IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 9 July 2022, accepted 17 August 2022, date of publication 1 September 2022, date of current version 12 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3203696

RESEARCH ARTICLE

A New Crossover Methods and Fitness Scaling for Reducing Energy Consumption of Wireless Sensor Networks

WAFA ALSHARAFAT[1], ALAA AL-SHDAIFAT[2], KHALED BATIHA[2], AND AKRAM ALSUKKER[3]

[1]Department of Information System, Al Albayt University, Mafraq 25113, Jordan
[2]Department of Computer Science, Al Albayt University, Mafraq 25113, Jordan
[3]Department of Industrial Engineering, Faculty of Engineering, The Hashemite University, Zarqa, Jordan

Corresponding author: Wafa Alsharafat (researcher2010@yahoo.com)

ABSTRACT Efficient energy in the wireless sensor networks (WSNs) is a critical issue because sensor nodes are equipped with one-time or low-energy batteries. In these networks, efficient energy saving methods involve clustering network nodes to avoid long-distance communications with base stations (BS) to conserve their energy over a long period of time and extend their lifetimes. In other words, the choice of cluster heads (CHs) to improve routing and energy efficiency plays a central role in extending the network lifetime. This study proposes a new central cluster algorithm based on an improved genetic algorithm (EGA) that finds appropriate numbers of CHs in networks. This enhancement concerns about the application of two new crossover methods: Whole Arithmetic Crossover (WOX), and Local Crossover (LX) methods. This study explored the impact of the two aforementioned crossover methods on WSN energy efficiency and the effect of applying the scaled fitness function on network lifetime. For evaluation, we conducted a comparison between the crossover methods WOX and LX with three crossover methods; Simple Arithmetic Crossover (SMX), Single Arithmetic Crossover (SNX), and Discrete Crossover (DX) considering fitness scaling or without fitness scaling to identify the method that influences energy consumption. The results were then compared with the Low Energy Adaptive Clustering Hierarchy protocol (LEACH). All the simulation experiments were performed in MATLAB. The simulation results reveal that WOX and SNX with a scaled fitness function lead to a longer network lifetime by selecting CHs with longer lifetimes than the SMX, DX and LX methods. As a result, the proposed method exhibited better performance in terms of the power consumption and throughput rate.

INDEX TERMS Fitness-scaling, genetic algorithm (GA), local crossover (LX), whole arithmetic crossover (WOX), wireless sensor network (WSN).

I. INTRODUCTION

The history of Wireless Sensor Networks (WSNs) usage has begun in the military and defense industries. This type of network has a wide range of applications owing to its affordability and multi-functionality. These WSNs have adaptable applications such as a real-time volcano warning system, which is used for high-risk and dangerous environmental monitoring. [1], monitoring enemy territory [2], healthcare systems [3] and the Intelligence Traffic Management System (ITMS) [4]. For such applications, several sensor networks (SNs) and a set of Base stations (BSs) are required to achieve the required mission. In general, WSN must be able to operate autonomously for extended periods of time in the majority of their applications.

One of the main advantages of WSN is that they can operate in high-risk environments, especially in humans, such as active volcanos and earthquakes. However, the deployment of thousands of low-cost SNs gives these sensing applications new advantages such as reducing the cost of nodes, improving network accuracy and reliability, and extending the scope of the sensor.

The associate editor coordinating the review of this manuscript and approving it for publication was Hisao Ishibuchi.

Furthermore, constraints on the size and cost of SNs generally lead to constraints on resources such as energy, memory, computational speed, and the network bandwidth needed to make resource use very efficient. WSN plays an important role in energy and life, and the main barrier to WSN applications is the threat of restricted energy resources.

SNs are usually equipped with one-time batteries, most of which are low-energy [6]. Furthermore, in most applications of WSNs, it is often infeasible or very difficult to recharge or replace the batteries attached to SNs once they are deployed. Therefore, the energy efficiency of such networks is crucial. In a WSN, the energy is consumed during data sensing, processing, and transmission. However, a node consumes almost 90% of the overall energy used in communications [7]. Therefore, the design and development of an energy-efficient routing protocol are the main goals for prolonging the lifetime of a WSN and improving its performance.

Energy-efficient, routing-based clustering is important to improve the lifetime of a WSN by virtue of clustering advantages, data aggregation or fusion, and CH selection techniques. LEACH is the main clustering-based, energy-efficient, distributed routing protocol that is commonly used by researchers to analyze, modify, and extend clustering routing protocols and compare their performance.

Many researchers have investigated ways to improve network lifetime through energy-efficient routing designs. However, the design of an energy-efficient routing algorithm for WSN is essentially an optimization problem. Several studies on energy consumption in WSN have been carried out using metaheuristic algorithms, as in [8], to present a survey that focuses on energy consumption in WSN and can guide researchers in term of energy efficiency.

Consequently, this study developed and presented a new centralized cluster algorithm based on the application of an enhanced genetic algorithm (CCA-EGA) to ensure that the network has an adequate CH to improve the efficiency of WSN energy. The major contributions of this study can be summarized as follows:

- Applying real coding in Genetic Algorithm (GA) based on the proposed clustering technique (CCA-EGA) to develop a more energy-efficient routing algorithm than that obtained from a binary representation.
- Developing a fitness-scaling function that facilitates highly energy-efficient routing for WSN.
- Examining the potential of several crossover operators to improve the GA's ability to find better solutions than the current solution.
- The ability of fitness scaling to improve the convergence efficiency of the genetic solution is greater than that of the traditional approach.

The remainder of this paper is organized as follows: Section II reviews previous related studies, and Section III presents the proposed GA for WSN clustering. Section IV analyses and discusses the experimental results of the GA for WSN clustering. Finally, Section V presents the conclusions of this study.

II. RELATED WORK

Several routing protocols have been developed for WSNs. The main criteria for the performance and quality of these protocols, and for differentiating between them are data collection efficiency, resource consumption reduction, and network lifetime maximization.

The Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol is the first, basic, energy-efficient routing protocol in cluster-based routing protocols. It was developed by Heinzelman in 2000 [9]. This was the pioneering protocols. This is the first protocol that applies a clustering technique for cost-effective energy management in WSNs. LEACH uses randomized rotation of the CHs such that the high-energy squandering caused by the communication of the WSNs with the BS is circularized to all sensor nodes in the network. Consequently, node selects CH by selecting a random number T between zero and one. In any round, the node classified as a CH is determined to be the node whose random number is less than a threshold value, which is calculated according to (1):

$$T(n) = \begin{cases} \frac{p}{1-p\left(r \bmod \frac{1}{p}\right)} & \text{if} n \in G \\ 0, & \text{Otherwise} \end{cases} \tag{1}$$

where p is the desired percentage of the CH nodes of the sensor population, In other words, it is the probability that a node is selected as a CH. For example, $p = 0.05$, r denotes the current or existing round number and G is the set of nodes that have not been CH in the previous 1/p rounds. Thus, the number of nodes in G is compared to T(n). If the number of the node is less than the threshold value T(n), then the node defines itself as a CH.

The LEACH protocol, as in [10], has some deficiencies in term of energy consumption and network lifetime, which prompted researchers to investigate its enhancement. One of these shortcomings is that the criterion for the selection of CHs in LEACH is not evenly positioned across the network, because the CHs are selected in this protocol based on the principle of probability. Consequently, during certain rounds, the selected CHs may be concentrated in a part or area of the network. Thus, LEACH randomly selects CHs. The amount of residual energy in the selected CH is not considered. This causes an imbalance in the energy load, which means that the nodes with low energy have the same likelihood of being selected as CHs, unlike nodes with high residual energy.

From another perspective, the distance factor is not considered in LEACH when selecting CHs. In general, the higher the distance between the CH and BS, the higher the energy consumption. Therefore, LEACH is unsuitable for large networks. When the BS is located far away, this leads to more energy consumption because the LEACH protocol assumes that all SNs can communicate with each other and that they have the ability to reach the BS, irrespective of where it is situated. Furthermore, in LEACH, the BS has no control over CH selection. This negatively affects the residual energy through the wastage of a large amount of energy if the CH is located far away from the BS. Moreover, a BS usually possesses high capabilities that should be utilized to assist in the routing process as much as possible. Furthermore, the failure of a CH sharply reduces the robustness of the network. The LEACH protocol uses dynamic clustering after the completion of each round, which results in extra overhead, that is, periodic changes in CH that adversely affect network performance by increasing energy consumption [11].

Therefore, it is necessary to prevent the rapid death of CHs, which mainly affects network lifetime [12]. In addition, there is a need to periodically change the CHs for each round in a specific manner to ensure a long network lifetime. However, optimal CHs can reduce energy consumption and extend the network lifetime. Subsequently, proposed CCA-EGA aims to identify and select the optimum CHs in terms of enhancing energy consumption and elongating the lifetime of the network, taking into account the residual energy, which plays a significant role in enduring the CHs.

Several researchers have focused in their works on improving the energy efficiencies of WSNs. For example, Heinzelman *et al.* [9] proposed one of the first and most popular standard clustering techniques, LEACH, and applied it to SNs. It is a dynamic, self-adaptive, probabilistic, and single-hop protocol. This protocol essentially forms clusters, and CH selection is based on weighted probability. The LEACH protocol employs distributed (or noncentralized) clustering method. The running process of the LEACH protocol depends on the round, which each round is divided into two phases. The first phase is the set-up phase, where the second phase is the steady-state phase. The main aim of the LEACH protocol is to distribute the energy load evenly among all sensor nodes in the network, elongate the network lifetime, and use data aggregation or fusion to compress data.

Heinzelman *et al.* [13] developed a Centralized Low Energy Adaptive Clustering Hierarchy (LEACH-C) protocol. The classical LEACH protocol applies a distributed clustering method to produce clusters and selects the CHs for the network through the nodes themselves. However, the LEACH-C protocol applies a centralized clustering method that sets the BS as a coordinator to organize the produced clusters and CH selection. The operation of the LEACH-C protocol was standardized in a manner similar to that the original LEACH protocol. The two protocols progress in two phases: set-up phase and steady-state phase. However, fundamentally, the set-up phase in LEACH-C is different from that in LEACH whereas the steady-state phase is the same as the set-up phase in the original LEACH protocol. During the set-up phase, at the beginning of each round of the LEACH-C protocol, each sensor node sends essential information about its residual energy and current location to the remote BS. After the BS receives the essential information from all the sensor nodes, the LEACH-C protocol calculates the average node energy based only on energy and checks or identifies it based on the following rules:

- A node whose energy is higher than the average energy can qualify as CH.
- A node whose energy is less than the average energy is eliminated as a member node (MN) that cannot be a CH.

Once the selected CHs and associated clusters are determined, the BS transmits a message containing the CH identifier (ID) to each node in the network. If the ID of the CH of the node matches its own ID and the other nodes are not selected as CH, then this node will join the nearest CH according to the CH identifier. Subsequently, nodes identified as CH are used to collect data from cluster members using Time Division Multiple Access (TDMA). The LEACH-C protocol was more effective than the LEACH protocol alone. It has a longer lifetime than the LEACH protocol. The LEACH-C protocol exhibits a better performance owing to the CH distribution across the network.

Liu and Ravishankar [14] developed an energy-efficient adaptive cluster protocol depend on GA, called the LEACH-Genetic Algorithm (LEACH-GA). It operates on the basis of the principles of the LEACH protocol. It should be noted that the original LEACH protocol requires the user to identify the anticipated probability for CHs and employs a threshold function to determine whether a node is a CH or not. However, in the LEACH-GA protocol, the GA is used to calculate and determine the optimal threshold probability of clustering in WSNs. Consequently, the optimal value of CH probability is assigned. The LEACH-GA protocol establishes a third phase in addition to the set-up and steady-state phases of the LEACH protocol. This phase was referred to as the preparation phase. In principle, the LEACH-GA protocol starts with the preparation phase, which is executed only once. During the preparation phase, similar to the procedure in the LEACH protocol, the nodes send information on their locations, IDs, and the probability of being CHs to the BS. After the messages are received by the BS, the GA is run to compute the optimal probability value for the nodes to serve as CHs and broadcasts this probability value to all the member nodes using an advertisement message for these nodes to be used in the formation of clusters in the following set-up phase. The set-up and steady-state phases of this protocol were similar to those of the LEACH protocol. The results of the study by Liu and Ravishankar [14] support the idea that the LEACH-GA protocol prolongs the lifetime of the network over the lifetime obtained when the original LEACH protocol is used, because it uses the optimal probability in the formation of CHs. Although LEACH-GA improves the CH threshold function, it does not consider the residual energy that is the selection of the CH is performed randomly.

Tabatabaei [15] proposed a novel algorithm, called the Social Spider Optimization (SSO) algorithm, for SN clustering. The algorithm is based on the use of social spider optimization and fuzzy logic to balance power consumption in WSNs. In SSO algorithms, nodes mimic spider groups and interact according to biological rules of a colony. Furthermore, the node is selected using fuzzy logic depended on two measures: the battery level and depth. This algorithm divides the energy consumption of clusters into intra- and inter-cluster electricity consumptions. In this protocol, the nodes compete to become a CH based on the lasting power and distance from the sinks. Nodes with higher power levels and shorter distances from the sinks were selected as the CHs. Nodes that were selected as CHs and those that were not selected as CHs joined the nearest CH. In addition, this protocol depends on the mobile sink because the nodes near the fixed sink share multihop routes and data and are integrated into the sink. In fact, these nodes tend to consume more battery energy than other nodes in the network. To evaluate the performance of the proposed protocol, a comparison was made between the Dynamic Command Response and Reply Protocol (DCRRP) and Novel Distributed Clustering Routing Protocol (NODIC). The comparison results revealed tha the proposed protocol has better performance in terms of power consumption than the two aforementioned protocols.

Wang *et al.* [16] presented a novel trajectory scheduling method that depends on the coverage of several mobility sinks in large-scale WSNs, abbreviated as TSCR-M. TSCR-M uses the Particle Swarm Optimization (PSO) heuristic algorithm and GA. PSO was integrated with mutation operators to search for optimal coverage-efficient parking positions. GA was used to determine the movement direction of various mobile sinks. In TSCR-M, all sensors prefer to use single-hop communications within their transmission ranges to transfer the monitored data to the mobile sink. In addition, mobile sinks travel along the loops and stop only at parking positions for data collection. The simulation results verified the efficiency of the TSCR-M method using the energy consumption and network lifetime. Furthermore, researchers in [17] employed PSO as an optimization algorithm combined with the Cuckoo Search (CS) algorithm to design Multi-Objective Optimization (MOO) algorithms for CH selection. Their results revealed that through network division, energy consumption was improved.

In [18] an improved LEACH protocol for mobile SNs was proposed to extend the life of the network and reduce packet loss through a fuzzy classification system called LEACH-mobile-fuzzy (LEACH-MF). The hierarchical clustering method combines various parameters with fuzzy logic for the CH selection. In addition to the residual energy, during CH selection, the speed of movement and break time (as movement) are introduced as vague descriptions. Therefore, the life cycle of the network is increased by balancing the energy consumption between the nodes. In addition, package loss is reduced by choosing more stable CHs with low mobility. The simulation results revealed that the LEACH-MF method improves network life and packet loss and that it is more efficient than other clustering algorithms such as the cluster-based routing protocol for mobile sensor nodes (CBR-mobile) and mobility-based clustering (MBC).

Alami *et al.* [19] proposed a novel scheme for an energy-efficient adaptive clustering algorithm for mobile WSNs, known as EEA. The proposed scheme offers a new form of node association to minimize the distance for data communication in WSNs to improve the energy efficiency of these networks. After the random deployment of SNs, cluster head selection is performed based on the difference between the SNs using a clustering method that considers their energy levels. Thus, the energy problem associated with cluster formation is solved. Furthermore, the EEA scheme binds sensor nodes to the CH in a free-adaptive association mode such as the backward mode. The transmission is eliminated and the total communication distance through which the locally collected data are traversed is minimized. Because the EFA associated mode reduces the total distance between the WSN data and messages, energy consumption is minimized and network life is maximized. Simulations were performed using 100 nodes that were randomly distributed in a network area of 100×100 m, where CCA-EGA was run on 600 nodes that provided a high diversity of nodes in the network area.

Idrissi *et al.* [20] proposed a new routing technique called optimal selection of a CH on the grid (OSCH-Gi). This method divides the network into multiple grids, where the CH of each grid is selected according to the residual energy and distance to the BS. The simulation results show that OSCH-Gi is more efficient in term of network lifetime and energy consumption than the other cluster algorithms.

Moridi *et al.* [21] proposed a fault-tolerant, clustering-based, multi-path algorithm (FTCM) that is based on a hybrid, energy-efficient, distributed clustering method called the HEED algorithm. The FTCM algorithm focuses on selecting a CH and backup node for each cluster to reduce the CH error. The backup nodes monitor the performance of CHs and store their data until they are delivered to their destination. In the FTCM, three pathways connecting sources CH and BS were identified. These routes are selected based on four parameters: the number of hops, residual energy, speed of spread, and reliability. According to the simulation results, the FTCM improved the energy consumption, reduced the end-to-end delay and packet loss rate, and increased the amount of correct data in the network. However, these researchers [21] did not discuss how to handle the connection and their method did not include a phase for routing maintenance.

Hajipour and Barati [22] proposed an energy-efficient, layered, routing protocol (EELRP) that divides the network area into eight equal sections, where every section is produced after performing a crossover between layers and sections. Each section contained a set of nodes, and the most appropriate node was likely to receive data from other nodes. Subsequently, every agent sends the aggregated information to the agent in the lower section of the same sector until the information arrives at the BS. The study results provide evidence that the EELRP protocol improves network lifetime.

Barati and Naghibi [23] investigated the possibility of reducing the power consumption by reducing the number of packets exchanged in WSNs. This task was accomplished by applying Secure Hybrid Structure Data Aggregation (SHSDA) based on tree and star topologies. In this method, the network is divided into four equal parts, each with a star structure. Each node transmits its data to a single parent. The BS then receives the data from the parent nodes using a tree structure. Barati *et al.* [24], [25] adapted artificial algorithms such as the Firefly algorithm (FA) and fuzzy logic (FL) to achieve the goals of reduced energy consumption and increased network lifetime. The method proposed in [24] consists of three phases, where FA and FL are applied in phases one and two, respectively. FA is used to cluster the WSN, whereas FL is employed to discover the paths between CHs by creating a primary path and backup path. The primary path is used to transmit data to the BS under normal conditions, whereas the backup path is used during failure of the primary path. The last phase is intended to maintain the network operation by restarting the route discovery process in the case of path failure.

GA has been widely used with WSNs to improve network lifetime [26], [27], [28], [29]. This reflects the significance of GA in the elongation of the lifetime of the network and the reduction in energy consumption. Researchers have also developed fitness functions to enhance the performance of these networks. The simulation results indicate that these functions achieved their goals. For example, in [26], the researcher applied a method called Genetic Algorithm Based Energy Efficient Clusters (GABEEC). This method proceeds in two phases: the set-up phase and the steady-state phase. Once clusters are created in the set-up phase, they remain static while the CHs change dynamically. In the steady-state phase, all nodes begin to communicate with their CHs for data aggregation. The data were then compressed and sent to the BS. After completion of the two phases, the BS evaluates the energy status of the CHs. If the CH energy is smaller than the MN's average energy, the CH is modified by selecting the MN with the highest residual energy. The GABEEC method uses binary encoding with a roulette wheel selection method with a random crossover point. Different parameters were considered to formulate a fitness function to enhance the network lifetime such as, the residual energy of the node, the distance metric between the node and BS, the probability of the node to become a CH, and the round in which the first node dies, as in [25] and [29]. In GABEEC, three parameters are considered in the fitness function: the first node dies round, the last node dies round, and the distance between clusters. The results of the evaluation showed that GABEEC was more effective than conventional LEACH and extended the network lifetime. However, the limitations of GABEEC is CH rotations, which reduce efficiency because clusters do not change during the WSN lifetime [27], [28].

Sabor *et al.* [29] applied a GA to develop a new protocol called the GA-based Energy-Efficient Adaptive Clustering Hierarchy Protocol (GAEEP). This protocol focuses on maximizing network lifetime by delaying the death of the first node. The main GAEEP processes were similar to those used in the LEACH protocol. In principle, the GAEEP protocol applies a binary encoding. The fitness function, however, is determined by several parameters, including the total energy of all alive SNs, the total dissipated energy of all alive SNs, and energy is dispersed in control packages of CH and MN. This algorithm incorporates the GA processes of selection, crossover, and mutation into GAEEP to obtain a solution that improves the stability of the network lifetime over that provided by LEACH.

III. THE GENETIC ALGORITHM PROPOSED FOR WSN CLUSTERING

Many cluster algorithms have been proposed to maximize the WSN energy efficiency. In the clustering algorithm, all nodes are organized within clusters where data transmission is initiated from every node to the BS through a CH. However, these protocols also have several limitations.

- the residual energy is not considered in CH selection.
- rotation of the CHs occurs, which negatively affects energy efficiency, where the CHs do not change throughout the lifetime of the WSN.
- In some protocols that apply GA, there is a need to apply an appropriate fitness function and crossover to maximize energy efficiency.

Considering the limitations of clustering algorithms, the CCA-EGA algorithm considers the residual energy in the selection of the CH. Furthermore, it employs a fitness function to improve the lifetime of the network. In addition, a node located far away from the BS was utilized to increase network robustness.

The algorithm proposed (CCA-EGA) in this study is a clustering algorithm based on a GA as a centralized system to utilize all BS operations. Therefore, the BS strives to optimize cluster formation and CH selection processes based on routing energy efficiently to cluster networks. Moreover, GA procedures are used to enhance the CH selection and network lifetime. This algorithm mainly takes into account the dynamic re-clustering (or clustering) to configure routes to prolng the lifetime of the network. In the CCA-EGA algorithm, operations are separated into rounds similar to the LEACH protocol. The setup phase includes CH selection and cluster construction, and data transmission is included in the steady-state phase. Although, the steady-state phase of the CCA-EGA protocol is essentially equivalent to that of LEACH, the setup phase differs from LEACH as it employs a GA-based method to perform the required computations in the BS to optimize the CH selection for each node to become a CH.

1) THE SET-UP PHASE

The set-up phase can be referred to as the cluster-building phase, which is the phase in which clusters are created and a set of CHs is specified. This phase relies on a centralized clustering method that provides the BS with computational resources and unlimited power and storage. Furthermore, a GA was applied to improve the cluster configurations in the WSN. To the best of our knowledge, all existing clustering protocols and algorithms have been designed in such a way as to progressively re-cluster the configuration processes. At the end of round r, the configuration $(r+1)$ of the round was determined without considering the configuration of the next round. This round-by-round re-clustering process creates local optimization problems. The set-up phase in the CCA-EGA protocol differs from the set-up phases in all existing clustering techniques in that the CCA-EGA protocol performs clustering configurations for the entire network during the first round of the set-up phase. This process consists of two steps: clustering and information collection. Both steps focus on selecting the optimal CH for the entire cluster. In the information collection phase, network information is only collected once. In other words, this step was performed only once before the start of the first round. Each alive SN in the network field sends information about the ID number and location (i.e., distance), and the amount of energy to the BS at the beginning of the network initialization. The clustering step then begins, and the BS uses the information received to run the central CH based on the GA and optimizes the selection of the CH depending on the design of the fitness function and the overall life time of the network. When selecting the CH, the Euclidean distance $d\,(i, j)$ is utilized when calculating the distance between nodes, and the distance between node and BS is considered as in (2):

$$d\,(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{2}$$

In this equation, variables (x_i, y_i) and (x_j, y_j) represent the cartesian coordinates of nodes i and j. GA begins by initializing a randomly selected population of possible solutions. Each population solution is represented by chromosomes containing gene sequences. Each of these genes was assigned to a CH in the WSN. The fitness function is then evaluated for each population member in every solution to determine its goodness. In every genetic iteration, the GA attempts to improve the on-hand population through selection, crossover and mutation operators. Subsequently, GA is operated to generate the next population and improve the lifetime of the network. When the BS selects the optimal CH set to identify all the cluster configurations, the first round transmits a message defining cluster configuration (ADV_CONFIGURATION message). This message contains the CH ID, each CH's associated MN, and TDMA calendar. In principle, the TDMA schedule assigned by BS or each cluster member is proportional to the cluster size. Thus, each node turns off its radio to conserve energy. It only activates (ON) to transmit data when it arrives at the TS. However, TDMA schedules organize intra-group communications, thereby reducing the interference of each group. After the clustering is completed, the TDMA schedule is constructed and distributed, and the steady-state phase begins. After the WSN's first round ends, and during the second round of the set-up phase, BS broadcasts a message ADV_CONFIGURATION containing the pre-defined configurations of the already calculated input round. However, because large and small CHs drain network energy, the percentage of ''optimal'' CHs is important and must be considered carefully.

The choice of the CHs is crucial for energy-efficient clustering. The following section explains the procedure that the GA employs to identify the best CHs sequence, which significantly affect the network lifetime.

- Genetic algorithm encoding

In WSNs, the optimization of clustering and the reduction of energy consumption using an energy-efficient routing algorithm elongates the lifetime of the network. A Real Coded Genetic Algorithm (RCGA) was used for this purpose instead of integer or binary coding. Integer coding is suitable for non-fractional numbers or values where this work depends on the real (fractional) number. In addition, binary coding requires mapping entire real values to be converted into binary code. Furthermore, real coding is appropriate in this study for the proposed crossover methods. Each solution represents the specific length of the chromosome, each chromosome being divided into several genes. A group of chromosomes is known as a population. This represents a collection of feasible solutions. In this representation, a chromosome is constructed as a special sequence of allocated CHs selected from SNs where the value of the gene within any chromosome specifies the node ID (a unique identifier of 1 to n is a number of SNs in a network.). The chromosome itself was divided into several equal parts equivalent to the maximum number of iterations (r_{max}), each part of this chromosome contained a fixed set of CHs used for a specific transmission iteration (percentage of CHs). Fig. 1 shows the structure of the chromosome that used in the proposed CCA-EGA protocol. The fixed length of each chromosome is specified as in (3):

$$Chr_{Lenght} = percentage\ of\ CHs \times r_{max} \tag{3}$$

where $r_{\max}$ is the network's maximum round count and $\mathrm{Chr}_{\mathrm{Length}}$ is the length of the chromosome. The clustering configurations are improved by allocating CHs in advance for each cycle, which also eliminates the expenses associated with periodic re-clustering. To demonstrate the representation of a solution, consider a 100 sensor network with 500 transmission cycles. Suppose that five nodes were selected as CHs for each cycle (5%). $\mathrm{Chr}_{\mathrm{Length}}$ can be calculated as 2,500 (50 x 50) genes. Each gene of a chromosome represents a CH ID in a network. as shown in Fig. 1.

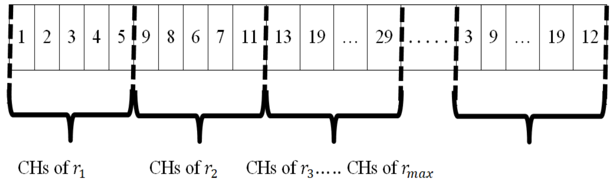


FIGURE 1. Chromosome structure.

In Fig. 1, nodes 1, 2, 3, 4, and 5 are the CHs in the first round whereas nodes 9, 8, 6, 7, and 11 are the CHs in the second round.

- Population Initialization

The populations were randomly initialized and stored in the BS. Each possible solution has a fitness value assigned to it based on the fitness function calculations and is stored in another vector matrix. In the beginning, during population (Pop) initialization, the elementary, or starting, solutions have to be smoothly pre-defined to form the population that is mostly produced in a random manner, which consists of randomly generated CH chromosomes. The size of the initial population (i.e., the number of chromosomes) is important for GA and is a pre-defined value. Population size is a critical variable based on the nature of the problem under consideration. A large $\mathrm{Pop}_{\mathrm{size}}$ will help to cover all points in the search space, but this will lead to slow convergence. Conversely, very small values of $\mathrm{Pop}_{\mathrm{size}}$ result in fast convergence to non-optimal solutions (premature convergence). To give all variables a chance to be included, it is recommended that the initial population be sufficiently large, which is determined in this study to be 50 possible solutions.

- Fitness function

In a WSN, clustering and energy-efficient routing are nondeterministic, polynomial-time hardness (NP-Hard) problems. As a result, CH selection for energy-efficient routing is important for prolonging the life of the network and must focus. Thus, the new fitness function $f(x)$ plays a significant role in increasing the useful life of a network. This is required to evaluate the quality of member solutions (organized CHs). The more appropriate the solution, the greater is the fitness value. Furthermore, an elitism strategy contributes to the preservation of the historically obtained elite chromosome, which is expected to enhance the performance of the GA in dynamic clustering. However, emphasis on transmission distance is crucial because it has an important impact on energy consumption, network life and data routing reliability.

The fitness function applied in this study was formulated as in (4):

$$N_{i,j} = \begin{cases} 0, & \textit{dead node} \\ 1, & \textit{alive node} \end{cases} \tag{4}$$

First, variable $N_{\mathrm{i,j}}$ is used as an indicator of whether node i in round j is alive or dead:

At a specific round as in (5):

$$g(\boldsymbol{x}, j) = \sum_{i=1}^{n} N_{i,j} \tag{5}$$

where $g(i)$ represents the number of alive nodes in specific (j) rounds, n is the number of nodes and i is the index of the node.

The GA then maximizes the number of transmissions while having a fraction of alive nodes. Each GA solution is generated to maximize the number of rounds, j, when a specific number of alive nodes is recorded as shown in (6).

$$\mathrm{f(x) = max(j)}\ \forall\ \mathrm{g(j)} >= \text{alive fraction}^{*}\mathrm{n} \tag{6}$$

Lastly, the fitness function is scaled as follows:

Scaling fitness function = 1/ (Rank(i))0.5

In the CCA-EGA protocol, users specify the wanted probability of CH, and use a threshold functions to determine whether node become CH. The threshold was specified as the percentage of alive nodes required before the transmission was stopped.

- Selection

The selection is done to open the way to subsequent pairing or crossover the selected chromosomes (CHs used as parent chromosomes) to extract the current population. Basically, the more suitable the chromosomes are, the greater their probability of transmission to the next generation.

In the algorithm proposed in the present study (CCA-EGA), the elitism and roulette wheel methods are used to evaluate chromosomes to select those that can produce the next generation. First, the elitism method was used to select 2% of the chromosomes with the best-fit fitness value as elites for propagation to the next generation. It is necessary to protect a good solution from being lost so that it will be incorporated into subsequent generations. Here, the selection of elitism contributes to tracking the best solution obtained in a given generation. In addition, we used the roulette wheel method to generate other chromosomes for the next generation. However, the roulette-wheel strategy causes a premature convergence problem of the GA, which sometimes results in an incorrect optimal solution [30]. Therefore, we employed a fitness scaling method to improve the performance of roulette wheel selection in the GA. In the classical roulette wheel method, the most fitting chromosome occupies the largest section of the wheel, further increasing the probability of an unfitting chromosome (i.e., least fitting) [30]. To address this problem, many fitness scaling methods have been developed, including linear, power, and exponential scaling [31]. The scaling method adopted in the present study preserves the population diversity by ensuring small differences between members with small and large fitness value. The main objective of this method is to avoid destroying unfitting chromosomes early in the process, which may have reasonably good characteristics.

- Crossover

Once the chromosomes selection process is completed, it is based on the use of roulette wheels and elite. The selected chromosomes are ready for crossover. In the GA, the decision to complete crossover process depends on the probability of crossover (P_{x}) which usually lies between 0.5 and 1.0 [32]. In this study, P_{x} equals 0.8 which indicates a high possibility of crossover. However, the crossover process ensures that genetic materials are exchanged between two members of the population called the ''parent chromosomes'' (Pr_{A} and Pr_{B}), and produces chromosomes that are probably better than the parents chromosomes, inheriting the parents' best characteristics and producing new children (new solutions). However, the selection of appropriate crossing operators should avoid early GA convergence as far as possible.

In this study, two crossover operators were used based on real parameters. These are Whole Arithmetic Crossover (WOX), and Local Crossover (LX) operators. These two crossover operators were compared with three crossover operators: the Simple Arithmetic Crossover (SMX) operator, Single Arithmetic Crossover (SNX), and Discrete Crossover (DX) operator which were used in [33] and [34]. These operators determine distinct parameter values to represent 0 – 1 random values, as described in [9]. However, these operators are not used with a real-parameter GA.

WOX is the most common operator that works by taking the weighted sum with the same $\alpha\alpha$ of the two parental gene values for every gene [9] to produce a new child as in (7):

$$Ch_i^{gene} = \alpha \times Pr_A^{gene} + (1 - \alpha) \times Pr_B^{gene} \tag{7}$$

The LX operator uses the same equation as the WOX operator, except that the WOX operator selects the value of α at random for every gene position [9].

- Mutation

A mutation process was implemented to ensure the population genetic diversity by introducing random changes to create new individuals. In general, mutation is carried out in sequence after crossover, depending on the mutation probability (P_m). This probability represents the percentage of genes with mutations. In general, GA has a P_m ranging from 0.001 to 0.05 [32]. High mutation rates interfere with evolutionary processes, while low mutation rates do not produce good changes.; thus, the value of P_m was determined to be 0.05. In the proposed CCA-EGA protocol, mutations are carried out based on P_m by randomly changing the gene values $[1, n]$, where n is the ID of the network node.

- Termination process

Basically, when the GA reaches one or more termination criteria, the algorithm stops the cycle and gives the best chromosomes in the current population. In the current study, the process was terminated when the maximum generation number (max_{gen}) was reached or when a generation reached a pre-defined value (r_{max}).

2) THE STEADY-STATE PHASE

The Steady- state phase typically begins after the setup phase is completed. In the current study, the steady-state phase was identical to that used in the LEACH protocol. It was started once the clusters had been organized and the TDMA schedule was fixed. Data transmission is controlled when each node transmits the data sensing to the CH corresponding to the TDMA plan received by the BS. After the CH receives all the required data, each CH processes the aggregation of the data received to discard redundant and uncorrelated data and transmit a single-hop communication of consolidated or fused data to the BS.

IV. EXPERIMENTAL RESULTS AND EVALUATION

A. SIMULATION SOFTWARE (MATLAB)

All GA-based cluster operations and simulations are performed using MATLAB software installed on a laptop with the following specifications: Intel Core i5 CPU (2.2GHz), 4.0 GB RAM and Windows 7 64-bit operating system. MATLAB was used because it has several toolboxes for modeling and simulation, data analysis and processing, and algorithm development.

B. ENERGY CONSUMPTION MODEL

Normally, radio models are composed of three main components: transmitter, power amplifier and the receiver. In this study, we used a simplified energy model (known as the first radio model) introduced by LEACH [9] to estimate the energy consumption in WSNs. In this model, both the free space (for d^2 power loss) and multi-path fading (for d^4 power loss) channels were relayed along the path extending from the transmitter node to the receiver node to route data transmission. Therefore, when the distance is less than a threshold value that is calculated using the equation: $d_0 = \text{Sqrt}(fs/mp)$, ε_{fs} then the free space (fsfs) is used, otherwise, the multi-path (mp) model is used when the distance is greater than the threshold value. The energy is distributed at the distance d between the transmitter (ETX) and the receiver (ERX) and transmits the l-bit data packet d can be mathematically calculated as in (8) and (9) [9]:

$$\boldsymbol{E_{TX}\,(l,d)} = \left\{\begin{matrix} \boldsymbol{l \times E_{elec} + l \times \varepsilon_{fs} \times d^2, if\ d \le d_0} \\ \hline \boldsymbol{l \times E_{elec} + l \times \varepsilon_{mp} \times d^4, if d > d_0} \end{matrix}\right\} \tag{8}$$

$$\boldsymbol{E_{RX}\,(l) = l \times E_{elec}} \tag{9}$$

where

d: the distance between the transmitter and the receiver.

d_0: the transmission distance threshold.

fs: the amplification coefficient of the free space signal.

mp: the multi-path fading signal amplification coefficient.

E_{elec}: the energy consumed to transmit or receive data packets of length l in bits. However, E_{elec} can be affected by digital coding, modulation, filtering, and signal scattering.

C. NETWORK MODEL

To test the proposed CCA-EGA, a set of network assumptions was first made. These are:

1. There is only one BS at a fixed location that is immobilized and has an unlimited memory, calculations, and battery resources.
2. The BS is located away from the sensor field.
3. All SNs are homogeneous, have known locations, have the same capabilities and characteristics, and have limited initial energy resources (equivalent amounts of energy). In addition, they are not rechargeable. Moreover, they always have data to be transmitted.
4. Each SN has a unique identifier (node ID).
5. Single-hop communication is applied whereby the MN transmits its sensory data to the CH, which, in turn, transmits these data to the BS at a very far distance.
6. The radio models of the transmitters and receivers for the energy consumption calculations are similar to those presented in [9].

D. PERFORMANCE MEASUREMENT

To evaluate the proposed CCA-EGA protocol, we compared the performance of the known LEACH with five metrics.

1) The Network lifetime

There is no universal agreement on the definition of the network lifetime [34]. However, it can be defined based on the network requirements. This can be described as [34]:

- The time when the first node dies (FND).
- The time when the half node dies (HND).
- The time when the last node dies (LND).
- Time when a certain fraction (percentage) of the nodes died (e.g., 10%, 60%, or 90%).

2. Total number of alive nodes per round. This metric is related to the network lifetime. This determined the number of alive nodes in each round.
3. Total number of dead nodes per round. This metric is also related to the network lifetime. The number of dead nodes in each round was measured.
4. Total residual energy per round. This metric computes the average energy remaining in all the nodes during the rounds.
5. The throughput is the total number of packets in the BS. This metric counts the total number of packets received by a BS.

E. THE SIMULATION SETTINGS

The proposed CCA-EGA simulation is based on the random distribution of 100 random SNs over 2D representations of network areas (i.e., sensor field) of 100 m * 100 m with a far BS located at the coordinates of 100 and 375 m. In addition, in the simulation, the SN is a homogeneous node. The parameters used in the simulation and their values are listed in Table 1.

TABLE 1. Simulation parameters and their values.

Simulation Parameter	Value(s)	Simulation Parameter	Value(s)
Network Size ($n\times n$)	100 m x 100 m	Energy to run radio electronics circuit (E_{elec})	50 nJ/bit
BS Location (X, Y)	(100 m, 375 m)	Threshold distance (d0)	87 m
Number of SNs	50, 100	Mutation probability	0.05
Percentage of CHs	5%	Crossover probability	0.8
Packet Size	4,000 bits	Primary population size	50
Control packet size	200 bits	Generation size	50
Maximum number of rounds	600	Elite	2
Initial energy (E_0) of SNs	0.5, 0.4 Joule/node	Selection operators	Roulette Wheel based fitness scaling, and elitism
Data aggregation energy (E_{DA})	5 nJ/bit/signal	Crossover operator	WOX, SNX, SMX,
			LX, and DX
Amplification energy for multi-path (ε_{mp})	0.0013 pJ/bit/m^4	Mutation operators	Uniform
Amplification energy for free space (ε_{fs})	10 pJ/bit/m^2		

TABLE 2. Comparison of network lifetimes without fitness scaling (E_0 = 0.40).

Crossover Operator	Number of Rounds until Death of Nodes							
	FND	10 %	20 %	30 %	50 %	70 %	80 %	90 %
SNX	30	60	78	91	124	170	203	310
WOX	27	47	66	80	113	162	206	371
SMX	24	59	78	92	118	153	173	211
LX	26	54	72	86	119	161	201	278
DX	29	58	73	88	114	147	174	213

F. SIMULATION RESULTS AND EVALUATION OF THE PROPOSED GA-BASED CLUSTERING ALGORITHM

The following paragraph deals with the simulation results of the proposed CCA-EGA algorithms by analyzing the lifetime of the network, total number of alive nodes per round, and total number of dead nodes per round. The results were compared with those obtained using the LEACH protocol. The results below are the average of 10 independent runs of 600 rounds per run.

1) COMPARISON OF NETWORK LIFETIMES

The proposed GA-based clustering algorithm (CCA-EGA) was primarily designed as a new method to provide an energy-efficient algorithm for maximizing WSN network lifetime. Intrinsically, GA operators, selection, crossover, and mutation are central to achieve the best performance. Therefore, we investigated the performance of five crossover operators: WOX, SNX, SMX, LX, and DX, and implemented them based on the two selection operators of elitism and roulette wheels. Simulations were conducted to evaluate the performance of each crossover operator with and without fitness scaling for various E_0 values. The results of these simulations are listed in Tables 3 to 5. The network life metrics considered were (i) FND and (ii) a certain percentage of dead nodes (10% - 90%).

Tables 3 to 4 show that there are differences between the various examined GA crossover operators when using elitism selection and roulette wheel selection with and without fitness scaling. Regarding network lifetime, the results

TABLE 3. Comparison of network lifetimes with fitness scaling (E_0 = 0.40).

Crossover Operator	Number of Rounds until Death of Nodes							
	FND	10 %	20 %	30 %	50 %	70 %	80 %	90 %
SNX	26	50	66	82	117	180	251	478
WOX	24	57	77	96	125	177	227	433
SMX	26	54	71	87	115	159	191	286
LX	25	56	73	90	120	167	203	332
DX	27	54	73	90	117	153	175	208

TABLE 4. Comparison of network lifetimes without fitness scaling (E_0 = 0.50).

Crossover Operator	Number of Rounds until Death of Nodes							
	FND	10%	20%	30%	50%	70%	80%	90%
SNX	38	71	92	111	148	194	234	310
WOX	39	71	93	111	141	181	213	289
SMX	45	77	100	118	150	191	225	282
LX	38	72	91	107	147	192	229	326
DX	38	71	93	110	141	181	209	254

TABLE 5. Comparison of network lifetimes with fitness scaling (E_0 = 0.50).

Crossover Operator	Number of Rounds until Death of Nodes							
	FND	10 %	20 %	30 %	50 %	70 %	80 %	90 %
SNX	36	62	87	106	150	217	297	496
WOX	36	71	92	108	147	200	247	439
SMX	43	77	96	113	146	187	227	298
LX	33	71	91	109	147	196	246	368
DX	41	78	101	122	158	196	220	264

summarized in Tables 2 and 4 demonstrate the impact of E_0 without fitness scaling when the value of E_0 is changed from 0.4. to 0.50. The results reflect an improvement in the FND for all crossover methods when E_0 was increased to 0.50. For example, the FND using SMX improved from 24 to 45 rounds, suggesting an approximately double lifetime improvement. Similar to SNX, the other methods improved the FND. This leads to the conclusion that charging nodes with a suitable initial amount of energy increases their lifetime. However, with fitness scaling, all crossover methods resulted in an improvement in the FND and a decrease in the percentage of nodes that died. The best results were obtained with fitness scaling and an E_0 value of 0.50. In the SNX and WOX cases, node deaths occurred during rounds 496 and 439, respectively. However, for the LX operator, Tables 4 and 5 show that 70%, 80%, and 90% of the nodes died during rounds 192, 229, and 326, respectively. However, with fitness scaling, these death percentages were reached in rounds 196, 246, and 368.

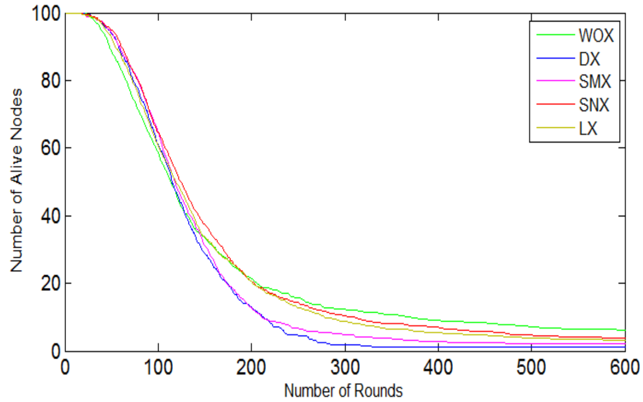


FIGURE 2. Comparison of alive nodes in each round without fitness scaling (E_0= 0.40).

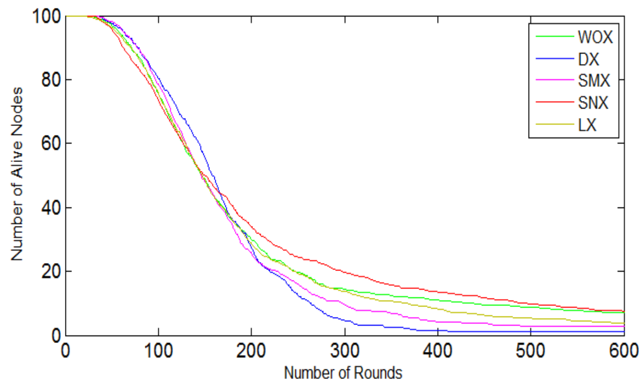


FIGURE 3. Comparison of alive nodes in each round with fitness scaling (E_0= 0.40).

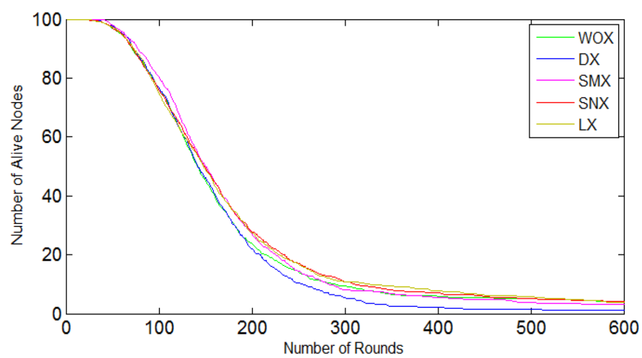


FIGURE 4. Comparison of alive nodes in each round without fitness scaling (E_0= 0.50).

In summary, the results reported in Tables 3 to 5 indicate that the SNX, WOX, and LX crossover methods produced better results than the SMS and DX methods.

2) COMPARISON OF ALIVE AND DEAD NODES

Figures 2 to 5 present a comparison of the numbers of alive nodes in each round between the case when fitness scaling was applied and the case when it was not, taking into account the residual energy when E_0 had a value of 0.40 and when it had a value of 0.50. These figures confirm that the proposed crossover methods last for the lifetimes of multiple nodes for

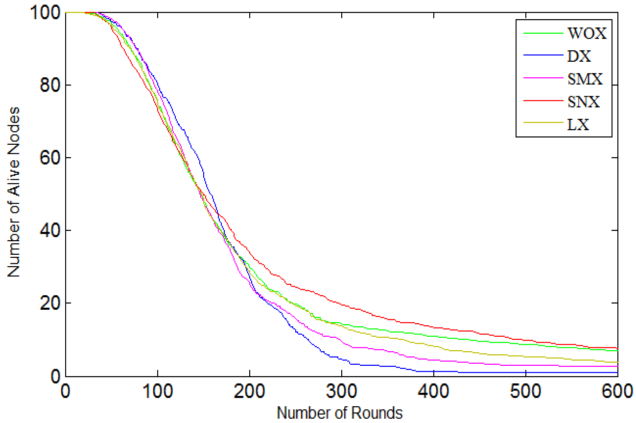


FIGURE 5. Comparison of alive nodes in each round with fitness scaling ($E_0 = 0.50$).

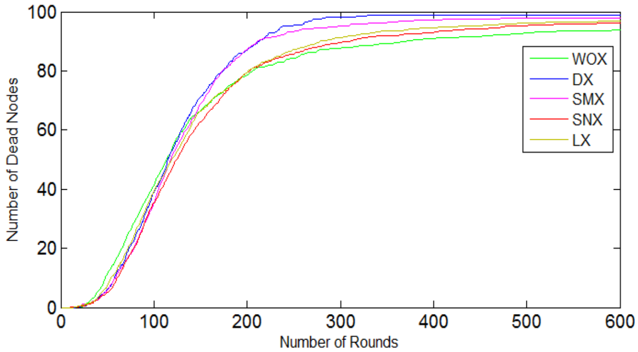


FIGURE 6. Comparison of dead nodes in each round without fitness scaling ($E_0 = 0.40$).

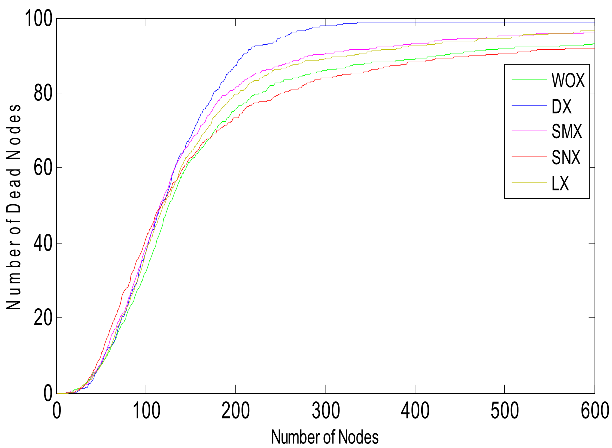


FIGURE 7. Comparison of dead nodes in each round with fitness scaling ($E_0 = 0.40$).

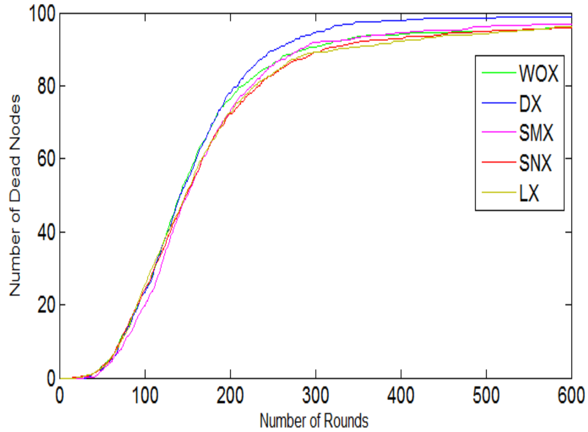


FIGURE 8. Comparison of dead nodes in each round without fitness scaling ($E_0 = 0.50$).

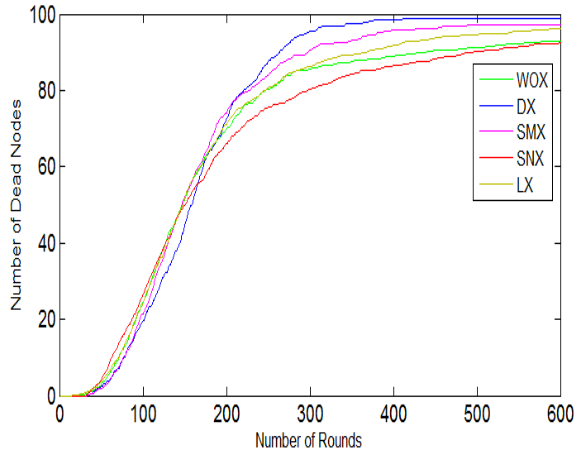


FIGURE 9. Comparison of dead nodes in each round with fitness scaling ($E_0 = 0.50$).

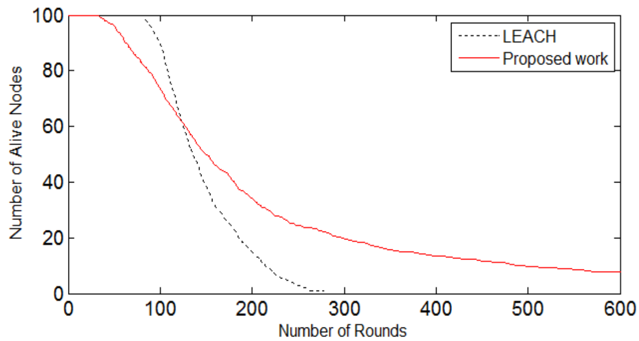


FIGURE 10. Comparison of alive nodes in each round.

more than 600 rounds, where WOX and SNX outperform the other crossover methods under fitness scaling with an E_0 value of 0.50. In addition, the WOX and SNX crossover methods keep the nodes alive longer than the SMX, DX, and LX methods.

Figures 6 through 9 present comparisons of the incidences of node deaths in each round between the case when fitness scaling was applied and the case when it was not under the condition of varying E_0 values. It can be observed that the SNX and WOX methods have fewer dead nodes during rounds with and without fitness scaling under varying E_0 values. In contrast, the DX crossover method suffers from a faster node death than the other crossover methods under study.

3) COMPARISON OF ALIVE AND DEAD NODES IN LEACH

A comparison of the performance of the CCA-EGA algorithm in terms of the number of alive SNs for different

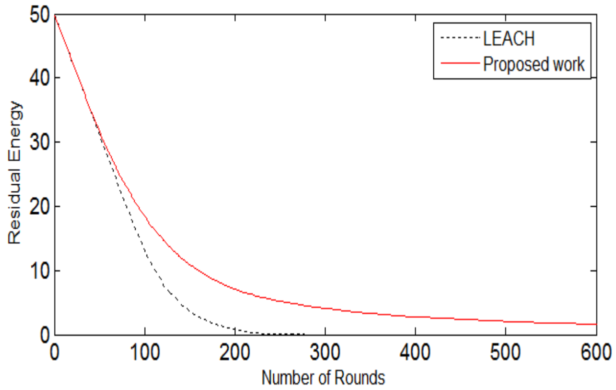


FIGURE 11. Total residual energy per round.

numbers of rounds is shown in Fig. 10. It can be observed in this figure that, in the protocol proposed here, the number of alive SNs is higher than the corresponding number in the LEACH protocol. This confirms that in the CCA-EGA algorithm, SNs remain alive in the network for a long period and that the probability of early death of SNs is reduced. The lifetime of the network is in reverse proportion to the number of SNs that died. Consequently, the number of dead SNs must be minimized to maximize the network lifetime.

4) COMPARISON OF TOTAL RESIDUAL ENERGIES

Fig. 11 shows a plot of the residual energy of the WSN per round for various numbers of rounds. Based on this plot, it can be concluded that the residual energy gradually increased in the proposed CCA-EGA clustering algorithm as the SNs died. This energy increase is owing to the successful selection of the objective function, which mainly considers the transmission distance that affects the residual energy of the SNs. However, the network still had some residual energy, even after 600 rounds.

Table 6 summarizes the results of a comparative analysis of the network lifetimes associated with the five different crossover methods that were applied to WSNs in terms of energy efficiency based on simulations using 100 randomly deployed SNs. This table lists the network lifetimes obtained using FND, HND, and LND performance measures. The proposed method proved its efficiency by delay the death of the last node. As seen in Table 6, the last node died before completing 600 rounds as in [38] and [39] whereas the proposed method prevents the death of the last node before 600 rounds. Compared to LEACH, LEACH suffers from the death of the last node earlier than the proposed method.

For further performance evaluation, the proposed clustering protocol was executed for 2000 rounds in response to the fact that the time required for the last node to die (LND) can be reached after 2000 rounds. For comparison, the artificial neural network (ANN) was used to extend the network lifetime [35]. Our results indicated that LND occurred in round 894. By implementing the GA in OGA [36], the last node died in round 1561 whereas in GABEEC [26], the death of the last node occurred in round 1571. However, the last node died in round 2280 of LEACH-T [37]. As a result, CCA-EGA achieved remarkable achievements in that it improved the WSN lifetime.

TABLE 6. Comparison of network lifetimes associated with various methods.

	Method result			LEACH		
Method	**FND**	**HND**	**LND**	**FND**	**HND**	**LND**
Proposed work	26	117	> 600	62	112	213
Fathi and Nazari [38]	22	118	547	53	97	251
Kim et al. [39]	557	-	598	245	-	590

The comparisons presented in Table 6 indicate what follows:

- FND occurred in round 62 in LEACH whereas HND and LND occurred in rounds 112 and 213, respectively. In contrast, CCA-EGA performed better in terms of HND and LND which occurred in rounds 117 and more than 600, respectively.
- In [38], FND occurred in round 53 with LEACH, whereas HND and LND occurred in rounds 97 and 251, respectively. However, in the protocol proposed by Fathi and Nazar [38], FND, HND, and LND occurred in rounds 22, 118, and 574, respectively. Therefore, the algorithm proposed by Fathi and Nazar [38] performed better in terms of HND and LND.
- Kim *et al.* [39] reported that FND and LND occurred in rounds 245 and 590, respectively, in the case of LEACH, whereas in the case of their proposed algorithm, FND and LND occurred in rounds 557 and 598, respectively, indicating that the proposed algorithm performs better than LEACH, especially in large networks.

V. CONCLUSION

The main objective of this study is to find ways to extend the lifetime of WSN. For this purpose, we developed CCA-EGA for energy efficiency in WSN. To handle the clustering and CH selection problems, some decisions were made for the eventual purpose of elongating the network lifetime. The major decisions are summarized as follows:

1- Real encoding is implemented to represent a possible solution to the existing problem of designing an energy-efficient routing algorithm. Real encoding improves the flexibility of the algorithm over binary encoding in representing real-world problems.

2- A new scaling fitness function was developed to avoid premature convergence and strengthen the global search.

3- In order to optimize the network life, we used two crossover operators and analyzed their performance to determine their impact on the network lifetime.

The simulation results show that real encoding, crossover operators, and scaling fitness functions affect the lifetime of the network. Using a GA with suitable tuning for all its operations can produce better results than traditional GA operations. Additionally, no specific crossover operator can be applied to any problem.

In view of the study's findings, several issues must be considered in future studies to address the limitations of the present study. First, this study did not discuss how to deal with a failed link between the BS and CH, which affects the stability of the network. Therefore, it is necessary to specify a maintenance policy during transmission between the BS and CH to reduce the risk of disconnection between stations. In addition, the values of the mutation and crossover probabilities were fixed during the simulation. Therefore, it is necessary to study the adaptive crossover and mutation probabilities to improve the performance of WSN in terms of energy efficiency. Finally, a CCA-EGA protocol is proposed for wireless immovable sensors. Future research could extend this approach to address mobile sensors and the optimal number of CHs. A mathematical model that represents clustering algorithms is proposed.

REFERENCES

[1] R. Lara, D. Benítez, A. Caamaño, M. Zennaro, and J. L. Rojo-Álvarez, "On real-time performance evaluation of volcano-monitoring systems with wireless sensor networks," *IEEE Sensors J.*, vol. 15, no. 6, pp. 3514–3523, Jun. 2015, doi: 10.1109/JSEN.2015.2393713.

[2] K. Jangam, V. Sunkara, M. Talluri, and Z. Shaik, "Wireless sensor network design for tactical military applications," *Int. J. Appl. Sci., Eng. Manag.*, vol. 5, no. 2, pp. 5–8, 2016.

[3] H. Elayan, R. M. Shubair, and A. Kiourti, "Wireless sensors for medical applications: Current status and future challenges," in *Proc. 11th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2017, pp. 2478–2482, doi: 10.23919/EuCAP.2017.7928405.

[4] R. Sandhya and D. N. Sengottaiyan, "A comparative study on sensor based intelligent rush-hour traffic mechanism," *Int. J. Latest Trends Engineer Technol.*, vol. 5, no. 3, pp. 216–222, Aug. 23, 2022.

[5] R. Banerjee and C. K. Bhattacharyya, "Cluster based routing algorithm with evenly load distribution for large scale networks," in *Proc. Int. Conf. Comput. Commun. Informat.*, Jan. 2014, pp. 1–6, doi: 10.1109/ICCCI.2014.6921779.

[6] D. Kumar and R. Sharma, "Energy proficient and security protocol for WSN: AODV," *Int. J. Comput. Appl.*, vol. 107, no. 13, pp. 27–33, Dec. 2014, doi: 10.5120/18813-0391.

[7] M. Abinaya and M. A. Asha, "Energy efficient routing protocol for homogeneous and heterogeneous wireless sensor networks," *Int. J. Future Innov. Sci. Eng. Res. (IJFISER)*, vol. 20, no. 20, pp. 45–50, 2015.

[8] N. Sharma and V. Gupta, "Meta-heuristic based optimization of WSNs energy and lifetime—A survey," in *Proc. 10th Int. Conf. Cloud Comput., Data Sci. Eng.*, Jan. 2020, pp. 369–374.

[9] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, Maui, HI, USA, Jan. 2000, p. 10, doi: 10.1109/HICSS.2000.926982.

[10] S. Bansal and V. K. Arora, "Distributed and centralized clustering approaches: Leach vs. C-leach," *HCTL Open Int. J. Technol. Innov. Res. (IJTIR)*, vol. 14, pp. 1–7, Apr. 2015.

[11] P. R. Mahapatra and K. R. Yadav, "Descendant of LEACH based routing protocols in wireless sensor networks," *Proc. Comput. Sci.*, vol. 57, pp. 1005–1014, Jan. 2015, doi: 10.1016/j.procs.2015.07.505.

[12] A. Kaur and G. Mehta, "Investigating optimization techniques for cluster head election in WSN," *Commun. Appl. Electron.*, vol. 5, no. 2, pp. 28–37, May 2016, doi: 10.5120/cae2016652226.

[13] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002, doi: 10.1109/TWC.2002.804190.

[14] J.-L. Liu and C. V. Ravishankar, "LEACH-GA: Genetic algorithm-basedenergy-efficient adaptive clustering protocolfor wireless sensor networks," *Int. J. Mach. Learn. Comput.*, vol. 1, pp. 79–85, Apr. 2011, doi: 10.7763/IJMLC.2011.V1.12.

[15] S. Tabatabaei, "A novel fault tolerance energy-aware clustering method via social spider optimization (SSO) and fuzzy logic and mobile sink in wireless sensor networks (WSNs)," *Comput. Syst. Sci. Eng.*, vol. 35, no. 6, pp. 477–494, 2020, doi: 10.32604/csse.2020.35.477.

[16] J. Wang, Y. Gao, C. Zhou, R. S. Sherratt, and L. Wang, "Optimal coverage multi-path scheduling scheme with multiple mobile sinks for WSNs," *Comput., Mater. Continua*, vol. 62, no. 2, pp. 695–711, 2020, doi: 10.32604/cmc.2020.08674.

[17] K. Vijayalakshmi and P. Anandan, "Global Levy flight of cuckoo search with particle swarm optimization for effective cluster head selection in wireless sensor network," *Intell. Automat. Soft Comput.*, vol. 26, no. 2, pp. 303–311, 2020, doi: 10.31209/2020.100000165.

[18] J.-S. Lee and C.-L. Teng, "An enhanced hierarchical clustering approach for mobile sensor networks using fuzzy inference systems," *IEEE Internet Things J.*, vol. 4, no. 4, pp. 1095–1103, Aug. 2017, doi: 10.1109/JIOT.2017.2711248.

[19] H. El Alami and A. Najid, "EEA: Clustering algorithm for energy-efficient adaptive in wireless sensor networks," *Int. J. Wireless Netw. Broadband Technol.*, vol. 7, no. 2, pp. 19–37, 2018, doi: 10.4018/IJWNBT.2018070102.

[20] N. El Idrissi, A. Najid, and H. El Alami, "New routing technique to enhance energy efficiency and maximize lifetime of the network in WSNs," *Int. J. Wireless Netw. Broadband Technol.*, vol. 9, no. 2, pp. 81–93, Jul. 2020, doi: 10.4018/IJWNBT.2020070105.

[21] E. Moridi, M. Haghparast, M. Hosseinzadeh, and S. J. Jassbi, "Novel fault-tolerant clustering-based multipath algorithm (FTCM) for wireless sensor networks," *Telecommun. Syst.*, vol. 74, no. 4, pp. 411–424, Aug. 2020, doi: 10.1007/s11235-020-00663-z.

[22] Z. Hajipour and H. Barati, "EELRP: Energy efficient layered routing protocol in wireless sensor networks," *Computing*, vol. 103, no. 12, pp. 2789–2809, Dec. 2021, doi: 10.1007/s00607-021-00996-w.

[23] M. Naghibi and H. Barati, "SHSDA: Secure hybrid structure data aggregation method in wireless sensor networks," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 12, pp. 10769–10788, Dec. 2021, doi: 10.1007/s12652-020-02751-z.

[24] A. N. Shahbaz, H. Barati, and A. Barati, "Multipath routing through the firefly algorithm and fuzzy logic in wireless sensor networks," *Peer-Peer Netw. Appl.*, vol. 14, no. 2, pp. 541–558, Mar. 2021, doi: 10.1007/s12083-020-01004-2.

[25] A. Singh, K. Babbar, and A. Malik, "Clustering and energy efficient routing protocol for wireless sensor network using genetic algorithm," *Int. J. Comput. Appl.*, vol. 119, no. 7, pp. 40–43, Jun. 2015, doi: 10.5120/21083-3766.

[26] S. Bayrakli and S. Z. Erdogan, "Genetic algorithm based energy efficient clusters (GABEEC) in wireless sensor networks," *Proc. Comput. Sci.*, vol. 10, pp. 247–254, Jan. 2012, doi: 10.1016/j.procs.2012.06.034.

[27] S. Das, S. Barani, S. Wagh, and S. Sonavane, "An exhaustive survey on nature inspired metaheuristic algorithm for energy optimization in wireless sensor network," *ICTACT J. Commun. Technol.*, vol. 6, pp. 1173–1181, Dec. 2015, doi: 10.21917/ijct.2015.0172.

[28] N. B. Nimbalkar and S. S. Das, "Genetic algorithm based energy aware clustering techniques: A survey," *Int. J. Comput. Appl.*, vol. 95, no. 10, pp. 11–16, Jun. 2014, doi: 10.5120/16629-6489.

[29] M. Abo-Zahhad, S. M. Ahmed, N. Sabor, and S. Sasaki, "A new energy-efficient adaptive clustering protocol based on genetic algorithm for improving the lifetime and the stable period of wireless sensor networks," *Int. J. Energy, Inf. Commun.*, vol. 5, no. 3, pp. 47–72, Jun. 2014, doi: 10.14257/ijeic.2014.5.3.05.

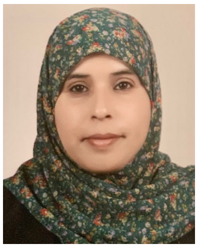
[30] B. Kishan, "Performance analysis of fitness scaling functions in genetic algorithm," *Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET)*, vol. 4, no. 7, pp. 653–656, 2016.

[31] S. Picek, D. Jakobovic, and M. Golub, "On the recombination operator in the real-coded genetic algorithms," in *Proc. IEEE Congr. Evol. Comput.*, Jun. 2013, pp. 3103–3110, doi: 10.1109/CEC.2013.6557948.

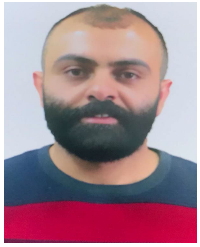
[32] D. K. Dey, "Mathematical study of adaptive genetic algorithm (AGA) with mutation and crossover probabilities," *COMPUSOFT Int. J. Adv. Comput. Technol.*, vol. 3, no. 5, pp. 765–768, 2014.

[33] A. Al-Shdaifat, W. Alsharafat, and K. Batiha, "Improving the energy of wireless sensor network using genetic algorithm," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 22, no. 1, pp. 684–696, Jan. 2022.

[34] A. Al-Shdaifat, W. Alsharafat, and K. Batiha, "Improving the energy of wireless sensor network using genetic algorithm," M.S. thesis, Dept. Comput. Sci., Al Al-bayt Univ., Mafraq, Jordan, 2018.
[35] S. Sharma, D. Sethi, and P. Bhattacharya, "Artificial neural network based cluster head selection in wireless sensor network," *Int. J. Comput. Appl.*, vol. 119, no. 4, pp. 34–41, Jun. 2015, doi: 10.5120/21058-3710.
[36] 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, Aug. 2015, doi: 10.14257/ijfgcn.2015.8.4.13.
[37] A. Yadav and S. Kumar, "A teaching learning based optimization algorithm for cluster head selection in wireless sensor networks," *Int. J. Future Gener. Commun. Netw.*, vol. 10, no. 1, pp. 111–122, Jan. 2017, doi: 10.14257/ijfgcn.2017.10.1.10.
[38] M. Nazari and M. Fathi, "An energy-efficient density-based clustering approach for wireless sensor networks," *Int. J. Mechtron., Electr., Comput. Technol. (IJMEC)*, vol. 7, pp. 3405–3410, Aug. 2017.
[39] J. Kim, Y. Kim, and Y. Yoo, "Energy-efficient routing using genetic algorithm in cluster-based wireless sensor networks," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2019, pp. 1–5, doi: 10.1109/VTCFall.2019.8891489.



WAFA ALSHARAFAT received the B.S. degree in computer science from Hashemite University, Zarqa, Jordan, in 2001, and the M.S. and Ph.D. degrees in computer information system from Arab Academy, Amman, Jordan, in 2004 and 2009, respectively. She is currently an Associate Professor with Al Albayt University, Mafraq, Jordan. Her research interests include the wireless sensor networks, artificial intelligence, genetic algorithm, optimization algorithms, and network security.



ALAA AL-SHDAIFAT received the B.S. and M.S. degrees in computer science from Al Albayt University, Mafraq, Jordan, in 2012 and 2017, respectively.

He is currently working with international non-profitable organization. His research interests include the wireless sensor networks and artificial intelligence.

KHALED BATIHA is currently a Professor with Al Albayt University, Mafraq, Jordan. His research interests include the wireless sensor networks, network security, and data mining.

AKRAM ALSUKKER is currently an Assistant Professor of industrial engineering with The Hashemite University, Zarqa, Jordan. His research interests include the artificial intelligence, neural networks, genetic algorithm, particle swarm optimization, ant colony optimization, differential evolution, wireless sensor networks, and optimization.