

## Research Article

# Energy-Efficient Clustering and Localization Technique Using Genetic Algorithm in Wireless Sensor Networks

Junfeng Chen , Samson Hansen Sackey , Joseph Henry Anajemba , Xuewu Zhang ,  
and Yurun He 

*College of Internet of Things Engineering, Hohai University, Changzhou 213022, China*

Correspondence should be addressed to Junfeng Chen; [chen-1997@163.com](mailto:chen-1997@163.com)

Received 18 February 2021; Accepted 20 July 2021; Published 3 August 2021

Academic Editor: Abd E.I.-Baset Hassanien

Copyright © 2021 Junfeng Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Localization is recognized among the topmost vital features in numerous wireless sensor network (WSN) applications. This paper puts forward energy-efficient clustering and localization centered on genetic algorithm (ECGAL), in which the residual energy, distance estimation, and coverage connection are developed to form the fitness function. This function is certainly fast to run. The proposed ECGAL exhausts a lesser amount of energy and extends wireless network existence. Finally, the simulations are carried out to assess the performance of the proposed algorithm. Experimental results show that the proposed algorithm approximates the unknown node location and provides minimum localization error.

## 1. Introduction

The application of wireless sensor networks for precise localization of devices and humans in bounded scenarios characterizes a significant feature in the provision of highly classified essential services in the discipline of logistics and administration [1]. The basic implementation force for various localization means is monitoring, sensing, and tracking. However, few instances include health monitoring, habitat monitoring, agriculture monitoring, environment monitoring, military investigation, intruder recognition, pollution control, and space handling on planet and moon. WSN is perceived to be in any environmental structure such as underwater and terrestrial nature. Presently, the persistent distribution of such systems is prohibited as a result of their gelatinous intricacy and their costs of maintenance [2]. Primarily, this is owing to the fact that, in interior environments, precise localization in the occurrence of unembellished multipath fading frequently necessitates a huge amount of static sensing nodes which are referred to as anchors (also called beacon) with a recognized location [3] and the utilization of complex signal processing algorithms for disseminated proximity approximation [4]. In a broader spectrum, accumulating more anchor nodes means higher

position accuracy. However, anchor nodes cost more than ordinary nodes and are 10 times more expensive [5]. By and large, if the position of an unidentified node is discovered, the anchor nodes will then be negligibly ignored. Consequently, with the decrease of quantity of anchor nodes, the positioning of the node is affected [6]. Reducing anchor nodes cuts cost, and the more the anchor nodes are slashed, the lesser the accuracy is. Lately, the approach used in WSN localization is based on the optimization problem of multidimensional and multimodal which is solvable by utilizing stochastic methods which are based on population [7]. Section 2 of our research reviews several metaheuristic algorithms used in the field of WSN localization. The results of those research studies show that these algorithms have successfully and drastically minimized errors in localization [8].

The two distinctive ways by which localization can be performed are distributed and centralized localization. In the former, each node finds the unknown nodes by themselves, while in the latter, information from each node is transmitted to a centralized component for further processing in order to obtain information about the position. Furthermore, the clustering process of the nodes in the second one is seen to be divided into the network that is interconnected, called clusters, with each cluster having

many energized sensor nodes led by a cluster center (cluster head) node which is a coordinator. The cluster center works in sequence with other cluster centers which are all provisional base stations. Most WSN nodes are categorized into several states including isolated, normal, cluster center, and gateway. Ultimately, isolated nodes preserve a secure neighbor table where neighbor node information is stored. The selection of cluster centers is considered very essential in clustering. Clustering in sensor networks is to a large extent power efficient. In a broader term, localization methods are categorized into range-free and range-based methods [9]. Range-based localization exploits the distance connecting an unidentified node and a recognized node (anchor) even though range-free methods solely utilize nodes' connectivity data. On the other hand, range-based methods are fine-grained localization methods, whereas range-free methods are coarse-grained [10]. Range-based methods can be classified into four major classes, namely, time-difference-of-arrival (TDOA), time-of-arrival (TOA), angle-of-arrival (AOA), and received signal strength (RSS). On the other hand, range-free localization methods include centroid algorithm, Monte Carlo localization, approximate point in triangle test (APTT), closest point-based method, assumption-based coordinate method, DV-HOP, and amorphous method.

Energized sensor nodes in hostile environments must stay alive for long time, but it is ambiguous or almost impossible to boost or change their batteries [11], and this necessitates inventing new energy-efficient alternatives to some of the existing traditional wireless networking challenges which include intermediate access control, self-organization, bandwidth distribution, security, and routing. Some factors to consider in prolonging the lifetime of networks is by utilizing the gains of trade-offs such as power, latency, and accuracy, coupled with using hierarchical (tiered) architectures. The position of a localized unknown node is important because it helps solve bigger WSN problems like routing and data aggregation. For this reason, it is necessary to focus on localization methods in WSN works [12]. Localization occurs once when considering static nodes, whereas tracking is nonstop localization of the cell node over time. Clustering is a preferred method for attaining competent and accessible overall performance in WSNs [13]. Clustering nodes into sets conserves power and minimizes disputes at the network since the nodes transmit their information to their cluster heads through reduced distances. Thus, the circulation of command through the transmission facilitates permits locality of transmission [14]. With respect to the difficulties faced by energized sensor nodes in localization, this paper's aim is to establish an elevated energy-efficient localization technique which is dependent on low energy depletion and a clustering structure. The proposed approach continues to show strength by dealing with location accuracy via clustering means in GA localization. Additionally, ECGAL successively reduces the whole WSN energy consumption.

In summary, the main contributions of this paper are summarized as follows:

- (1). We utilized genetic algorithm (GA) with an energy-efficient clustering approach to solving localization problems in WSN.
- (2) The performance of the proposed ECGAL (energy-efficient clustering in genetic algorithm localization) is analyzed and compared with DV-HOP (distance vector-hop), CENTA (centroid algorithm), EDV-HOP (evolutionary distance vector-hop), and CGAL (clustering in genetic algorithm localization).
- (3) The results have shown that our proposed approach outperforms the existing localization algorithms with respect to energy efficiency, localized distance error, and coverage connection.

The remaining part of this paper is structured as follows. Section 2 elaborates several previous research exertions relating to localization. In Section 3, a brief description of the utilized genetic algorithm (GA) is presented, while Section 4 analyzes a sensor node localization setup and cluster formation in a wireless sensor network localization scenario. Section 5 provides details on the fitness function proposal, including the definitions of energy efficiency, distance estimation, and coverage connection. Section 6 presents the performance evaluation of ECGAL. Finally, the research summary and derived conclusions are presented in Section 7.

## 2. Related Works

Lately, there are several algorithms in the field of optimization used in reworking the drawbacks in the localization of WSN nodes [15]. Some of these existing related research studies are briefly discussed in this section.

The authors in [16] proposed an efficient hybrid bio-inspired optimization in localization methods which is applicable in industrial WSNs. In their research, they proposed particle swarm optimization (PSO) and dragonfly algorithm (DA) which earns slight time of computation and extensive accuracy. On the other hand, Kanoosh et al. suggested a node localization design which is dependent on a current bio-inspired algorithm known as salp swarm algorithm (SSA). The performance of their results is measured against similar optimization algorithms, specifically, particle swarm optimization (PSO), firefly algorithm (FA), grey wolf optimizer (GWO), and butterfly optimization algorithm (BOA) under distinctive wireless sensor network positioning [17]. Based on the current localization and monitoring setups, the localization of sensor nodes and the application of WSN target trailing technology had been examined from the perspectives of accuracy perfection, extending the WSN natural life with respect to coordination theory, particle filter, range-free theory, and different computing approaches. Zhang et al. merged an energized sensor triggering algorithm and dynamic clustering process to prolong a parallel particle filter algorithm and further target monitoring system [18]. The authors adopted a two-object tracking strategy used in WSNs primarily based on cluster algorithms which have been combined together to perform many features in the proposed algorithm. Musaffer et al.

benefited from using cluster algorithms to count and detect node in a cluster by reporting an event to the cluster center (also cluster head) node according to a query, conveying all audible information to the base station [19].

Numerous device-free localization methods are launched in wireless systems. Alippi et al. proposed a radio tomographic imaging- (RTI-) based localization technique imaging the received signal strength (RSS) reduction resulting from the targets with inexpensive and standard hardware [20, 21]. Xu et al. employed device-free wireless localization in WSNs which utilizes the RSS differences between sensor nodes in order to find solutions to problem [22, 23]. Wang et al. explained that when a target is found in a WSN deployment area, the existence of the target will imitate, strew, and engage the WSNs radio signals [24]. The outcomes of localization are determined with several RSS values using (3). Zhang and Wong exercised the genetic algorithm for gathering ecological questions within a WSN for successfully localizing its sensor nodes [25]. However, all the coordinates of the grid network offer random perturbations of the quality of the received signal. Furthermore, genetic algorithm can enormously acquire information about the location and minimize the likely errors related with the RSSI estimation assumed for every coordinate. Sackey et al. showed the implementation of GA to practically localize WSN nodes by means of three coordinates or more anchors [26].

The research of Wang et al. [27] centered on range-free localization as a cheap choice in comparison to range-based methods. But the localization based on range-free undergoes greater localization mistakes in contrast to the range-based algorithms. Furthermore, Sivakumar et al. offered an expanded form of DV-HOP which is a famous range-free method that is reliant on hop-proximity calculation [28]. Primarily, the enhancement in the DV-HOP algorithm is done depending on GA. Song et al. provided simulation results to prove the superior precision performance of the proposed localization algorithm in localization against the performance of other localization algorithms in [29]. Kumar discussed the localization of sensors in motion to deliver the advancement of connectivity, security, and energy consumption [30]. This needs to be an uncomplicated, dynamic, and adaptable localization method. Wang et al. presented a mobile sensor localization algorithm which is independent and has the capacity of coping with the ambiguities, connection breakdowns, and node flexibility in the network [31]. Nonetheless, computational intelligence (CI) possess qualities identical to such algorithms. Sharma et al. showed that CI methods can function in a setting of imprecision and ambiguity [32, 33]. CI consists of methods that can be altered. These methods can act logically in composite situations.

### 3. Brief Introduction of GA Algorithm

The proposed methodology relies on efficient clustering and strong global search for GA in order to increase accuracy and efficiency. Subsequently, the ensuing subsection presents the GA.

**3.1. Genetic Algorithm.** Genetic algorithm (GA) is derived from biological behaviors used in the field of optimization. GA is an existing metaheuristic driven by the approach of evolutionary algorithms for natural selection. The assumed population size is said to consist of  $N^p$  competitors (candidates) on the possible solution which is made up of decoding and encoding chromosomes to a fixed dimension of binary numbers. The given interval contains 0s and 1s with  $N$  bits. The GA approach follows a specific procedure which is apportioned along with the genetic algorithm operators into initialization, selection, crossover, and mutation. Figure 1 shows a sample structure of a chromosome.

**3.1.1. Initialization.** Firstly, the countless mixed-up candidate solutions created tend to shape the preliminary population. The general population range is subjected to natural adversities but characteristically comprises a number of hundreds or more viable solutions. Conventionally, the now scattered populace produces an overlay which completes the varying feasible options (the search space). The search space entails all likely options to the question. Seldomly, the options might be “seeded” in the search space with high-quality solutions.

**3.1.2. Selection.** The technique of decision making relies on the chromosomes’ fitness capacities to control the mating process for every individual. The chromosome holds the results in the shape of genes and is chosen in accordance to a particular selection method. The better the fitness value of these chromosomes, the greater the probability of being selected. Solutions with higher fitness have more chance to duplicate. However, ranking takes place after the chromosome with the highest fitness value attains the most appropriate chromosome. The defined function portrays the nearness a solution can change to, providing best results. Based on the value of likelihood selection, one or more individuals multiply to bring forth offspring. The probability of selecting is  $P^a$  of which each individual is determined by

$$P^a = \frac{\text{Fitness}^a}{\sum_{b=1}^{N^p} \text{Fitness}^b}, \quad (1)$$

where  $a \in \{1, \dots, N^p\}$ , and the fitness of the selected individual  $a^{\text{th}}$  is denoted as  $\text{Fitness}^a$ . The selection of a chromosome denoted as  $a$  is based on  $r \in (0, 1)$  random numbers. The cumulative probability  $C^a$  is well defined in (2), and it satisfies a chromosome selected at random within  $C^{i-1} < r \leq C^a$ .

**3.1.3. Crossover.** Crossover operator pairs two formerly selected chromosomes to copulate and produce offspring that share positive characteristics of both parents. Copulating comprises choosing two arbitrary crossover points  $c^1$  and  $c^2$  along the stretch of the chromosome. As a result, the encoded binary numbers are surrounded by some points that can be swapped between carefully chosen chromosomes interchangeably.

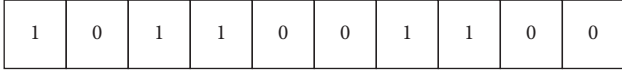


FIGURE 1: Structure of a sample chromosome.

$$C^a = \sum_{b=1}^{N^p} \text{Fitness}^b. \quad (2)$$

(1) *Single-Point Crossover*. In single-point crossovers, two parents can produce a cut point and re-merge the first fragment of the original parent crosses with the second fragment of the subsequent parent to create one offspring. The second fragment of the original parent is then fused with the first fragment of the subsequent parent to create another offspring. In Figure 2, a random point is preferred from two parents. Parents are then divided at the crossover point. Finally, children are created by exchanging tails.

(2) *Two-Point Crossover*. The single-point crossover and the two-point crossover share many similarities exempting the number of cut points they both create. In the two-point crossover, two offspring are created, but in the single-point crossover, only one cut point is made. As observed from the representation in Figure 3, there are two arbitrary numbers differing between 0 and 1 (length of the chromosome). The genes found between these two arbitrary chosen numbers are delivered from the initial parent to offspring and the complementary genes in the second parent are neglected. The vacant cells of the offspring are then singly filled with the unused genes from the second parent.

(3) *N-Point Crossover*. In Figure 4,  $n$  random crossover points are chosen from a chromosome sample, which is then fragmented along with those points. Then, exclusive parts are joined, which are alternating between parents. Finally, there is generality of 1 point (still some point preference).

3.1.4. *Mutation*. Immature concurrences in the algorithm are avoided since the mutation follows a GA mechanism which brings out uncharted results in the GA population. The arbitrary binary changes in chromosomes direct the process of mutation. The sequence from selection, mutation, and crossover is looped. The global optimum of a perfect but average individual is closely realized after multiple consecutive iterations of an expanding population.

(1) *Displacement Mutation*. The displacement mutation process arbitrarily selects two genes and transposes them after the parents are chosen. Figure 5 illustrates the implementation of the displacement mutation.

(2) *Shift Mutation*. Following the selection of the parent chromosome, two different points are carefully chosen randomly in a 1 to  $n$  interval (chromosome span) and the genetic factors positioned amidst these two points are moved towards the left corner, rotationally. An illustration of this shift mutation is presented in Figure 6.

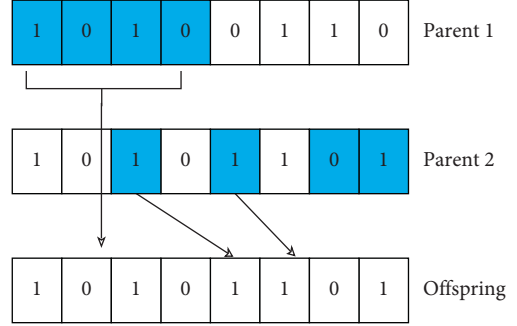


FIGURE 2: Single-point crossover.

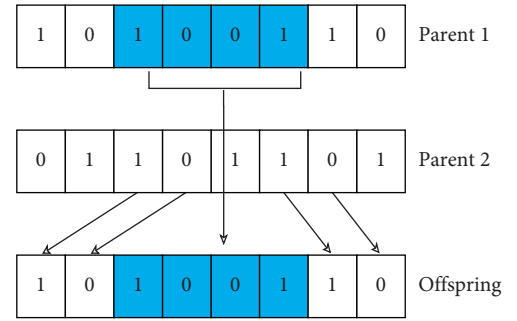


FIGURE 3: Two-point crossover.

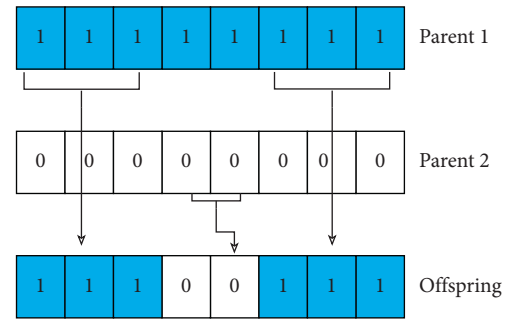


FIGURE 4: N-point crossover.

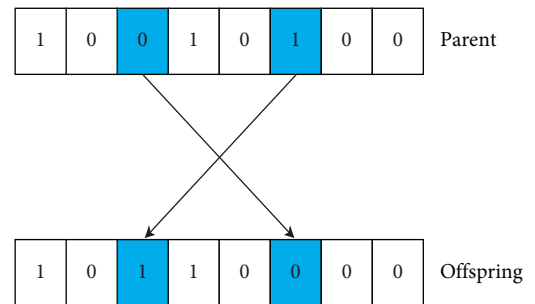


FIGURE 5: Displacement mutation.

3.1.5. *Termination*. The process of generation is recurrent until an end state is attained. Specifically, the termination criteria include the following: the value of the objective function reaches a certain predefined value, the number of iterations reaches the preset maximum iteration, the time or calculation cost of the budget allocation is reached, the

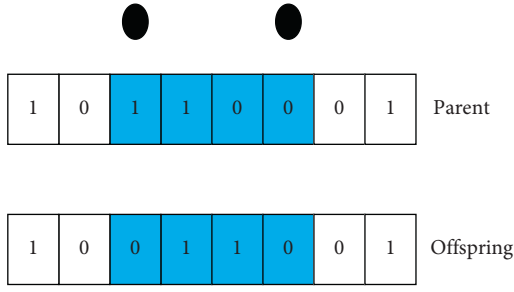


FIGURE 6: Shift mutation.

objective function value does not improve within a certain number of iterations, and various mixed termination criteria of the above termination conditions.

**3.2. GA Pseudocode.** The GA process is concisely analyzed in this section. Firstly, Figure 7 demonstrates the GA flow chart which describes the step-by-step approach of the proposed genetic algorithm. Secondly, the pseudocode for the GA is shown in Algorithm 1.

#### 4. Design of Localization Problem for WSN

**4.1. Description of Localization.** Since the received signal strength (RSS) measurements are comparatively low-priced and simple to implement in hardware, they are widely used in real-world localization applications and are also approved as a preserved confined measurement in most research projects. The RSS at a reliable place within a localization area can be stated as follows:

$$\text{RSS}(d_o) = P_t + K_e - 10\eta \log\left(\frac{d_o}{d_1}\right) + \alpha + \beta, \quad (3)$$

where  $P_t$  symbolizes the nominal transmission power (dBm),  $K_e$  stands for a constant depending on the system,  $\eta$  signifies the path loss coefficient,  $d_1$  is a reference distance for the antenna in far field,  $\alpha$  denotes the fast fading effect, and  $\beta$  represents the random attenuation sparked by shadowing. The RSS is analyzed by adjusting  $d_o$  which is the real transmitter-receiver distance.

The anchor nodes are nodes whose precise locations are recognizable prior to localization process. These can also be referred to as known nodes. When  $A$  represents the number of anchor nodes, a group of all WSN nodes with known locations is represented as  $KN$ . Thus, a known node position  $K_x$  is denoted by  $(p_{kx}, q_{kx})$ . Furthermore, unknown nodes are nodes whose location is calculated using a particular localization algorithm. The group of unrecognized WSN nodes is represented using  $UN$ :

$$\left. \begin{aligned} KN &= K_x | x = 1, 2, \dots, A \\ UN &= U_x | x = 1, 2, \dots, B - A \\ RN &= E_x | x = 1, 2, \dots, C \end{aligned} \right\} \quad (4)$$

Consider  $B - A$  as the quantity of unrecognized nodes. In real-time request, the actual positions of  $U_x$  defined by  $(p_{ux}, q_{ux})$  are undiscoverable. Let us assume the radius of the communication range to be  $R$ . Given that two energized sensor nodes are represented by  $p_x$  and  $p_y$ , if  $p_x$  is placed in

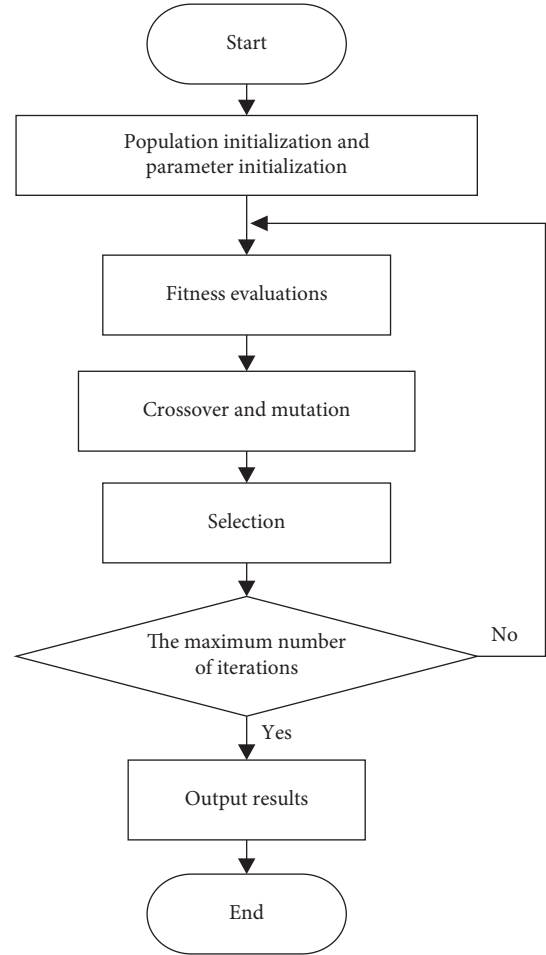
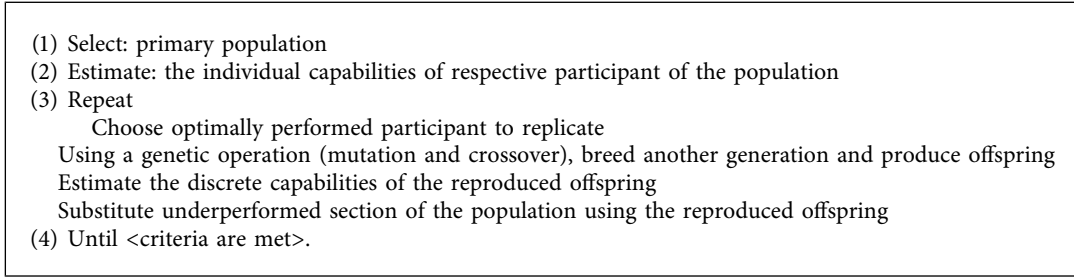


FIGURE 7: Flowchart for the genetic algorithm.

the broadcasting range of  $p_y$ ,  $p_x$  is directly considered a neighbor of  $p_y$ . As long as all the energized WSN nodes are endowed with very similar range of transmission,  $p_y$  is correspondingly a neighbor of  $p_x$ .

The likelihood of locating a node without a specific location is given as an estimated point because other unknown nodes are also being located during the positioning. The estimated position, actual position, and reference position are shown in Figure 8. The estimated position of  $U_x$  is represented by  $(p_{ux}^0, q_{ux}^0)$ . Apparently, the rationale behind localization is to make  $(p_{ux}^0, q_{ux}^0) = (p_{ux}, q_{ux})$  for  $U_x$ . The reference nodes comprise localized unknown nodes and known sensor nodes in the course of trying to locate an energized sensor node. The array of reference nodes is symbolized by  $RN$ , where  $B \geq C \geq A$ . The reference node point  $E_x$  with an exact position is defined by  $(p_{ex}, q_{ex})$ . In fact, if  $E_x$  is anchoring  $K_y$ , then  $(p_{ex}, q_{ex}) = (p_{ky}, q_{ky})$ . However, if  $E_x$  contains localized unknown nodes  $U_k$ , at the moment  $(p_{ex}, q_{ex}) = (p_{uk}^0, q_{uk}^0)$ .

The actual distance which is represented by  $d_{xy}$  is the distance covered concerning the actual points of  $U_x$  as well as  $E_x$ . The measurement distance  $d_{xy}^l$  is obtained by a certain measurement method based on the error  $z$ , and this error is determined by the random used measuring instrument. For convenience, follow-up studies generally use random value instead of this error. We presuppose that



ALGORITHM 1: Pseudocode for the GA.

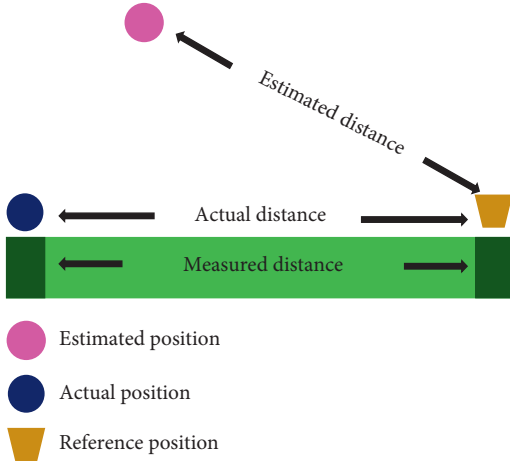


FIGURE 8: Distance between sensor nodes.

$d_{xy}^1 = d_{xy} + N(0, d_{xy}z)$ , where  $(N(0, d_{xy}z))$  is the Gaussian function with a mean value of 0 and a variance of  $d_{xy}z$ .

The distance covered by the estimated positions  $U_x$  and  $E_y$  is denoted by  $d_{xy}^0$  (estimated distance). Suppose unknown node  $U_x$  has  $m$  neighbor reference nodes  $E_1, E_2, \dots, E_m$ , where  $y = 1, 2, \dots, m$ . We can derive the following equations to get  $(p_{ux}^0, q_{ux}^0)$ :

$$d_{xy}^1 = \sqrt{(p - p_{ey})^2 + (q - q_{ey})^2}, \quad (5)$$

where  $(p, q)$  is an unknown dimension to resolve and  $p_{ey}, q_{ey}$  is the position of  $E_y$ . Due to the existence of the distance measurement error  $z$  and the estimated position  $(p_{ux}^0, q_{ux}^0)$  it is impossible to know the actual point (position) of  $E_y$  in a broad sense. Making use of the estimated position  $(p_{ux}^0, q_{ux}^0)$ , the estimated distance  $d_{xy}^0$  at that point is expressed as

$$d_{xy}^0 = \sqrt{(p_{ux}^0 - p_{ey})^2 + (q_{ux}^0 - q_{ey})^2}. \quad (6)$$

Because of the uncertainty of  $d_{xy}$ , that is, the actual distance is different, and the goal of positioning is to achieve an infinitesimal distance from  $d_{xy}^1$  to  $d_{xy}^0$ . Finally, we construct the location problem denoted  $U_x$  as

$$\sum_{y=1}^n w_y (d_{xy}^0 - d_{xy}^1)^2, \quad (7)$$

$$\sum_{y=1}^n w_y \left( \sqrt{(p - p_{ey})^2 + (q - q_{ey})^2} - d_{xy}^1 \right)^2,$$

where  $w_y = (1/d_{xy}^1) \sum_{x=1}^m (1/d_{xy}^1)$ , which gives better understanding pertaining to the reference point closer to  $U_x$ . In factual terms, the distance covering the communication range of an energized sensor is a half-way circle caused by multipath fading, asymmetrical message delivery, and redundant noise.

In finding the minimal localization error of unidentifiable (unknown) location points  $U_x$ , the change in estimated and actual location point should be always be considered which is found in the equation below.

$$LE_x = \frac{1}{R} \sqrt{(p_{ux}^o - p_{ux})^2 + (q_{ux}^o - q_{ux})^2}. \quad (8)$$

**4.2. Clustering Model.** During the course of node clustering, the most approximate energized sensors are observed to be in the same locality (cluster) which tries to save energy by reducing the transmission range and the closest point amidst the energized sensors. Figure 9 shows a setup of the suggested clustering scheme. The key concern is to discover the precise location which depends on several decisions on how to locate it. With the goal of finding the preferable location for a particular energized sensor, the distance of a sensor node is calculated using (5). Our new approach for well-organized clustering splits the entire WSN nodes using Euclidean distance connecting sensor nodes into numerous clusters. However, equal cluster size must be assured at some point in the clustering process. Taking one cluster into consideration, the sensor nodes are however placed in order to minimize (14), that is, the Euclidean distance between the location points and their immediate central point. Therefore, if the location point is initiated with a sensing range  $R$  in a deployment area consisting of energized sensors at the central point, then it is said to be covered. Consequently, the distance of a location point  $p_x$  and the central sensor node at a point  $q_c$  should be less than or equal to the distance between a location point  $p_x$  and any energized sensor node at point  $q_y, \forall y = 1, 2$  to  $R$ , and it is mathematically represented as  $d(p_x, q_c) \leq d(p_x, q_y)$ .

## 5. Proposed Fitness Approach for ECGAL

In this section, we derive the fitness function for the proposed energy-efficient clustering and localization using a genetic algorithm.



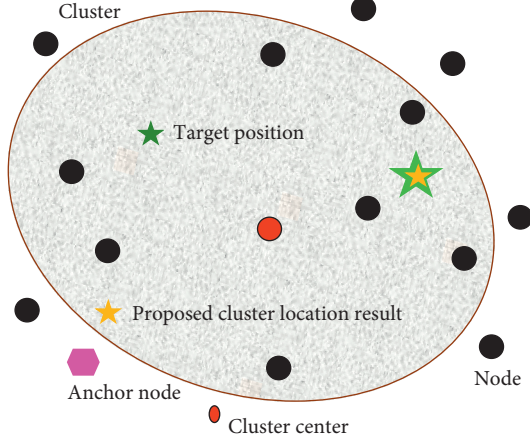


FIGURE 9: Expected WSN cluster structure.

**5.1. Energy Efficiency.** The ability for a chromosome to withstand all conditions helps it in lowering the energy exhausted and to maximize lifespan of the network system. The channel description for free-space and multipath fading used still considered the sum of distance connecting the receiver and the transmitter. Suppose the upper threshold value  $d_o$  is greater than the node distance pairs  $d$ , then the energy amplification consumption assumes a free-space model, but if  $d$  is greater than or equal to  $d_o$ , a multipath decay model is implemented. Therefore, the amount of energy needed by the radio to convey a  $h$ -bit message over a distance  $d$  is given in (9). The radio also uses up energy to accept a  $h$ -bit message given in (10).  $E_{elec}$  depends on factors for instance modulation, filtering, digital coding, and combining the dispersion of signals, but the amount of energy to amplify the system,  $\varepsilon_{fs}d^2$  or  $\varepsilon_{mp}d^4$ , relies on the receiving structure per the distance travelled and the suitable error per bit.  $E_{elec}$  is defined as the electronic energy required by the electronic circuit and  $d_0 = \sqrt{\varepsilon_{fs} \setminus \varepsilon_{mp} \cdot \varepsilon_{fs}}$  and  $\varepsilon_{mp}$  are the amplifier energies in free space and multipath, respectively. We denote  $E_i$  to be the residual energy once a message is communicated through  $h^{\text{th}}$ -bits at a distance  $d$  from the receiver in (11).  $E$  is the node's recent energy. Then,  $E_1 = E_T(h, d) + E_R(h)$  is the energy needed to send a message plus the energy consumed while receiving a message.

$$E_T(h, d) = \begin{cases} h \times E_{elec} + h \times \varepsilon_{fs}d^2, & d < d_o, \\ h \times E_{elec} + h \times \varepsilon_{mp}d^4, & d \geq d_o, \end{cases} \quad (9)$$

$$E_R(h) = h \times E_{elec}, \quad (10)$$

$$E_i = E - E_1. \quad (11)$$

**5.2. Distance Estimation.** The total distance covered starts with an energized node point to another sensor point which is assumed to be the distance between two neighboring sensor nodes. It is expressed as  $\text{dist}(p_a, q_b)$ . However, the distance from a locational node to the central cluster position should

be insignificant in order to get closer to our target compared to the distance from the cluster center to another node. The latter is represented by  $\text{dist}(p_a, q_c)$  and the former is denoted as  $\text{dist}(p_c, q_b)$ . This should be ensured orderly so that the wastage of energy of each node in a large network is minimal. This boosts the cluster strength and reduces the lack of sensor node involvement. For all node points  $p_x \in N$ , where  $N$  is the set of all nodes, we compute the sum of the distance  $D_i$  with all its neighboring points  $q_y$ . However, these energetic neighboring points could be the position of a node without location which could be activated using the position of a known node. This distance is given in (14).

$$D_G^0 = \sum_{p_a \in N} \text{dist}(p_a, q_b), \quad (12)$$

$$D_G^1 = \sum_{q_b \in N} \text{dist}(p_a, q_c) + \text{dist}(p_c, q_b), \quad (13)$$

$$D_i = \sum_{p_y \in N} \text{dist}(p_x, q_y). \quad (14)$$

**5.3. Coverage Connection.** Every WSN can be considered as a connected undirected figure denoted by  $G = (V, E)$ , where  $V$  consists of vertices comprising of  $\{v_1, v_2, \dots, v_u\}$  which denotes the energized node point found in the WSN along with  $E$  which is the edge set  $\{e_1, e_2, \dots, e_f\}$  representing the distance between the energized sensor nodes. This approach considers the weighted values depending on energy efficiency, distance estimation, and coverage connection which are represented on the edges. Furthermore, every edge in the network possesses a finite real number which is represented as  $w_i$ . Let the sensing range of a node be denoted by  $S_r$ . Let  $c = c_1, c_2, \dots, c_m$  be the connectivity variables associated with the energized sensor nodes  $p_x$  and  $p_y$ . However,  $C_y$  is the area covered by  $y^{\text{th}}$  cluster center node,  $N$  is the sum of all recognized energized sensor nodes, and  $C$  is the WSN area.

$$C_L = \begin{cases} 1, & \text{if } \|p_x - p_y\| \leq S_r, \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

$$C_i = \bigcup_{y=1}^N C_y \frac{C_y}{C}.$$

The final fitness function that demands to be computed in minimization is given below, and it constitutes the previous fitness minor objectives:

$$F_i = w_1 E_i + w_2 D_i + w_3 C_i, \quad (16)$$

where  $w = \{w_1, w_2, \dots, w_f\}$  is the distance associated with the edges. We define  $w_1, w_2, w_3$  to be the weight coefficients attached to the fitness function in order to quantify its contribution to each of the other subfunctions, and it is expressed by  $\sum_{i=1}^3 w_i \geq 0, w_i \in (0, 1)$ .

## 6. Performance Evaluation

The performance of our approach is evaluated in this section. The device used for the evaluation was Intel(R) Core (TM) i5-3317U CPU PC with 6144 RAM, which was accomplished using MATLAB 2014a. It is compared with existing approaches such as DV-HOP (distance vector-hop), CENTA (centroid algorithm), EDV-HOP (evolutionary distance vector-hop), and CGAL (clustering in genetic algorithm localization). The network scenario is considered realistic in nature with 200 energized sensor nodes randomly deployed with 25% anchor nodes, several unknown nodes, and reference nodes. The experimental parameters deployed in this work are further presented in Tables 1 and 2.

Figure 10 provides a solid evidence to conclude that the new approach performs impressively compared to other location-based algorithms in error location. Almost all the applied approaches work robustly in the same configuration. ECGAL drops gently because of the additional anchor nodes in the network which provided more reference points for the target nodes. However, the network is boosted when there are sufficient anchor nodes because the distance joining the unknown nodes and the anchor nodes gets smaller. In our simulation output, CGAL, EDV-HOP, and CENTA showed less localization errors as well.

In Figure 11, it is assumed that as the transmission range expands, the number of seconds of continuous simulation operation increases. Simultaneously, the transmission range starts from 5 m and steadily increases by 5 m so as to evaluate the performance of our approaches. The ability to locate a node improves and could be achieved when the radius of transmission inclines which reduces the error in localization. Finally, when the transmission range increases, ECGAL obtains better results in terms of location accuracy.

The experimental results in Figure 12 show the task of location error computed against varying node numbers. On top of it all, as the number of the energized sensor nodes increases, the localization error for all the algorithms decreases slowly. Among all our localization approaches, ECGAL shows fewer points for its localization error. As the number of nodes reaches 200, more reference points are found which help to localize the node with less error. However, as there is increase in energized nodes, there is also slack in lifetime contributing factor for CENTA and DV-HOP.

Figure 13 shows the localization error against the number of clusters considering different algorithms. The clustering technique proposed in this paper improves the energy efficiency in the network. With the increase of cluster number, the localization error decreased. ECGAL and CGAL dropped slowly because when the number of clusters is high, fewer nodes will be found in their clusters, which makes it easier to locate an unknown node, thereby reducing the localization error drastically. However, CENTA is seen to perform almost similar to EDV-HOP because of its special clustering abilities. The energy depletion level of a network enhances if there are some reasonable number of clusters. In DV-HOP, the transmission scale has enough energized

TABLE 1: Parameters used in sensor field and GA.

Simulation parameters	Value
Total number of nodes	200
Deployment field area	200 * 200 m <sup>2</sup>
Communication range $R$	40 m
Number of anchor nodes	50
Number of unknown nodes	100
Maximum iterations	300
Number of clusters	5
Population size	50
Length of chromosome	5
Number of generations	150
Mutation rate	0.5
Crossover % rate	0.8

TABLE 2: Parameters used in the energy model.

Parameters	Values
Initial energy	2 J
Distance ( $d_0$ )	87 m
Packet size	200 bits
Energy for transmitting ( $E_T$ )	50 nJ/bit
Energy for receiving ( $E_R$ )	50 nJ/bit
Energy for data aggregated ( $E_D$ )	50 nJ/bit/signal
Energy consumption of power amp in free space ( $\epsilon_{fs}$ )	10 pJ/bit/m <sup>2</sup>
Energy consumption of power amp in multipath fading ( $\epsilon_{mp}$ )	0.0013 pJ/bit/m <sup>4</sup>

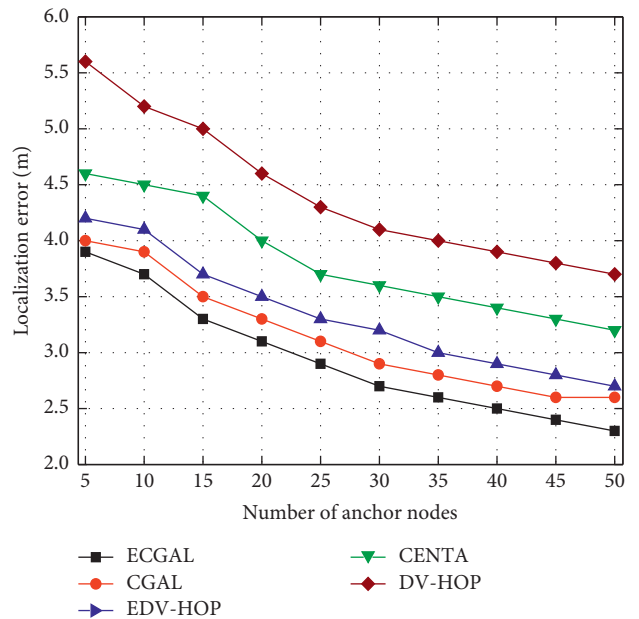


FIGURE 10: Localization error against the number of anchor nodes.

nodes, which indicates that more of these nodes are found in each cluster.

Figure 14 depicts the residual energy against the number of iterations. The energy savings are meaningful in ECGAL compared to CGAL, EDV-HOP, CENTA, and DV-HOP. All the approaches dropped more and more until 80 iterations where they start to drop significantly. The amount of residual



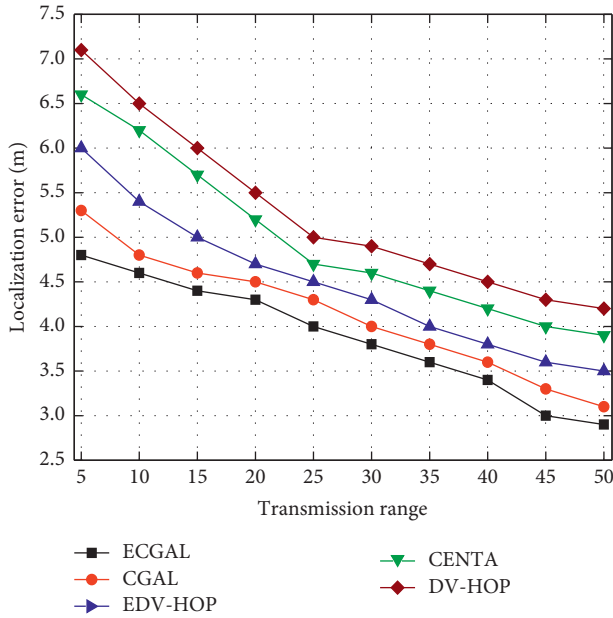


FIGURE 11: Localization error against transmission range.

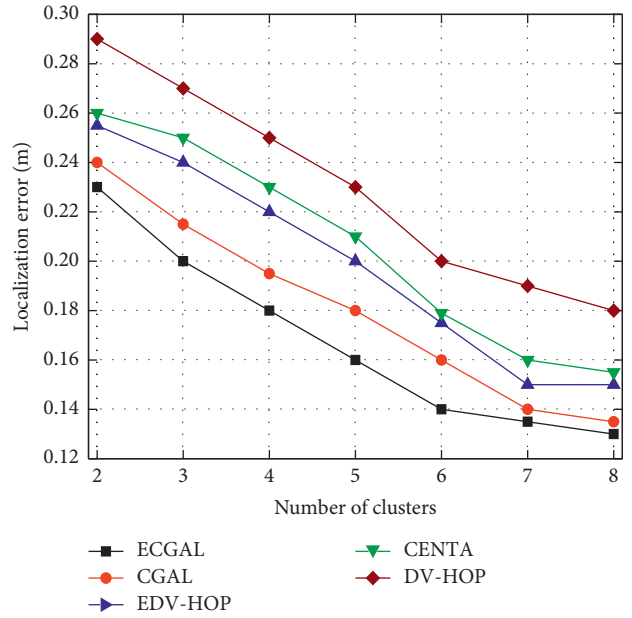


FIGURE 13: Localization error against the number of clusters.

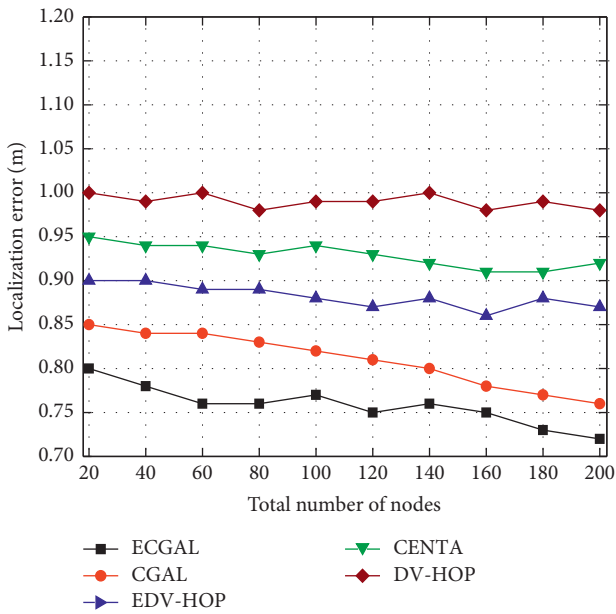


FIGURE 12: Localization error against the total number of nodes.

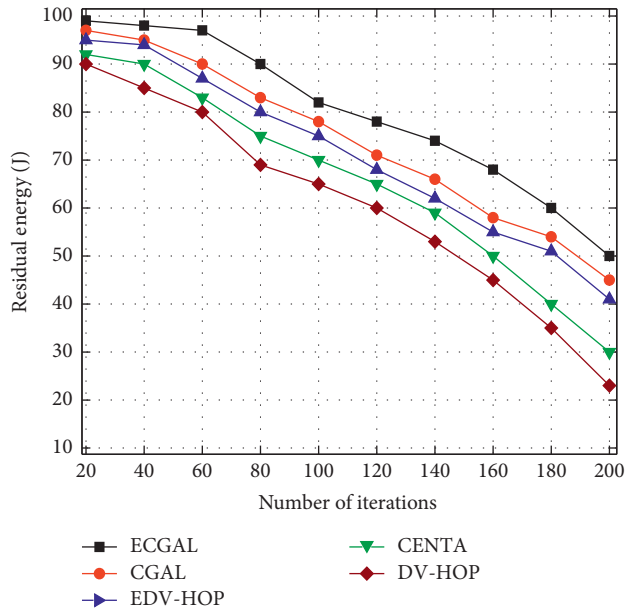


FIGURE 14: Residual energy against the number of iterations.

energy found in ECGAL surpasses CGAL because the energy left behind after 120 iterations is about 70 J. This could be due to the optimal election of cluster centers and the equidistance between the intracluster and intercluster. At the initial position, the performances of CGAL and EDV-HOP are the same. According to our graph, the energy remaining for EDV-HOP and CENTA after 160 iterations was 55 J and 50 J, respectively. Finally, the decline in residual energy affects the life expectancy of the network which in turn increases the number of exchanged control packets (overhead).

In Figure 15, the time taken to process the algorithm is computed for 200 iterations over seconds. The number of messages generated to messages sent to the final point is described as the total success rate of packets delivered. Finally, ECGAL performs better in terms of the convergence rate which is best compared to CGAL, EDV-HOP, CENTA, and DV-HOP. It is clear that the ECGAL proves its success in transporting about 90% information to its final destination. With an increase in the number of iterations, CGAL and EDV-HOP showed better execution compared to CENTA and DV-HOP. The performance output of CENTA and DV-HOP is almost similar.

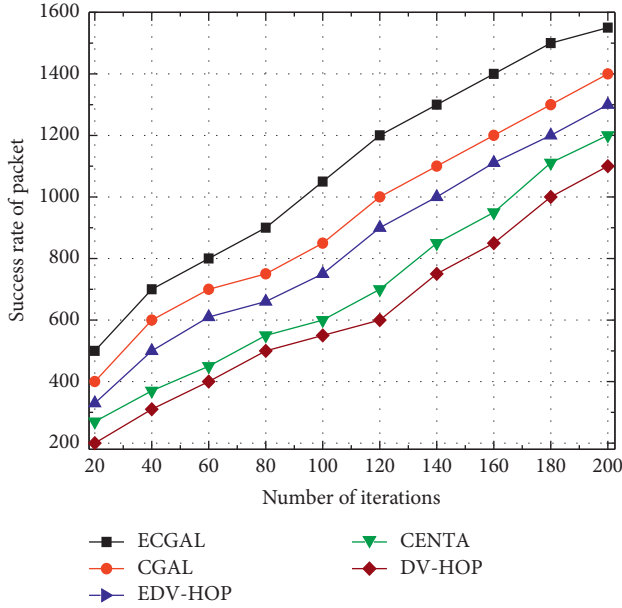


FIGURE 15: The success rate of packets against the number of iterations.

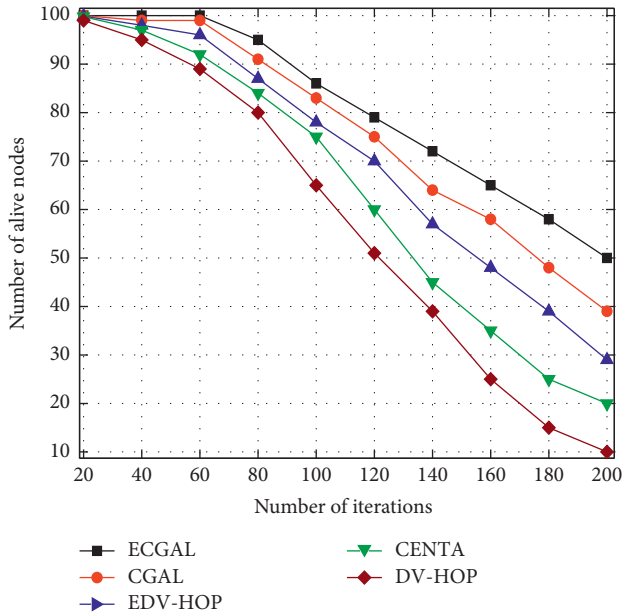


FIGURE 16: The number of alive nodes against the number of nodes.

Figure 16 shows the lifetime of the network. CGAL and the proposed technique radically improved the life of the network compared to EDV-HOP, CENTA, and DV-HOP techniques. The live nodes are evaluated by increasing the number of iterations to 200. The energy level of the networks' energized sensor nodes drains after several iterations. When the number of iterations reaches 180, compared with CGAL's 50 active nodes, ECGAL has 60 active nodes. At the same time, EDV-HOP and CENTA have only about 40 and 25 active nodes, respectively. The lines of the active nodes in the figure indicate that our method has a longer life span compared with other methods. On account of the developed

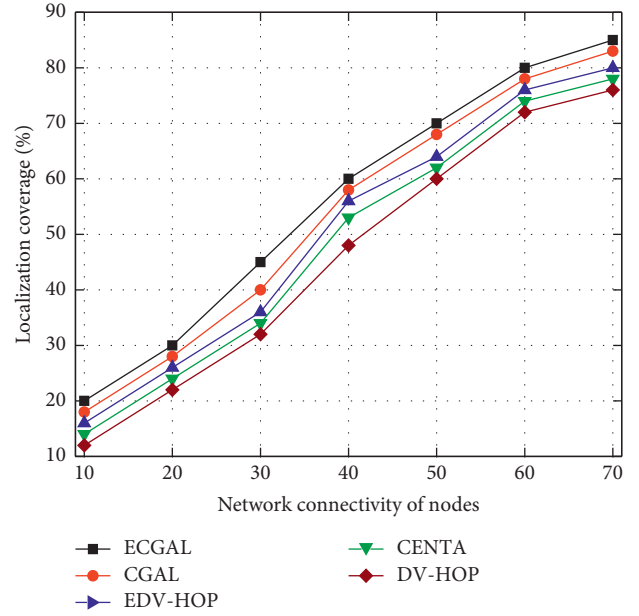


FIGURE 17: Localization coverage against network connectivity of nodes.

energy-efficient clustering localization approach based on genetic algorithm, ECGAL showed better output than that of the CGAL, EDV-HOP, CENTA, and DV-HOP algorithms.

In Figure 17, the influence of connected nodes is studied by analyzing the coverage against the network node correlation starting from 10 through 70. In our experiment, 200 energized sensors occupy 200 by 200 m<sup>2</sup> deployment area. With an increase in connected nodes, the localization coverage for ECGAL, CGAL, EDV-HOP, CENTA, and DV-HOP also increases. When the readings on horizontal axis reach 50 connected nodes, the network coverage still increases for ECGAL and CGAL. When the coverage location indicator accumulates 70 bars, the strength of the node to node relationship still goes higher for EDV-HOP, CENTA, and DV-HOP. The density of neighboring energized sensors warrants cost-effective and well-built connection between known and unknown node points. In general, our improved positioning method is superior to other existing methods in terms of accuracy.

## 7. Conclusion

The proposed ECGAL approach for ultimate location problem shows significant results after employing, and it proved that the energy-efficient clustering based on genetic algorithm localization approximates the node that demands to be identified and later assures a minimal location error when matched with DV-HOP, CENTA, EDV-HOP, and CGAL. ECGAL is better due to its efficient energy clustering strategy. In fact, our improved approach for better localization reconstructs our solution to quickly detect the location of the unidentifiable sensor node. However, nodes with known location point are randomly dispersed in an exact WSN because of the randomly deployed energized node point. For that reason, anchor nodes assist in locating

wherever unknown sensors are even though the converse is true, and thus excessive neighboring known points cause more nodes to be unlocalized. In conclusion, we can acknowledge that the proposed ECGAL performs effectively when studied with other approaches in relation to true position point and minimal error in terms of location.

## Data Availability

The data used to support the findings of this study can be accessed from the following link: [https://www.researchgate.net/publication/349836503\\_ECGAL](https://www.researchgate.net/publication/349836503_ECGAL).

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the National Key R&D Program of China (no. 2018YFC0407101) and the Fundamental Research Funds for the Central Universities (no. 2019B22314).

## References

- [1] E. Fernandes, A. Rahmati, K. Eykholt, and A. Prakash, "Internet of things security research: a rehash of old ideas or new intellectual challenges?" *IEEE Security & Privacy*, vol. 15, no. 4, pp. 79–84, 2017.
- [2] R. Kaur and S. Arora, "Nature inspired range based wireless sensor node localization algorithms," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 6, pp. 7–17, 2017.
- [3] W. Lan, W. Zhang, and J. Luo, "Design and implementation of adaptive intelligent trilateral localization algorithm," *Chinese Journal of Sensors and Actuators*, vol. 30, no. 7, pp. 1089–1094, 2017.
- [4] Y. Liu and J. Chen, "A K-means based firefly algorithm for localization in sensor networks," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 34, no. 4, pp. 364–379, 2018.
- [5] P. Singh, A. Khosla, A. Kumar, and M. Khosla, "Optimized localization of target nodes using single mobile anchor node in wireless sensor network," *AEU - International Journal of Electronics and Communications*, vol. 91, pp. 55–65, 2018.
- [6] S. K. Rout, A. K. Rath, P. K. Mohapatra, P. K. Jena, and A. Swain, "A fuzzy optimization technique for energy efficient node localization in wireless sensor network using dynamic trilateration method," *Advances in Intelligent Systems and Computing in Progress in Computing, Analytics, and Networking*, vol. 1, pp. 325–338, 2018.
- [7] A. Azadeh, S. Goldansaz, and A. Zahedi-Anaraki, "Solving and optimizing a bi-objective open shop scheduling problem by a modified genetic algorithm," *The International Journal of Advanced Manufacturing Technology*, vol. 85, no. 5-8, pp. 1603–1613, 2016.
- [8] S. Anthrayose and A. Payal, "Comparative analysis of approximate point in triangulation (APIT) and DV-HOP algorithms for solving localization problem in wireless sensor networks," in *Proceedings of 7th IEEE International Advanced Computing Conference*, pp. 372–378, Hyderabad, India, January 2017.
- [9] T. Najeh, H. Sassi, and N. Liouane, "A novel range free localization algorithm in wireless sensor networks based on connectivity and genetic algorithms," *International Journal of Wireless Information Networks*, vol. 25, no. 1, pp. 88–97, 2018.
- [10] S. P. Singh and S. C. Sharma, "A PSO based improved localization algorithm for wireless sensor network," *Wireless Personal Communications*, vol. 98, no. 1, pp. 487–503, 2018.
- [11] S. H. Sackey, J. A. Ansere, J. H. Anajemba, M. Kamal, and C. Iwendi, "Energy efficient clustering based routing technique in WSN using brain storm optimization," in *Proceedings of 2019 15th International Conference on Emerging Technologies (ICET)*, pp. 1–6, Peshawar, Pakistan, May 2019.
- [12] S. Sreenivasamurthy and K. Obraczka, "Clustering for load balancing and energy efficient in IoT applications," in *Proceedings of 2018 13th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, pp. 319–332, Atlanta, GA, USA, August 2018.
- [13] J. Wang, Y. Gao, K. Wang, A. K. Sangaiah, and S. J. Lim, "An affinity propagation-based self-adaptive clustering method for wireless sensor networks," *Sensors*, vol. 19, no. 11, p. 2579, 2019.
- [14] S. H. Sackey, J. Chen, J. A. Ansere, G. K. Gapko, and M. Kamal, "A bio-inspired technique based on knowledge discovery for routing in IoT networks," in *Proceedings of 2020 IEEE 23rd International Multitopic Conference*, pp. 1–6, Bahawalpur, Pakistan, November 2020.
- [15] Z. Yang, C. Liu, and L. Jin, "A clustering-based algorithm for device-free localization in IoT," in *Proceedings of 2018 IEEE 4th International Conference on Computer and Communications*, pp. 769–773, Philadelphia, PA, USA, December 2018.
- [16] P. T. Daely and D.-S. Kim, "Bio-inspired cooperative localization in industrial wireless sensor network," in *Proceedings of 15th IEEE International Workshop on Factory Communication Systems*, Sundsvall, Sweden, May 2019.
- [17] H. M. Kanoosh, E. H. Houssein, and M. M. Selim, "Salp swarm algorithm for node localization in wireless sensor networks," *Journal of Computer Networks and Communications*, vol. 2019, Article ID 1028723, 12 pages, 2019.
- [18] F. Zhang, H.-F. Xue, Y.-H. Zhang, and F. You, "A new localization and tracking algorithm for wireless sensor networks based on the internet of things," *Sensors and Transducers*, vol. 154, no. 7, pp. 56–61, 2013.
- [19] H. MUSAFAER, R. Abdulhameed, E. Abdelfattah, and K. Elleithy, "A dynamic clustering algorithm for object tracking and localization in WSN," in *Proceedings of 27th International Conference on Computer Applications in Industry and Engineering, at New Orleans*, New Orleans, LA, USA, September 2014.
- [20] C. Alippi, M. Bocca, G. Boracchi, N. Patwari, and M. Roveri, "RTI goes wild: RTI goes wild: radio tomographic imaging for outdoor people detection and localization," *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2585–2598, 2016.
- [21] T. Liu, X. Luo, and Z. Liang, "Enhanced sparse representation-based device-free localization with radio tomography networks," *Journal of Sensor and Actuator Networks*, vol. 7, no. 1, p. 7, 2018.
- [22] C. Xu, B. Firner, Y. Zhang, and R. E. Howard, "The case for efficient and robust RF-based device-free localization," *IEEE Transactions on Mobile Computing*, vol. 15, no. 9, pp. 2362–2375, 2016.

- [23] L. Lin and L. Donghui, "An energy-balanced routing protocol for a wireless sensor network," *Journal of Sensors*, vol. 2018, Article ID 8505616, 12 pages, 2018.
- [24] J. Wang, D. Fang, Z. Yang et al., "E-HIPA: an energy-efficient framework for high-precision multi-target-adaptive device-free localization," *IEEE Transactions on Mobile Computing*, vol. 16, no. 3, pp. 716–729, 2016.
- [25] L. Zhang and T. N. Wong, "An object-coding genetic algorithm for integrated process planning and scheduling," *European Journal of Operational Research*, vol. 244, no. 2, pp. 434–444, 2015.
- [26] S. H. Sackey, J. Chen, A. J. Henry, and X. Zhang, "A clustering approach based on genetic algorithm for wireless sensor network localization," in *Proceedings of 2019 15th International Conference on Computational Intelligence and Security (CIS)*, pp. 54–58, Macao, Macao, May 2019.
- [27] J. Wang, X. Zhang, Q. Gao, H. Yue, and H. Wang, "Device-free wireless localization and activity recognition: a deep learning approach," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6258–6267, 2017.
- [28] S. Sivakumar, "Error minimization in localization of wireless sensor networks using fish Swarm optimization algorithm," *International Journal of Computer Applications*, vol. 159, no. 7, pp. 39–45, 2017.
- [29] L. Song, L. Zhao, and J. Ye, "DV-hop node location algorithm based on GSO in wireless sensor networks," *Journal of Sensor*, vol. 2019, Article ID 2986954, 9 pages, 2019.
- [30] G. Kumar and M. K. Rai, "An energy-efficient and optimized load-balanced localization method using CDS with the one-hop neighborhood and genetic algorithm in WSNs," *Journal of Network and Computer Applications*, vol. 78, no. 73–82, 2017.
- [31] J. Wang, Y. Gao, X. Yin, F. Li, and H.-J. Kim, "An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 9472075, 9 pages, 2018.
- [32] G. Sharma and A. Kumar, "Modified energy-efficient range-free localization using teaching-learning-based optimization for wireless sensor networks," *IETE Journal of Research*, vol. 64, no. 1, pp. 124–138, 2018.
- [33] J. H. Anajemba, T. Yue, C. Iwendi, M. Alenezi, and M. Mittal, "Optimal cooperative offloading scheme for energy efficient multi-access edge computation," *IEEE Access*, vol. 8, pp. 53931–53941, 2020.