} \quad (13)$$

After the MDSAC protocol succeeded in determining the optimum number of clusters in the network area, the next process was cluster formation. The cluster formation process is based on a modified Differential Search algorithm (MDSA).

In the process of forming each cluster, we assume there are  $NS_L$  active sensor nodes (this number does not include relay nodes if MDSAC protocol builds a virtual backbone) in the network, which is divided into  $n_{copt}$  clusters. MDSA's task is to determine clusters that can support energy consumption savings. Since the MDSAC protocol forms clusters before CH selection, it is necessary to find the best virtual CH position. The term virtual CH is used as an alternative to CH before the actual CH is selected. We define the virtual CH set as  $VCH =$

$\{vch_1, vch_2, vch_3, \dots, vch_{n_{copt}}\}$ . MDSA begins cluster determination by initializing  $G$  artificial-superorganisms and  $D$  the dimensioning of the problems. Artificial-superorganism refers to the term population in DSA. Every artificial-superorganism has  $n_{copt}$  artificial-organisms. This artificial organism's position is generated randomly within the upper limit of the  $ub$  and the lower limit of the  $lb$ . In this case,  $ub$  is the highest sensor node position, while  $lb$  is the lowest sensor node position among sensor nodes operating on the network. We express  $ub$  and  $lb$ , as follows:

$$ub = \max\{p_{ns_1}, p_{ns_2}, p_{ns_3}, \dots, p_{ns_n}\}, \quad (14)$$

$$lb = \min\{p_{ns_1}, p_{ns_2}, p_{ns_3}, \dots, p_{ns_n}\}. \quad (15)$$

Here,  $p_{ns}$  denotes the sensor node's position. Furthermore, MDSA evaluates the eligibility of each virtual CH candidate to be the best virtual CH. The best feature of virtual CH is that it is close to the sensor nodes. This closeness has advantages. The first advantage is that communication costs between sensor nodes are lower because sensor nodes do not need to spend extra energy transmitting data to one another. The second advantage, the observation data between sensor nodes, has a high correlation. The eligibility evaluation of this virtual CH candidate is carried out based on the following fitness functions:

$$F_{VCH} = \min \sum_{i=1}^{NS_L} \sum_{j=1}^{n_{copt}} dist(p_{ns_i}, p_{vch_j}). \quad (16)$$

Here,  $p_{vch}$  represents VCH position.

The MDSA evaluation result is the best virtual CH position and cluster-ID. Once MDSA ends its search process, the sink emits a message to each sensor node. This message contains the virtual CH position and cluster-ID. After the sensor nodes receive the message, the sensor nodes update their configuration and form clusters.

### 4.3 Selection of cluster head and load distribution

After the sensor nodes are organized into  $n_{copt}$  clusters, each cluster performs a cluster head selection process. Because the CH has a more solemn task than the cluster members, the CH selection system must determine the best sensor node as the CH in the cluster. However, the election of the CH is an NP-hard problem. Thus, a metaheuristic approach is required to select the best CH. This study proposes a modified DS algorithm to select CH. To evaluate the best sensor node as CH, we adopted the

formulation in [33] with adjustments. This formulation considers the position of the sensor node and the energy level of the sensor node.

The first consideration in choosing a CH is how close the CH candidate is to the virtual CH. Virtual CH occupies a point that is closest to all sensor nodes in a cluster. So that the position of the CH candidate close to the virtual CH allows the cluster members to save transmission energy. Assume that the position of the sensor node  $i$  as the CH candidate in cluster  $c$  is  $p_{ns_{i-c}}$  and the virtual position of CH in cluster  $c$  is  $p_{vch_c}$ , this first consideration is expressed in:

$$F_{ns-vch} = DVCH(p_{ns_{i-c}}, p_{vch_c}). \quad (17)$$

A second consideration is the distance of the CH candidate from the sink. The closer the CH distance to the sink, the less energy required for data transmission. Assuming the sink position is  $p_{bs}$  This second consideration is defined as:

$$F_{ns-sink} = DBS(p_{ns_{i-c}}, p_{bs}). \quad (18)$$

The third consideration is the energy remaining if the sensor node  $i$  is selected as CH. These considerations are essential to ensure that CH does not suddenly run out of energy while CH is performing tasks. Suppose that the energy demand in cluster  $c$  is  $E_{cluter}$  and the energy available at the sensor node as a CH candidate is  $E_{av-ns}$ , then the remaining energy of the sensor node if selected as CH is:

$$F_{ns-E} = E_{av-ns} - E_{cluter}. \quad (19)$$

The three considerations need to be combined and minimized to determine the appropriate sensor node as CH. The following is a fitness function to evaluate the feasibility of a sensor node:

$$\text{Minimize } F_{CH} = F_{ns-vch} + F_{ns-sink} - F_{ns-E}. \quad (20)$$

Based on this objective function, if a sensor node has more energy to fulfill the cluster's energy demands and is closer to the sink and the virtual CH, it has a significant chance of being selected as CH.

After the MDSAC protocol has successfully selected a sensor node as CH, the MDSAC protocol also determines how many rounds of a sensor node can return following the CH selection. This rotation is necessary to avoid overloading specific sensor nodes. Thus, there is an even distribution of the load among all sensor nodes in one cluster. The study in

---

**Algorithm 1:** Virtual backbone formation, cluster formation, and CH selection

---

**Result:** Relay node, cluster, and CH

```

1:   $n \leftarrow$  assign total number of sensor nodes in
    network area
2:   $NS \leftarrow$  define a set of sensor nodes spread over in
    network area
3:
4:  While the number of alive sensor nodes  $> 0$  do
5:     $NS_L \leftarrow$  define a set of alive sensor nodes
6:     $Maxd_{tos} \leftarrow$  find the maximum  $d_{tos}$  in  $NS_L$  and
    assign
7:    If  $Maxd_{tos} > d_t$ 
8:      Build virtual backbone
9:      Relay nodes transmit announcement message
    to other sensor nodes
10:   End
11:   Calculate mean  $d_{tos}$ 
12:   Calculate  $n_{copt}$ 
13:   Cluster topology construction
14:   Sink assigns and transmits cluster-ID and
    VCHs position to each sensor node
15:   While  $n_{copt} > 0$  do
16:     CH selection
17:     Insert  $ns_h$  as selected CH to CH list
18:      $CHCount \leftarrow CHCount + 1$ 
19:     Calculate limits on the number of rounds to
    participate in the CH elections
20:     CH transmits advertisement message to
    CMs
21:     CH broadcasts TDMA Schedule to CMs
22:      $n_{copt} = n_{copt} - 1$ 
23:   End
24: End

```

---

[3] inspired a rotational mechanism applicable to the MDAC protocol. Assuming the number of alive sensor nodes in the cluster is  $H_c$ , the number of rounds for the sensor node to be reselected as CH is:

$$F_{ro} = H_c - 1 \quad (21)$$

In subsequent rounds, this value is reduced by one.

In the final process of selecting CH in the MDSAC protocol, each selected CH broadcasts an announcement message to each cluster member. Besides, CH distributes the TDMA schedule to each sensor node. This TDMA schedule regulates the time allocation when the sensor nodes transmit data to the CH.

#### 4.4 Modified differential search algorithm (MDSA) for WSN

The differential search algorithm (DSA) is a metaheuristic optimization algorithm developed by mimicking living organisms' migration patterns. DSA uses the Brownian-like random-walk



Table 2. Symbol list in proposed protocol.

Symbol	Meaning
$NR$	Set of relay nodes
$d_{toS}$	Distance from node sensor to sink.
$E_{av-ns}$	Residual energy of sensor node.
$\overline{E_{av-ns}}$	Average residual energy of sensor nodes.
$H$	Total number of alive sensor nodes.
$R_s$	Transmission range distance.
$n_c$	The number of clusters.
$\overline{L_c}$	Average cluster area.
$S \times S$	Size of network area.
$NS_L$	Set of alive sensor nodes.
$R_c$	Radius of cluster.
$\rho$	The density of a cluster.
$x_{ch}$	CH position at the x coordinate.
$y_{ch}$	CH position at the y coordinate.
$E[d_{toCH}^2]$	The expected quadratic transmission distance from the cluster member to CH.
$n_{copt}$	The optimum number of clusters.
$VCH$	Set of virtual CHs.
$ub$	Upper bound limit.
$lb$	Lower bound limit.
$F_{VCH}$	Object function to find the minimum distance from alive sensor nodes to virtual CHs in the network area.
$dist(p_{ns_i}, p_{vch_j})$	Distance between the position of sensor node $i$ to the position of virtual CH $j$ in the network area.
$DVCH(p_{ns_{i-c}}, p_{vch_c})$	Distance between the position of sensor node $i$ to the position of virtual CH in the $c$ cluster.
$DBS(p_{ns_{i-c}}, p_{bs})$	Distance between the position of sensor node $i$ in the $c$ cluster to sink.
$E_{cluter}$	Consumed energy in a cluster.
$F_{CH}$	Object function to find the optimum CH.
$F_{ro}$	The number of rounds for the sensor node to be reselected as CH
$H_c$	The number of alive sensor nodes in a cluster.
$G$	The maximum generation.

$O$	The number of elements in the superorganism
$D$	The dimension of the problem.
$randw()$	Weibull random number.
$rand()$	Uniform random number.

movement model as a migration pattern. DSA needs to select this strategy so that DSA can solve problems related to nonlinear, multimodal, and multivariable optimization [34].

Several kinds of organisms, such as birds [35], bees, fire ants, whales, and monarch butterflies, have a habit of migrating based on a specific cycle of time or season. Changes in climate or seasons throughout the year significantly affect the number of food sources available in nature. Living organisms need to migrate to new habitats that are rich in food sources to maintain survival. In this migration process, these organisms organize themselves into a large group called a superorganism. Then, superorganisms move from one habitat to another in search of a habitat that is rich in food sources. The movement of superorganisms to find new habitats can be modeled in a Brownian-like random-walk [36].

In DSA, random solutions to optimization problems are generated and organized in the population. The population is analogous to migrating artificial superorganisms. Artificial-superorganism is denoted as:

$$Superorganism_g, g = \{1, 2, 3, \dots, G\}. \quad (22)$$

Here  $G$  is the maximum generation. Meanwhile, random solutions that are members of artificial-superorganisms are called artificial-organisms. This artificial-organism can be expressed as:

$$Org_i, i = \{1, 2, 3, \dots, O\}, \quad (23)$$

The number of artificial-organisms in a superorganism depends on the size of the problem, viz:

$$org_{ij}, j = \{1, 2, 3, \dots, D\}. \quad (24)$$

Where  $O$  represents the number of elements in the superorganism, and  $D$  represents the dimension of the problem. At the upper limit of  $ub$  and lower limit of  $lb$ , each of these artificial organisms has an initial position which is defined as follows:

$$org_{ij} = rand() \times (ub_j - lb_j) + lb_j. \quad (24)$$

Individual artificial-organisms were randomly selected to move towards a suitable stopover site. The random selection of artificial-organism individuals uses the following equation:

$$donor = Org_{random\_shuffling(i)}. \quad (25)$$

Meanwhile, to acquire the position of the stopover, the DS algorithm uses the following equation:

$$Stopover = Superorganism + Scale \times (donor - Superorganism). \quad (26)$$

Here, *Scale* is the value generated by the Gamma random number generator. The following is the formula for generating *Scale*:

$$Scale = randg() \times [2 \cdot rand()_1] \times (rand()_2 - rand()_3). \quad (27)$$

Random number generation is used in heuristic and metaheuristic algorithms and other fields such as cryptography [37] and simulation and complex systems modeling. Generating random numbers is an essential parameter in metaheuristics. Even the performance of intelligence algorithms can be improved by selecting quality random number generators [38]. This quality random number has characteristics, namely an incredibly long sequence, uniform, and uncorrelated [39]. Research in [40] shows that the Weibull random number is of good quality. Also, this random number generator can improve the performance of the metaheuristic algorithm [40]. Besides, the modeling cases in terms of durability or reliability make use of Weibull. Thus, this study proposes a Weibull random number generator to replace the gamma random number generator in the DS algorithm.

$$Scale = randw() \times [2 \cdot rand()_1] \times (rand()_2 - rand()_3). \quad (28)$$

We list some of the symbols that are used in this section in Table 2 to ease understanding of MDSAC protocol.

## 5. Experiment and performance analysis

The performance measurement of the MDSAC protocol was carried out by simulation under the MATLAB environment. We compared the MDSAC protocol with HFAPSO [29] and LEACH-GA [41] protocol. We also compared MDSAC protocol with two widely recognized in research at WSN, i.e., LEACH [14], and PEGASIS [18]. This evaluation

Table 3. The details of simulation parameters.

Simulation parameters	Value
Number of nodes	150, 200
WSN Area	100 m × 100 m, 200 m × 200 m, 300 m × 300 m
$E_{ec}$	50 nJ/bit
$E_{fs}$	10 pJ/bit/m <sup>2</sup>
$E_{mpf}$	0.0013 pJ/bit/m <sup>4</sup>
$E_{da}$	5 nJ
$E_0$	2 J
Packet's size	4000 bits
$p$ (LEACH)	0.05

used the parameters of the total energy consumption per round, the average residual energy per round, and the network lifetime. The total energy consumption per round is the amount of energy consumed by all sensor nodes in the network for transmitting, receiving, and aggregating data in each round [15, 30]. The average residual energy per round is the average residual energy of all sensor nodes in the network for one round. Meanwhile, network lifetime is when sensor nodes in the network start operating until the last sensor node runs out of energy [30].

Network lifetime can be measured based on the parameters of the first node dies (FND), half of the nodes die (HND), and the last node dies (LND). These three parameters can be defined as follow:

- FND is the number of rounds when the sensor node turns off for the first time.
- HND is the number of rounds when half of the sensor nodes on the network die.
- LND is the number of rounds when one last sensor node runs out of energy in the network.

In other literature, FND also represents a period of stability. After the sensor nodes pass through this period of stability, WSN enter a period of instability [30]. In this period of instability, the sensor nodes die off relatively quickly. Thus, one dead sensor node has high implications for the load of the other sensor nodes.

We simulated the experiment in two scenarios. In the first scenario, 150 and 200 sensor nodes were randomly distributed to the observation area measuring 100 m × 100 m. Furthermore, in the second scenario, 150 sensor nodes were randomly distributed to the observation area measuring 100 m × 100 m, 200 m × 200 m, dan 300 m × 300 m. In both scenarios, the sink position was precisely in the center of the network area. The parameters concerning the sensor node energy were adopted from research [14][32]. The values of the simulation

parameters used to evaluate the performance of the MDSAC protocol are shown in detail in Table 3.

### 5.1 Scenario 1: performance comparison in different node density level

The experiment in this first scenario aims to measure the MDSAC protocol's performance on the density level of the sensor nodes. The performance of this MDSAC protocol is compared to other protocols under the same conditions.

The first performance to evaluate is network lifetime. This parameter is significant because it indicates the protocol's ability to make sensor nodes operate within a specific period. The longer operating time for sensor nodes will certainly provide benefits for WSN users. We consider three parameters to assess network lifetime, namely FND, HND, and LND.

Table 4 shows that the MDSAC protocol produces the highest FND value among the comparison protocols. The FND value achieved by the sensor node using the MDSAC protocol was 4888. In comparison, the LEACH protocol provides FND at round 2184, PEGASIS at round 481, LEACH-GA at round 3164, and HFAPSO at round 2487. In addition to the FND factor, the MDSAC protocol can produce the highest LND value, namely 9995. Whereas the LEACH protocol only produces LND at round 4227, PEGASIS at round 4840, LEACH-GA at round 4318, and HFAPSO at round 6834.

Then, the experiment was continued by deploying 200 sensor nodes in the same area. The results of this experiment show that the MDSAC protocol still outperforms other protocols. The MDSAC protocol experienced FND at round 5134. Meanwhile, FND for the other four protocols, namely LEACH at round 2245, PEGASIS at round 414, LEACH-GA at round 3102 and HFAPSO at round 1784. The same applies to the LND parameter; the MDSAC protocol has the highest LND value, which is in the 9999 round. However, the four comparison protocols had less

LND value than the MDSAC protocol. The LEACH, PEGASIS, LEACH-GA, and HFAPSO protocols had LND values at rounds 7343, 4906, 4631, and 7100, respectively.

Fig. 4 shows the change in the number of alive sensor nodes in each round for each protocol when the number of initial sensor nodes is 150. As shown in the figure, the MDSAC protocol is superior to the LEACH, PEGASIS, LEACH-GA, and HFAPSO protocols in terms of FND and LND. The high FND value also indicates that the use of the MDSAC protocol can extend the period of stability. When the sensor nodes enter a period of instability, the decrease in the number of alive sensor nodes occurs slowly. In contrast, the LEACH, PEGASIS, and LEACH-GA protocols experience a sharp decrease in the number of live sensor nodes when the three protocols were in a period of instability. The period of instability for the MDSAC protocol is 5107 rounds. The LEACH, PEGASIS, LEACH-GA, and HFAPSO protocols have lower periods of instability than MDSAC, namely 2043 rounds, 4359 rounds, 1154 rounds, and 4347 rounds, respectively. Besides, the MDSAC protocol can increase network lifetime compared to the LEACH, PEGASIS, LEACH-GA, and HFAPSO protocols, namely 5768 rounds, 5155 rounds, 5677 rounds, and 3161 rounds, respectively.

The experimental results with the initial sensor node number of 200 also show the superiority of the MDSAC protocol, as shown in Fig. 5. The MDSAC protocol has a high period of stability compared to LEACH, PEGASIS, LEACH-GA, and HFAPSO. Besides, the decrease in the number of sensor nodes due to energy exhaustion is taking place slowly. Thus, the MDSAC protocol produces a high period of instability, which is at 4865 rounds. The period of instability of the LEACH, PEGASIS, LEACH-GA, and HFAPSO protocols was only 5098 rounds, 4492 rounds, 1529 rounds, and 5316 rounds respectively. The long period of instability provides the advantage of increased network lifetime for the MDSAC

Table 4. The results of the evaluation of scenario 1 for network lifetime parameters

Protocols	Network lifetime parameters					
	150 nodes			200 nodes		
	FND	HND	LND	FND	HND	LND
MDSAC	4888	9088	9995	5134	9172	9999
LEACH	2184	2978	4227	2245	3346	7343
PEGASIS	481	4302	4840	414	4369	4906
LEACH-GA	3164	3861	4318	3102	3816	4631
HFAPSO	2487	5098	6834	1784	4993	7100

protocol. Compared to LEACH, PEGASIS, LEACH-GA, and HFAPSO, the MDSAC protocol can increase network lifetime by 2650 rounds, 5093 rounds, 5368 rounds, and 2899 rounds, respectively.

Two factors cause the performance of the MDSAC protocol, which can extend the period of stability, the period of instability, and network lifetime. The first factor is low energy consumption. The second factor is load distribution. CH rotation among cluster members provides a relatively fair load balance provided the cluster is stable. The CH rotation that applies in LEACH and LEACH-GA depends on the percentage value of CH, which is  $\frac{1}{p}$ . The LEACH protocol stipulates that to enter CH selection, and the sensor node has never been a CH in the last  $\frac{1}{p}$  rounds [3,19]. This rotation system can create a condition in which only a few sensor nodes qualify as CH. So that the burden of CH increases because CH has many cluster members. Whereas in the PEGASIS protocol, the chance for a sensor node to become CL again depends on the number of sensor nodes spread in the network area. In PEGASIS, CL load is light because CL receives data that has been aggregated by its member nodes. Thus, each sensor node in the chain also spends energy on data aggregation in addition to transmitting and receiving data except for the first sensor node that transmits data.

Apart from the network lifetime parameters, the MDSAC protocol also shows advantages in the parameters of total energy consumption per round and average residual energy per round. Fig. 6, and Fig. 7 indicate that the total energy consumption per round protocol MDSAC has increased slowly for the number of sensor nodes 150 and 200. Whereas in Fig. 8, and Fig. 9 also indicate that the MDSAC protocol has an enormous average residual energy per round

compared to the LEACH, PEGASIS, LEACH-GA, and HFAPSO protocols for the number of sensor nodes 150 and 200. These four graphs indicate that the MDSAC protocol can save energy consumption better than LEACH, PEGASIS, LEACH-GA, and HFAPSO.

Three factors influence the energy efficiency of the MDSAC protocol. The first factor is the stability of the number of clusters and the optimal number of CH. The number of clusters and CH change when the number of sensor nodes decreases due to running out of energy. In the LEACH and LEACH-GA protocols, the number of CH and clusters was unstable. The number of CHs and clusters may vary at each round. This instability is influenced by the threshold value and random value generated by each sensor node in the setup phase. The second factor is the selection of the optimal CH. CH selection on LEACH and LEACH-GA only considers the threshold value without calculating other parameters such as residual energy and sensor node distance. So that when the sensor node selected as CH has a very far distance from the sink and cluster members, it can cause an increase in energy consumption. The third factor is the modified DSA as the basis of the MDSAC protocol, which can produce the optimal virtual CH and CH in scenario 1.

### 5.2 Scenario 2: effect of increasing network size on protocol performance

The experiment in the second scenario aims to evaluate the performance of the MSDAC protocol if the network area is increasing; however, the number of sensor nodes is constant. In this condition, the distance between the sensor node to the sink and other sensor nodes is getting further away. Also, the network density level is low.

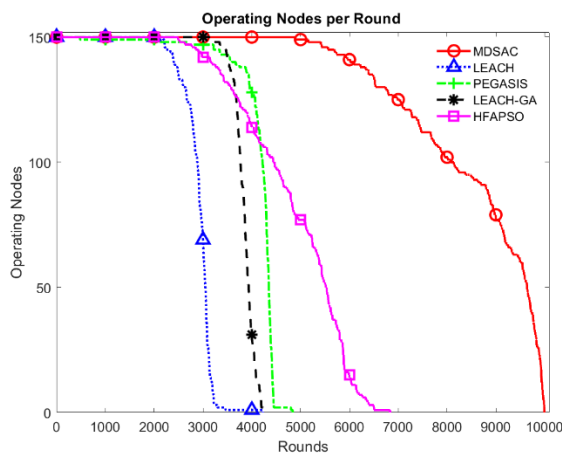


Figure. 4 Operating nodes per round for scenario 1 with 150 nodes

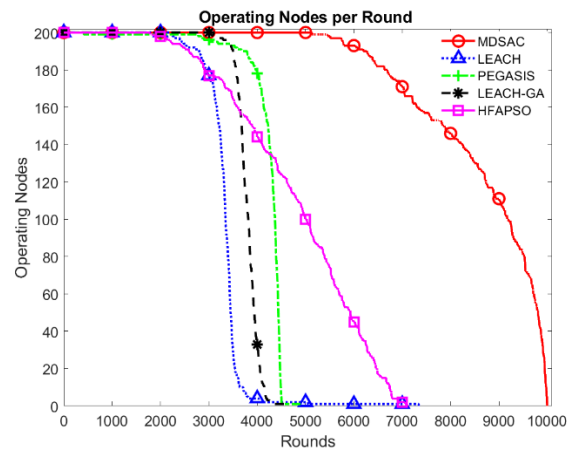


Figure. 5 Operating nodes per round for scenario 1 with 200 nodes

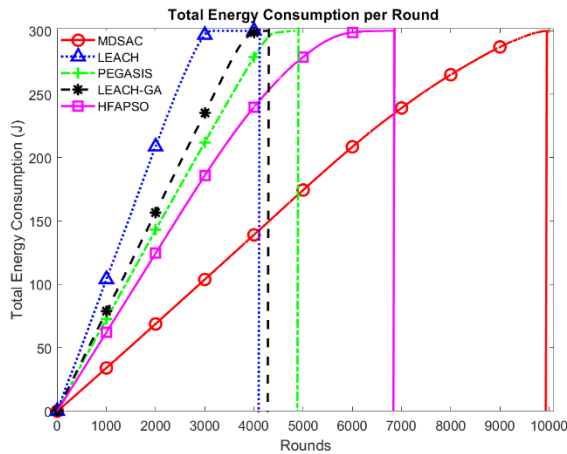


Figure. 6 Total energy consumption per round for scenario 1 with 150 nodes

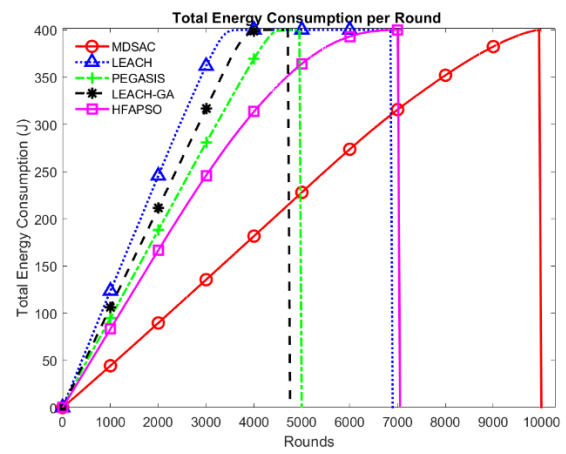


Figure. 7 Total energy consumption per round for scenario 1 with 200 nodes

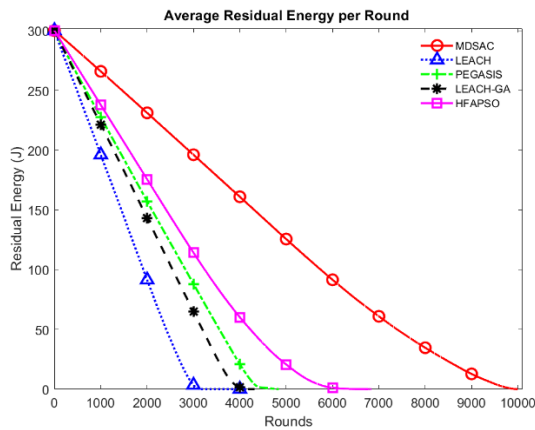


Figure. 8 Average residual energy per round for scenario 1 with 150 nodes

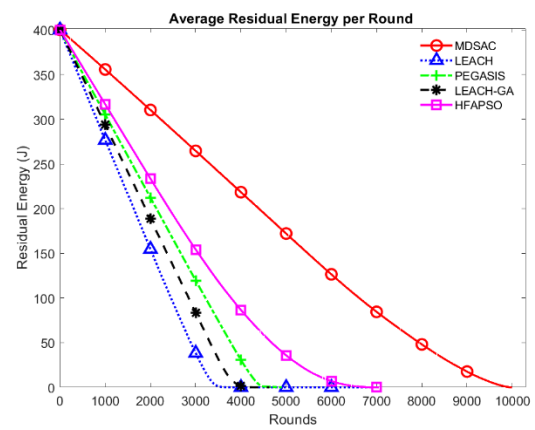


Figure. 9 Average residual energy per round for scenario 1 with 200 nodes

Table 5. The results of evaluation of scenario 2 for network lifetime parameters

Protocols	Network lifetime parameters					
	200m × 200m			300m × 300m		
	FND	HND	LND	FND	HND	LND
MDSAC	1554	<b>6577</b>	<b>19732</b>	266	<b>9172</b>	<b>12470</b>
LEACH	687	1285	4006	341	3346	5483
PEGASIS	1057	3410	4222	68	2257	3274
LEACH-GA	602	1634	3805	132	3816	3440
HFAPSO	<b>2521</b>	4314	4834	<b>812</b>	2374	4786

Table 5 shows the results of the second experiment scenario on the network lifetime parameters, which include FND, HND, and LND. In scenario 2, with a network area size of 200m × 200m, the MDSAC protocol is still superior in terms of HND, and LND. The MDSAC protocol produces FND at 1554 rounds. This value is indeed lower than HFAPSO. This significant decrease in FND value MSDAC protocol occurred because the cluster

structure changed frequently. This change affects a change in the value of the sensor node rotation, which acts as CH. Thus, the sensor nodes can be selected multiple times in a short period. However, the FND value for the MSDAC protocol is higher than LEACH, PEGASIS, and LEACH-GA. The LEACH, PEGASIS, and LEACH-GA protocols have FND values of 687, 1057, and 602, respectively.

In addition to the FND value, the MDSAC protocol can achieve the highest LND value, i.e., 19732, and has the most prolonged period of instability, namely 18178 rounds. LEACH, PEGASIS, LEACH-GA, and HFAPSO are only able

to reach LND at 4006 rounds, 4222 rounds, 3805 rounds, and 4834 rounds. Furthermore, four comparison protocols have relatively short periods of instability, namely 3319 rounds, 3165 rounds, 3203

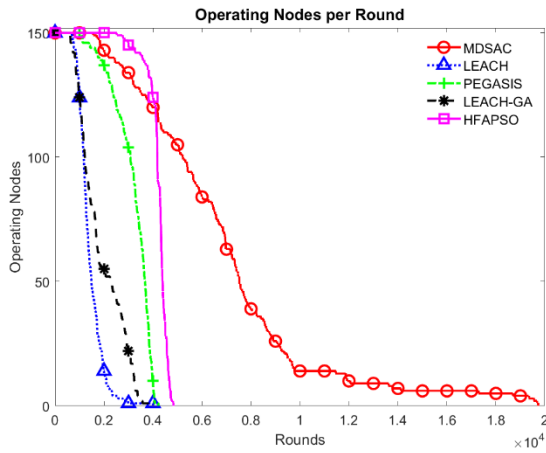


Figure. 10 Operating nodes per round for scenario 2 with 200m x 200m

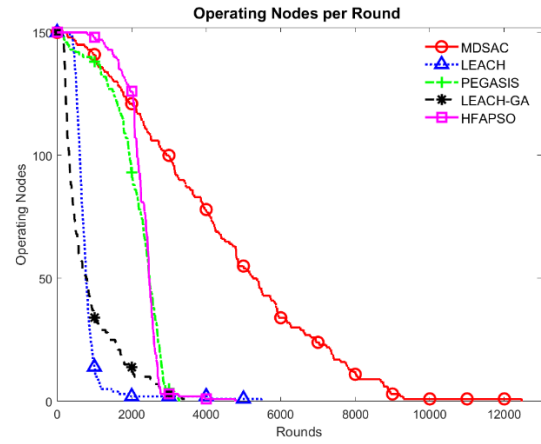


Figure. 11 Operating nodes per round for scenario 2 with 300m x 300m

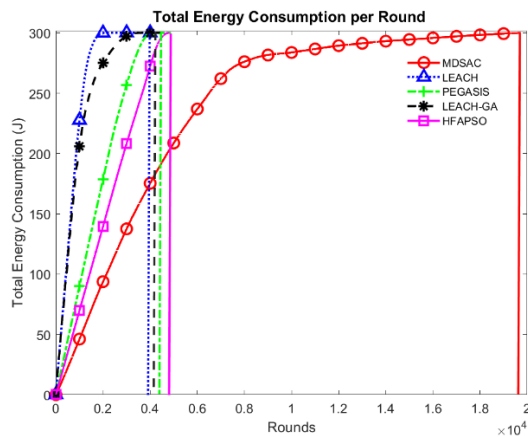


Figure. 12 Total energy consumption per round for scenario 2 with 200m x 200m

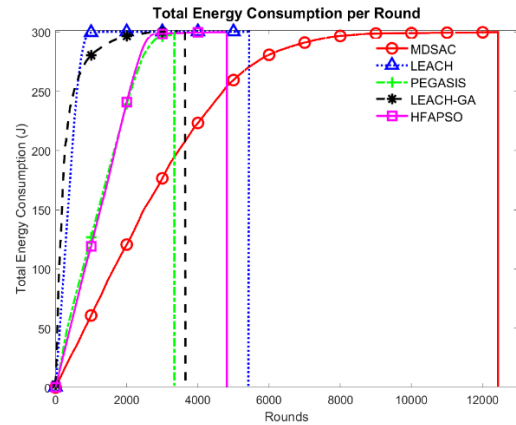


Figure. 13 Total energy consumption per round for scenario 2 with 300m x 300m

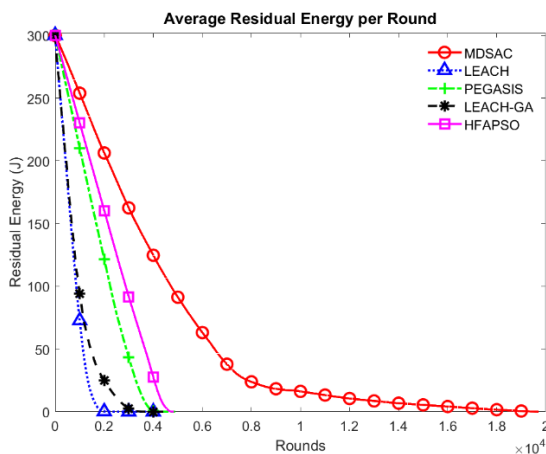


Figure. 14 Average residual energy per round for scenario 2 with 200m x 200m

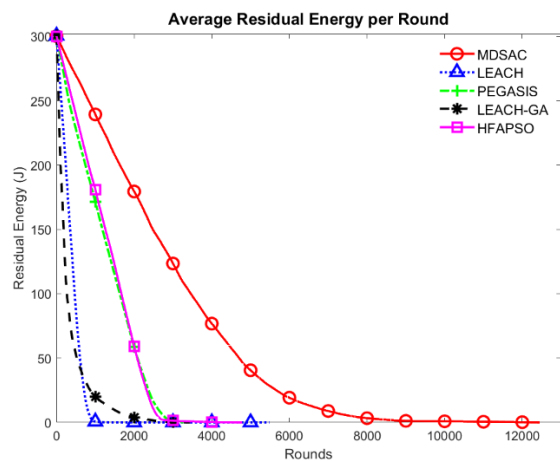


Figure. 15 Average residual energy per round for scenario 2 with 300m x 300m



rounds, and 2313 rounds, respectively. Overall, the MDSAC protocol was able to increase network lifetime compared to LEACH, PEGASIS, LEACH-GA, and HFAPSO, respectively, namely: 15726 rounds, 15510 rounds, 15927 rounds, and 14898 rounds. This increase in network lifetime is also clearly illustrated in Fig. 10. In the MDSAC protocol, the number of alive sensor nodes decreases slowly. This means that the sensor nodes can operate for a long time. In contrast, the other four protocols experienced a rapid decline in the number of nodes.

We continued the experiment by increasing the network area size to  $300\text{m} \times 300\text{m}$ . Based on the data presented in Table 5, an increase in network area size affects the decrease in the FND, HND, and LND values of all protocols. The FND values for the MDSAC, LEACH, PEGASIS, LEACH-GA, and HFAPSO protocols fell to 266, 341, 68, 132, and 812, respectively. Even though HFAPSO has a highest FND value, MDSAC protocol has the highest LND value and the longest network lifetime. The increase in the MDSAC network lifetime protocol to LEACH, PEGASIS, LEACH-GA, and HFAPSO is 6987 rounds, 9196 rounds, 9030 rounds, and 7684 rounds, respectively. Fig. 11 featuring the performance of the MDSAC protocol, which can maintain sensor nodes to operate over a long period. Meanwhile, the other four protocols are only able to operate in a relatively shorter period. These results indicate that the MDSAC protocol is still able to save energy use even though the network area size increases.

Fig. 12, and Fig. 14 confirm the total energy consumption per round and the average residual energy over a network area size of  $200\text{m} \times 200\text{m}$ . Based on these two figures, the LEACH protocol consumes the most energy. This immense level of energy consumption is due to the CH selection system. This system allows the CH position far from the sink to be selected. The same condition applies to LEACH-GA. The PEGASIS protocol also consumes much energy because the distance between the nodes is getting farther. Due to this massive energy consumption, these four protocols have a small average residual energy per round. In contrast to the three comparison protocols, the MDSAC protocol has a low energy consumption per round. Thus, the average residual energy per MDSAC round protocol is relatively high.

Fig. 13 and Fig. 15 display the total energy consumption per round and the average residual energy per round for the four protocols over a network area size of  $300\text{m} \times 300\text{m}$ . In this experiment, the LEACH-GA protocol consumed the most energy compared to LEACH, PEGASIS, HFAPSO and MDSAC. The LEACH-GA protocol's high energy

consumption level can be caused by the optimization results of the binary genetic algorithm, which are trapped at the local optimum. Thus, the LEACH-GA protocol may provide one CH. However, because CH also aggregates data from many sensor nodes, the network lifetime protocol LEACH-GA is higher than PEGASIS. Multi-hop data transmission, the further distance between nodes, and data aggregation at each node along the chain can accelerate node death. Therefore, multi-hop data transmission can be less efficient for large network areas and low node density. Compared to LEACH, PEGASIS, LEACH-GA, and HFAPSO, the MDSAC protocol has the lowest energy consumption rate per round and the high average residual energy per round. Although the size of the network area was increased, the MDSAC protocol was still able to operate efficiently. The MDSAC protocol can achieve energy efficiency because the presence of a relay node can save energy for CHs that have a long distance from the sink. Multilevel and adaptive relay node positions may provide better results. Relay nodes that move can also be an alternative strategy to increase efficiency in a large network area and low node density.

## 6. Conclusion

This paper describes the MDAC protocol, which is a clustering protocol based on a modified DSA algorithm to form clusters and determine the optimal CH to minimize energy consumption during WSN operations. In addition, the MDSAC protocol regulates CH rotation to ensure even load distribution between sensor nodes. Under certain conditions, the MDSAC protocol also defines relay nodes as virtual backbones for load sharing with CHs located far from the sink.

Under different experimental conditions, namely node density and network size, MDSAC protocol can save overall energy consumption and distribute the load well. Thus, the network lifetime can be significantly increased. Experimental results indicate this. On average, the MDSAC protocol increased FND by 1596 rounds, 2456 rounds, 1211 rounds, and 1060 rounds compared to LEACH, PEGASIS, LEACH-GA, and HFAPSO. In terms of HND, the MDSAC protocol increased the mean HND by 5764 rounds, 4918 rounds, 5221 rounds, and 4308 rounds compared to LEACH, PEGASIS, LEACH-GA, and HFAPSO. In addition, the MDSAC protocol increased the mean LND score, namely 7784 rounds, 8739 rounds, 9001 rounds, and 7161 rounds, respectively, compared to other comparison protocols. Therefore, the MDSAC protocol can operate for a long time in congested environments

and large network areas. Future research will lead to mobile relay or data ferry nodes and adaptive multilevel relay nodes to solve wireless sensor network problems.

### Conflicts of Interest

The authors declare no conflict of interest.

### Author Contributions

Conceptualization, KNP. Pamungkas, W. Wibisono, and S. Djanali; methodology, KNP. Pamungkas; software, KNP. Pamungkas; formal analysis, KNP. Pamungkas; writing—original draft preparation, KNP. Pamungkas; writing—review and editing, W. Wibisono and S. Djanali; supervision, W. Wibisono and S. Djanali.

### References

- [1] K. A. Darabkh, M. Z. El-Yabroudi, and A. H. El-Mousa, “BPA-CRP: A balanced power-aware clustering and routing protocol for wireless sensor networks”, *Ad Hoc Networks*, Vol. 82, pp. 155–171, 2019.
- [2] M. Arghavani, M. Esmaeili, M. Esmaeili, F. Mohseni, and A. Arghavani, “Optimal energy aware clustering in circular wireless sensor networks”, *Ad Hoc Networks*, Vol. 65, pp. 91–98, 2017.
- [3] W. Wibisono, T. Ahmad, R. M. Ijtihadie, K. Monika, and D. Pertiwi, “A node density-based approach for energy-efficient data gathering protocol in wireless sensor network environments”, *International Journal of Innovative Computing, Information and Control*, Vol. 16, No. 2, pp. 681–700, 2020.
- [4] A. Ghazi, B. Ahiod, and M. Abbad, “TLBO-Based Routing Approach for Wireless Sensor Networks”, *International Journal of Intelligent Engineering and Systems*, Vol. 11, No. 1, pp. 94–103, 2018.
- [5] S. Jothi and M. Chandrasekaran, “Cluster based compressed data aggregation and routing in WSN”, *International Journal of Intelligent Engineering and Systems*, Vol. 9, No. 4, pp. 69–78, 2016.
- [6] P. R. Babu and P. V. Krishna, “Ant colony and tolerable delay based modified virtual backoff algorithm for wireless sensor networks”, *International Journal of Intelligent Engineering and Systems*, Vol. 10, No. 2, pp. 78–87, 2017.
- [7] M. Radhika and P. Sivakumar, “Energy optimized micro genetic algorithm based LEACH protocol for WSN”, *Wireless Networks*, Vol. 8, 2020.
- [8] L. Chan, K. Gomez Chavez, H. Rudolph, and A. Hourani, “Hierarchical routing protocols for wireless sensor network: a compressive survey”, *Wireless Networks*, Vol. 26, No. 5, pp. 3291–3314, 2020.
- [9] A. Zeb, A. K. M. Muzahidul Islam, M. Zareei, I. Al Mamoon, N. Mansoor, S. Baharun, Y. Katayama, and S. Komaki, “Clustering analysis in wireless sensor networks: The ambit of performance metrics and schemes taxonomy”, *International Journal of Distributed Sensor Networks*, Vol. 12, No. 7, pp. 1–24, 2016.
- [10] X. Liu, “An optimal-distance-based transmission strategy for lifetime maximization of wireless sensor networks”, *IEEE Sensors Journal*, Vol. 15, No. 6, 2015.
- [11] M. M. Afsar and M.-H. Tayarani-N, “Clustering in sensor networks: A literature survey”, *Journal of Network and Computer Applications*, Vol. 46, pp. 198–226, 2014.
- [12] S. Tyagi and N. Kumar, “A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks”, *Journal of Network and Computer Applications*, Vol. 36, No. 2, pp. 623–645, 2013.
- [13] B. Jan, H. Farman, H. Javed, B. Montrucchio, M. Khan, and S. Ali, “Energy efficient hierarchical clustering approaches in wireless sensor networks: A survey”, *Wireless Communications and Mobile Computing*, Vol. 2017, 2017.
- [14] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks”, In: *Proc. of the 33rd Annual Hawaii International Conf. on System Sciences*, Hawaii, USA, pp. 1–10, 2000.
- [15] M. Elshrkawey, S. M. Elsherif, and M. Elsayed Wahed, “An Enhancement Approach for Reducing the Energy Consumption in Wireless Sensor Networks”, *Journal of King Saud University - Computer and Information Sciences*, Vol. 30, No. 2, pp. 259–267, 2018.
- [16] S. K. Singh, P. Kumar, and J. P. Singh, “A Survey on Successors of LEACH Protocol”, *IEEE Access*, Vol. 5, pp. 4298–4328, 2017.
- [17] M. M. Afsar and M. H. Tayarani-N, “Clustering in sensor networks: A literature survey”, *Journal of Network and Computer Applications*, Vol. 46, pp. 198–226, 2014.
- [18] S. Lindsey and C. S. Raghavendra, “PEGASIS: Power-efficient gathering in sensor information systems”, In: *Proc. of IEEE Aerospace Conf. Proc.*, Vol. 3, pp. 1125–1130, 2002.

- [19] X. Liu, "Routing Protocols Based on Ant Colony Optimization in Wireless Sensor Networks: A Survey", *IEEE Access*, Vol. 5, pp. 26303–26317, 2017.
- [20] P. C. S. Rao, P. K. Jana, and H. Banka, "A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks", *Wireless Networks*, Vol. 23, No. 7, pp. 2005–2020, 2017.
- [21] M. Krishnan, S. Yun, and Y. M. Jung, "Enhanced clustering and ACO-based multiple mobile sinks for efficiency improvement of wireless sensor networks", *Computer Networks*, Vol. 160, pp. 33–40, 2019.
- [22] A. Adamou Abba Ari, B. Omer Yenke, N. Labraoui, I. Damakoa, and A. Gueroui, "A power efficient cluster-based routing algorithm for wireless sensor networks: Honeybees swarm intelligence based approach", *Journal of Network and Computer Applications*, Vol. 69, pp. 77–97, 2016.
- [23] P. Civicioglu, "Transforming geocentric cartesian coordinates to geodetic coordinates by using differential search algorithm", *Computers & Geosciences*, Vol. 46, pp. 229–247, 2012.
- [24] P. Hemavathi and A. N. Nandakumar, "Novel scheme for minimal iterative PSO algorithm for extending network lifetime of wireless sensor network", *International Journal of Electrical and Computer Engineering*, Vol. 8, No. 2, pp. 1084–1091, 2018.
- [25] N. Srikanth and M. S. Ganga Prasad, "Efficient clustering protocol using fuzzy K-means and midpoint algorithm for lifetime improvement in WSNs", *International Journal of Intelligent Engineering and Systems*, Vol. 11, No. 4, pp. 61–71, 2018.
- [26] A. Yadav, S. Kumar, and S. Vijendra, "Network Life Time Analysis of WSNs Using Particle Swarm Optimization", *Procedia Computer Science*, Vol. 132, pp. 805–815, 2018.
- [27] P. T. Karthick and C. Palanisamy, "Optimized cluster head selection using krill herd algorithm for wireless sensor network", *Automatika*, Vol. 60, No. 3, pp. 340–348, 2019.
- [28] H. Ouldzira, H. Lagraini, A. Mouhsen, M. Chhiba, and A. Tabyaoui, "MG-leach: An enhanced leach protocol for wireless sensor network", *International Journal of Electrical and Computer Engineering*, Vol. 9, No. 4, pp. 3139–3145, 2019.
- [29] B. Pitchaimanickam and G. Murugaboopathi, "A hybrid firefly algorithm with particle swarm optimization for energy efficient optimal cluster head selection in wireless sensor networks", *Neural Computing and Applications*, Vol. 32, No. 12, pp. 7709–7723, 2020.
- [30] M. Abo-Zahhad, S. M. Ahmed, N. Sabor, and S. Sasaki, "Mobile Sink-Based Adaptive Immune Energy-Efficient Clustering Protocol for Improving the Lifetime and Stability Period of Wireless Sensor Networks", *IEEE Sensors Journal*, Vol. 15, No. 8, pp. 4576–4586, 2015.
- [31] M. Koç and I. Korpeoglu, "Coordinated movement of multiple mobile sinks in a wireless sensor network for improved lifetime", *EURASIP Journal on Wireless Communications and Networking*, Vol. 2015, No. 1, p. 245, 2015.
- [32] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks", *IEEE Transactions on Wireless Communications*, Vol. 1, No. 4, pp. 660–670, 2002.
- [33] K. N. P. Pamungkas, W. Wibisono, and S. Djanali, "An Optimum Clustered Grid-Based Particle Swarm Optimization to Enhance Efficiency Energy in Wireless Sensor Networks", In: *Proc. of 2019 3rd International Conf. on Informatics and Computational Sciences (ICICoS)*, Semarang, pp. 1–6, 2019.
- [34] P. Civicioglu and E. Besdok, "A conceptual comparison of the Cuckoo-search, particle swarm optimization, differential evolution and artificial bee colony algorithms", *Artificial Intelligence Review*, Vol. 39, No. 4, pp. 315–346, Apr. 2013.
- [35] S. A. Cabrera-Cruz, T. J. Mabee, and R. Villegas-Patracá, "Patterns of nocturnal bird migration in southern Mexico", *Revista Mexicana de Biodiversidad*, Vol. 88, No. 4, pp. 867–879, 2017.
- [36] V. Trianni, E. Tuci, K. M. Passino, and J. A. R. Marshall, "Swarm Cognition: An interdisciplinary approach to the study of self-organising biological collectives", *Swarm Intelligence*, Vol. 5, No. 1, pp. 3–18, 2011.
- [37] C. Kösemen, G. Dalkılıç, and Ö. Aydin, "Genetic programming-based pseudorandom number generator for wireless identification and sensing platform", *Turkish Journal of Electrical Engineering and Computer Sciences*, Vol. 26, No. 5, pp. 2500–2511, 2018.
- [38] A. Ozpinar and E. S. Kucukasci, "Use of Chaotic randomness numbers - metaheuristic and artificial intelligent algorithms", in *Intelligent Techniques for Data Analysis in Diverse Settings*, N. Celebi, Ed. Hershey, PA: IGI Global, pp. 207–227, 2016.

- [39] M. Cárdenas-Montes, M. A. Vega-Rodríguez, and A. Gómez-Iglesias, “Sensitiveness of Evolutionary Algorithms to the Random Number Generator”, In: *Proc. of Part I: Adaptive and Natural Computing Algorithms: 10th International Conf. (ICANNGA 2011)*, A. Dobnikar, U. Lotric, and B. Ster, Eds. Springer Berlin Heidelberg, pp. 371–380, 2011.
- [40] C. J. Montiel Moctezuma, J. Mora, and M. González Mendoza, “A self-adaptive mechanism using weibull probability distribution to improve metaheuristic algorithms to solve combinatorial optimization problems in dynamic environments”, *Mathematical Biosciences and Engineering*, Vol. 17, No. 2, pp. 975–997, 2020.
- [41] J. Liu and C. V. Ravishankar, “LEACH-GA : Genetic Algorithm-Based Energy-Efficient Adaptive Clustering Protocol for Wireless Sensor Networks”, *International Journal of Machine Learning and Computing*, Vol. 1, No. 1, pp. 79–85, 2011.