Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 47 (2026), No. 5-B, pp. 1945–1954
DOI : 10.47974/JIOS-2286

KBBOCVI : A hybrid biogeography-based optimization algorithm for optimal cluster head selection for HWSNs

Sumit Kumar *
Hukum Chand Saini †
Department of Computer Science and Engineering
Jagannath University
Jaipur 303901
Rajasthan
India

Madan Lal Saini §
Department of Computer Science & Engineering
Apex Institute of Technology
Chandigarh University
Mohali 140413
Punjab
India

Sandeep Kumar ‡
Department of AI and Data Science Engineering
CHRIST (Deemed to be University)
Bangalore 560074
Karnataka
India

Abstract

In wireless sensor networks Energy-efficient communication is one of the key requirements for a long lifetime. Clustered organization of the sensor nodes is widely acceptable as an energy-efficient routing technique. Finding optimal cluster heads can be considered an NP-hard problem where classical algorithms fail. Therefore, in this paper, a hybrid algorithm biogeography-based optimization (BBO) called KBBOCVI is proposed which combines K-means with BBO and uses cluster validity index (CVI) for measuring

* *E-mail:* **sumitakashmathur@gmail.com (Corresponding Author)**
† *E-mail:* **hukumchand.saini@jagannathuniversity.org**
§ *E-mail:* **mlsaini@gmail.com**
‡ *E-mail:* **sandpoonia@gmail.com**

the quality of the solutions. The suggested KBBOCVI performance is compared with other cutting-edge techniques such as stable election protocol (SEP), evolutionary routing protocol (ERP), intelligent hierarchical routing (IHCR), and KBBO in terms of residual energy, overall lifespan and network stability. The simulation results validated that the KBBOCVI outperforms others.

Subject Classification: *Primary 68M15, Secondary 68M12.*

Keywords: *Wireless sensor network, Biogeography-based optimization, Cluster validity indices, Clustering.*

1. Introduction

A variety of small electro mechanical terminals called sensor nodes are used to make Wireless Sensor Networks (WSN). It presents a great performance in variety of complex problem [1]. Temperature, pressure, humidity, and other parameters are detected, calculated, and transmitted from the target region such as agricultural field using these sensor nodes. These tiny nodes are capable of data collection, aggregation, and transmission due to their sensing, communication, and computation capabilities [2].

By clustering the sensor nodes and choosing the cluster heads (CH) to transmit the data to the sink, the energy usage can be further decreased. In literature, clustering is one of the well-accepted techniques to data transmission in WSNs [3]. One of the challenge in WSN is energy efficiency and fault recovery [4]. Many hybrid approach combined to investigate performance of WSN in energy efficiency [5]. In clustering, randomly distributed nodes over the given area form the groups based on various parameters, like distance, energy consumption, and many other parameters. After a predetermined number of iterations, the most powerful node in each group is chosen to act as the CH. Clustering is an effective technique widely used in designing energy efficient routing protocols.

If there is a WSN with n sensor nodes, then there will be 2n-1 possible combinations of cluster heads in which we need to select the best one. Various researchers have used Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Genetic Algorithm (GA) [13], Biogeography-based Optimization (BBO) and other meta-heuristics optimization algorithms to find optimal sets of cluster heads [5-6].

In this paper, the most suitable set of cluster heads is chosen using BBO. To enhance the performance of the BBO algorithm, it is hybridized with K-means clustering. Moreover, a new fitness function has been proposed based on the cluster validity index. The proposed work's primary research contribution can be split into two categories:

- A variant of BBO is proposed by giving the seed population using K-means
- A new fitness function is proposed based on the cluster validity index

The paper is presented as follows: The literature survey of metaheuristic algorithms for routing in WSN is presented in the next section. In section III, an overview of BBO is discussed. In sections IV the proposed protocol is described and simulation results mentioned in section V. Section V also concludes some suggestions for future work to increase performance of the algorithm.

2. Literature Review

In the literature a large number of heuristic and meta-heuristic algorithms are used for clustering in WSNs. This section presents some of the well-known algorithms in both categories.

1) **Heuristic-based algorithms:** LEACH [7] is the basic classical clustering protocol in WSN. Setup and steady-state are the two phases of LEACH. The set-up is responsible for clustering while the steady-state phase is used for data transmission. The cluster heads are selected based on probability and it has been rotated after a certain number of epochs. Due to the rotation of CHs, an overhead may be added. However, in this case energy consumption is balanced. Many other approaches are introduced in literature.
2) **Meta-heuristic based algorithms:** A survey of different nature-inspired clustering techniques for mobile WSNs was discussed in paper [8]. This paper elaborates different meta-heuristic methods used for cluster head selection considering various parameters for objective functions.

Paper [9] used genetic algorithms (GA) for selecting optimal clusters by defining a fitness function considering node distance and residual energy of a node. Further, in [10] author used the earthworm optimization algorithm (EOA) for better routing and selection of optimal CH. Moreover, in case of optimal routing, EOA selects the next hop based on the fitness function used for solution quality measurement that includes energy efficiency, distance, and delay. Further, paper [11] used Harmony Search Algorithm (HSA) for clustering and routing. The CHs are selected randomly based on harmony memory and pitch adjustment. The selected CH should have high energy and assigned members should have less distance between each other. While in routing, nodes with high residual energy, high inter-cluster distance, low intra-cluster distance and high node degree are taken into account when choosing the path. Paper [12] applied Chicken Swarm Optimization (CSO) to find optimal CHs which leads to decreased energy consumption. The CH selection phase, cluster formation phase and data transmission phase are the three stages in which the CSO based Clustering Algorithm (CSOCA) operates.

Out of various, evolutionary meta-heuristic methods, biogeography-based optimization (BBO) method has shown its efficacy over other similar methods in various application areas including wireless sensor networks [15]. BBO is used to select optimal CHs as well as find the optimal route for energy saving in WSNs [16], [17]. Moreover, in the literature, a lot of similar work has been done for selecting optimal CHs based meta-heuristic methods with various objective functions. Most of the considered objective functions have used residual energy of the node, inter and intra cluster distances and many other similar parameters.

However, there is still scope of improvement in finding a better objective function. Therefore, in this paper, we have designed a new objective function by considering the cluster validity index (CVI) for better cluster formation along with residual energy of the node for its claim to be a CH. The proposed fitness function is used with BBO for finding optimal CHs.

3. An overview of BBO algorithm

BBO [14] is based on the concept of Island biography where immigration and emigration rates are used to migrate the species from one island to another. The migration is dependent on the living conditions like temperature, rainfall and others). The concept of migration is utilized as the exploitation in the optimization process. Other than migration, BBO also has two more operators, namely elitism and mutation. Elitism is generally used to keep the best solutions preserved so that they can be transferred to the next generation. Moreover, mutation is used to implement the exploration process in the optimization. BBO is very much similar to a genetic algorithm (GA) with some basic differences in migration process. In BBO, the migration operator is similar to the crossover operator in GA.

4. Proposed method

In the proposed clustering method, BBO is used to find the optimal clusters in WSN. BBO is an evolutionary algorithm. One of the demerits of these algorithms is the random initialization of the population which may select the CHs in a region where sensor nodes are sparsely distributed. The CHs may be allocated farther to the sink and other CHs in the network. This may result in early decay of the energy for this node. Due to this reason, more energy consumption in the network will take place which may lead to poor routing performance. To avoid this problem, the initial seed of BBO is provided by the K-means clustering algorithm. K-means is one of the fast and efficient clustering algorithms. In WSN, K-means returns the CHs based on the mean value of the sensor node positions. Hence, hybridizing K-means with BBO leads to the best seed for the BBO even if the sensor node positions are biased.

Moreover, to generate the efficient clusters, a novel objective function is proposed which is inspired from CVIs and has been used to measure the quality of solution in the BBO algorithm. The proposed algorithm (KBBOCVI) is described as follows: Section 4.1 describes the formulation of objective function while section 4.2 describes KBBOCVI in detail.

4.1 *Formulation of objective function*

The main objective of the proposed fitness function can be divided into two parts: (i) optimal clustering in WSN (ii) to maximize lifetime of a network by minimizing the energy consumption. Optimal clustering means a set of clusters that are compact and well separated. Moreover, the network lifetime is the parameter which increases the functioning of the WSN. Clustering being a NP hard problem, BBO based clustering technique is applied in the proposed work to find out an optimal number of clusters and extend the network lifetime by minimizing total intra-cluster distance. The transmission energy (ETx) is directly proportional to distance d. If d is minimized, the total intra-cluster distance is also minimized. Then, ETx is also minimized resulting in increasing the network lifetime.

The two clustering parameters, intra- cluster distance (compactness) and inter-cluster distance (separation) are blended into a single fitness function which is defined by Eq. (1).

$$DI = \frac{1}{N}\sum_{i=1}^{N} max_{i\neq j}\{\frac{d(x_i)+d(x_j)}{d(c_i,c_j)}\} \quad (1)$$

Where N is the number of clusters, d (X_i) and d (X_j) are measures of compactness within clusters i, and j, d (C_i,C_j) measures the separation between CHs of cluster i and j. The smaller the value of DI, the better the separation of the clusters and the compactness within the clusters.

Moreover, the residual energy of the sensor nodes also plays a key role in selection of the promising CHs.

$$E = \sum_{i=1}^{N} RE(x_i) \quad (2)$$

Therefore, the overall objective function is formulated and given in Eq. (3)

$$Max\, f = {}^{E}/_{DI} \quad (3)$$

Where E is the overall residual energy of the network and we put it in the numerator as it needs to be maximized. As the overall objective of the fitness function is kept as maximized. Therefore, we keep the DI variable in the denominator as it needs to be minimized. The newly designed objective function in Eq. (3) will consider both the distance and energy factor of each node while selecting that node for the CH.

4.2 *KBBOCVI*

In the proposed KBBOCVI method, BBO is hybridized with K-means by modifying its population initialization phase described below. Moreover, for fitness calculation of each solution, a newly designed objective function (Eq. (3)) is used. There is no change in the migration and mutation operator of BBO. The population initialization phase is discussed below.

4.2.1 Population Initialization

In BBO, each solution in a population is represented by a binary string of zeros and ones. These solutions are also called Islands. A single Island structure in the context of clustering in WSN should represent cluster centroids and sensors. For example, let us consider the chromosome structure shown below:

0 0 0 1 1 1 0 1 0 1 0

Here, 1 represents a cluster centroid, 0 represents a sensor node. The size of a chromosome structure is the size of a sensor network. In KBBOCVI, population initialization is hybridized with K-means. The K-means clustering algorithm is provided in Algorithm 1.

Algorithm 1: The K-means clustering algorithm

1. Given a WSN (S) having N sensor nodes $\{x_1, x_2, \ldots, x_N\}$ which needs to be clustered into K non-overlapping clusters C = $\{C_1, C_2, \ldots, C_K\}$, such that $C_i \neq \emptyset$, $\forall$ i = 1, 2, 3, ….., K.

$$C_i \cap C_j = \phi \text{ if } i \neq j ,$$

$$C_1 \cup C_2 \cup \ldots \cup C_k = S$$

2. For each (N – K) sensor nodes assign a node x_i, i = 1, … N to cluster C_j, j $\in$ {1, 2, …, K} iff

$$\|x_i - C_j\| < \| x_i - C_q \|, \; q = 1,2\ldots K \text{ and } j \neq q$$

3. To compute the updated cluster centers, $C_1^*, C_2^*, \ldots, C_k^*$ the following equation is used.

$$C_i^* = \frac{1}{n_i} \sum_{x_j \in c_j} x_j \qquad i = 1, 2,.., K$$

where, n_i is the number of nodes belonging to a cluster C_i

4. Repeat Step 2 and 3 until C = i =1,2 ...K.

Since K-means converges faster, K-means algorithm runs first and its result is seeded with the initial population of BBO. The CHs returns by K-means are represented by 1s in a Island and the rest of the sensor nodes are represented by 0s denoting normal sensor nodes.

The proposed algorithm (KBBOCVI) is utilized in finding the optimal set of CHs in the set-up phase. We keep the steady-state phase similar to the standard algorithms like LEACH and SEP. The overall approach is discussed in Algorithm 2.

Algorithm 2: BBO based clustering algorithm in KBBOCVI

Input: A sensor network with N nodes distributed randomly in (XMax AND Ymax) sensing field. The sink is positioned in the middle of the sensor network.

Output: A set of optimal CHs

Step 1. Basic sensor network parameters are initialized based on the standard network model and energy model

Step 2. Initialize BBO Parameters such as number of generations, population size, elitism rate, and mutation rate.

Step 3. Set-up phase started

3.1 Initialize the population for BBO using Algorithm 1
3.2 Calculate fitness for each solution using **Eq**. (3)
3.3 **While** current iteration is less than maximum iteration **do**
3.4 Apply Elitism
3.5 Apply Migration
3.7 Apply Mutation
3.8 Calculate fitness value for updated population

EndWhile

Step 4. Save the best solution having optimal CHs after Step 3.

Step 5: Steady-state phase

5.1 The GPS location of CHs are advertised over the network

5.2 Member nodes joined the CHs based on the distances to from clusters

5.3 The data transmission will be done based on the assigned TDMA (time-division multiple access) slots.

5. Simulation Results

We have used MATLAB 2020a to simulate the WSNs on a core i5 processor with 8 GB of RAM. The region of the sensor network is considered as 100×100 where 100 nodes are deployed randomly. The performance of the proposed algorithm is compared and analyzed with other state-of-the-art methods, namely SEP, ERP, IHCR, and KBBO in terms of various performance metrics.

The considered network has following assumptions:

- The nodes are deployed randomly
- Some of the sensor nodes are having more energy than others initially. So the network has energy heterogeneity.
- The energy consumption depends upon two things, one is distance between the nodes and packet size.
- The distance is calculated based on the Euclidean distance
- The sink is placed at the middle of the network
- The sensor nodes do not have any mobility and the communication is omnidirectional and wireless.

Moreover, the energy model for the sensor network is kept as the literature [9]. The results for each considered method is compared in terms of overall lifetime of the network, stability, remaining energy, and number of packets sent to sink. All of these matrices are measured in terms of number of rounds. The definitions of all the considered performance parameters can be found in the literature [18].

The simulation results have been evaluated in two scenarios, one is when the network contains 10% more advanced nodes (m=0.1) and other is when networks contain 20% more advanced nodes (m=0.2). The proposed algorithm is executed round-by-round.

We analyzed the considered algorithms in terms of their stability period in terms of rounds. The stability period is the duration in which the first sensor died after functioning of the network. The duration is measured in terms of rounds where with m=0.1, SEP, IHCR, ERP, KBBO, KBBOCVI has stability period of 864, 914, 1012, 888 and 4144 respectively and when m=0.2, SEP, IHCR, ERP, KBBO, KBBOCVI has stability period of 999, 920, 1079, 1034 and

3783 respectively. It can be observed that the KBBOCVI has a larger stability period in both the scenarios.

The overall network lifetime is considered as the duration between the first node died and last node died in the network. The complete network is dead at 14265th round in case of KBBOCVI, 3952th round in KBBO, 3317th round in ERP, 3220th round in IHCR and 1563th round in SEP.

The next important performance criterion is residual energy.

We evaluated the proposed strategy with other considered methods in different intervals of the network in terms of the residual energy. Experimental setup is done for WSN with 10% heterogeneity and 20% heterogeneity respectively. The comparison is depicted for each 600 rounds interval till maximum of 6000 rounds. The results for KBBOCVI is far better than other considered methods as it preserves more energy in the networks. With 10% heterogeneity for 600 round interval it preserve 51.055 and 16.918 energy for 6000 round. Most of the considered methods are exhausted in the very early stage (approx. 3600th round last with KBBO with energy 0.118) except KBBOCVI. With 20% heterogeneity for 600 round interval it preserve 56.17 and 22.78 energy for 6000 round. Most of the considered methods are exhausted in the very early stage (approx. 3600th round last with KBBO with energy 0.46) except KBBOCVI.

6. Discussion and Conclusion

This study split up the entire network into clusters by using hybrid BBO and K-means techniques. The CHs for each cluster are chosen at the end of each iteration using a new fitness function derived from cluster validity indices. The suggested algorithm (KBBOCVI) outperforms other cutting-edge techniques in terms of stability period, energy dissipation, and network life. The right CHs are chosen in a load-balanced manner to maximize network longevity. Two scenarios with varying degrees of network heterogeneity are used to model the outcomes. This study can be expanded to mobile sensor networks, and the multi-hop approach can save even more energy.

References

[1] I. H. Abdalla and R. F. Yaser, "WSN recruitments for encrypted medical image transmission securely," *J. Discrete Math. Sci. Cryptogr.*, vol. 26, no. 7, pp. 1981–1990 (2023), doi: 10.47974/JDMSC-1838.

[2] S. P. Singh and S. C. Sharma, "Genetic algorithm based energy efficient clustering (GAEEC) for homogeneous wireless sensor networks," *IETE J. Res.* (2017).

[3] S. Famila and A. Jawahar, "Improved artificial bee colony optimization–based clustering techniques for WSNs," *Wireless Pers. Commun.* (2019).

[4] D. G. Takale, "Hybrid algorithm for fault node recovery and energy efficiency in wireless sensor networks," *J. Inf. Optim. Sci.*, vol. 45, no. 8, pp. 2347–2367 (2024), doi: 10.47974/JIOS-1831.

[5] D. Singh, "Energy-efficient wireless sensor networks for disaster management using hybrid technique," *J. Inf. Optim. Sci.,* vol. 45, no. 8, pp. 2249–2259 (2024), doi: 10.47974/JIOS-1799.

[6] H. Sadeghian and M. S. Aghae, "Improved cuckoo search-based clustering protocol for wireless sensor networks," *Majlesi J. Telecommun. Devices,* vol. 8, no. 2 (Jun. 2019).

[7] R. Sudarmani and R. Vanithamani, "Minimum spanning tree for clustered heterogeneous sensor networks with mobile sink," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res*. (2015).

[8] M. Sudha and J. Sundararajan, "Biologically inspired clustering algorithms in mobile wireless sensor networks: A survey," *Adv. Natural Appl. Sci.* (2017).

[9] T. Ganesan and P. Rajarajeswari, "Genetic algorithm based optimization to improve the cluster lifetime by optimal sensor placement in WSNs," *Int. J. Innov. Technol. Explor. Eng.* (Jun. 2019).

[10] V. R. Pasupuleti, "Efficient cluster head selection and optimized routing in wireless sensor networks using bio-inspired earthworm optimization algorithm," *J. Adv. Res. Dyn. Control Syst.,* vol. 11, Special Issue (2019).

[11] P. Lalwani, S. Das, H. Banka, and C. Kumar, "CRHS: Clustering and routing in wireless sensor networks using harmony search algorithm," *Neural Comput. Appl.,* vol. 30, pp. 639–659 (2018), doi: 10.1007/s00521-016-2662-4.

[12] W. Osamy, A. A. El-Sawy, and A. Salim, "CSOCA: Chicken swarm optimization based clustering algorithm for wireless sensor networks," *IEEE Access,* vol. 8, pp. 60676–60688 (2020), doi: 10.1109/ACCESS.2020.2983483.

[13] T. Sharma, G. S. Tomar, R. Gandhi, S. Taneja, and K. Agrawal, "Optimized genetic algorithm (OGA) for homogeneous WSNs," *Int. J. Future Gener. Commun. Netw.,* vol. 8, no. 4, pp. 131–140 (2015), doi: 10.14257/ijfgcn.2015.8.4.13.

[14] D. Simon, "Biogeography-based optimization," *IEEE Trans. Evol. Comput.,* vol. 12, no. 6, pp. 702–713 (2008).

[15] H. Ma, D. Simon, P. Siarry, Z. Yang, and M. Fei, "Biogeography-based optimization: A 10-year review," *IEEE Trans. Emerg. Topics Comput. Intell.,* vol. 9, no. 5 (2017).

[16] R. Pal, H. M. Avinash Pandey, and M. Saraswat, "BEECP: Biogeography optimization-based energy efficient clustering protocol for HWSNs," in *Proc. 9th Int. Conf. Contemp. Comput. (IC3)* (2016).

[17] P. Lalwani, H. Banka, and C. Kumar, "BERA: A biogeography-based energy saving routing architecture for wireless sensor networks," *Soft Comput.* (2016).

[18] B. A. Attea and E. A. Khalil, "A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks," *Appl. Soft Comput.,* vol. 12, no. 7, pp. 1950–1957 (2012).

Received April, 2025