Wireless Sensor Network Optimization Using Genetic Algorithm

Aseel B. Alnajjar 1*, Azhar M. Kadim 1, Ruaa Abdullah Jaber 1, Najwan Abed Hasan 1, Elhsan Qahtan Ahmed 1, Mohammed Sahib Mahdi Altaei 1, Ahmed L. Khalaf 2

1 Computer Science Department, College of Science, Al Najran University, Jaddiya, Baghdad, Iraq
2 Department of Communication Technology Engineering, College of Information Technology, Imam Ja’afar Al-Sadiq University, Baghdad, Iraq
Email: 1 aseel.alnajjar@nahrainuniv.edu.iq
*Corresponding Author

Abstract—Wireless Sensor Network (WSN) is a high potential technology used in many fields (agriculture, earth, environmental monitoring, resources union, health, security, military, and transport, IoT technology). The band width of each cluster head is specific, thus, the number of sensors connected to each cluster head is restricted to a maximum limit and exceeding it will weaken the connection service between each sensor and its corresponding cluster head. This will achieve the research objective which refers to reaching the state where the proposed system energy is stable and not consuming further more cost. The main challenge is how to distribute the cluster heads regularly on a specified area, that’s why a solution was supposed in this research implies finding the best distribution of the cluster heads using a genetic algorithm. Where using an optimization algorithm, keeping in mind the cluster heads positions restrictions, is an important scientific contribution in the research field of interest. The novel idea in this paper is the crossover of two-dimensional integer encoded individuals that replacing an opposite region in the parents to produce the children of new generation. The mutation occurs with probability of 0.001, it changes the type of 0.05 sensors found in handled individual. After producing more than 1000 generations, the achieved results showed lower value of fitness function with stable behavior. This indicates the correct path of computations and the accuracy of the obtained results. The genetic algorithm operated well and directed the process towards improving the genes to be the best possible at the last generation. The behavior of the objective function started to be regular gradually throughout the produced generations until reaching the best product in the last generation where it is shown that all the sensors are connected to the nearest cluster head. As a conclusion, the genetic algorithm developed the sensors’ distribution in the WSN model, which confirms the validity of applying of genetic algorithms and the accuracy of the results.

Keywords— Wireless Sensor Network (WSN), High Sensors Range (HSR); Low Sensors Range (LSR); Genetic Algorithm (GA)

I. INTRODUCTION

Wireless Sensor Network (WSN) is a high potential technology used in many fields, such as agriculture, earth or environmental monitoring, resources union, health, security, military, and transport [1], WSN also is a crucial aspect in IoT technology [2]. Wireless Sensor Network (WSN) refers to community dedicated sensors that screen and record the physical situations of the environment and ahead the gathered records to an important vicinity. Environmental situations which include (temperature, sound, pollution levels, humidity, and wind) can be measured the usage of WSNs, which can be used broadly in lots of fields: weather forecasting, earthquakes detection, tsunami, water great control, fuel leakage, forest hearth prediction, volcanoes, site visitors tracking, and chemical threat detection [3]. WSNs have large quantity of sensors which connect with every other over short-variety wirelessly the usage of low electricity. A single sensor node is working on accumulating information relies upon on their domain programs, that is the reason why big wide variety of sensors are used to shape WSN consists of a few cluster head (CH) communicated with base station (BS), CH is then ship and receive the facts with the outside sensors that either being high sensors range (HSR) or low sensors range (LSR) as proven in Fig. 1 [4].

Sensor nodes were mounted randomly within the field for monitoring and detection duties the usage of the programs. energy efficiency is the main concern in WSN because a major amount of electricity is utilized in records transmission and thus schemes to select the exceptional direction need to follow for efficient energy intake [5]. Routing calls for most beneficial course from supply station to vacation spot to satisfy first-rate of carrier (QoS) needs in terms of delay, jitter, throughput, and bandwidth for quit customers with specific actual time applications [6]. The Wi-Fi networks including hierarchical primarily based optimized routing used distinct routing algorithms including genetic algorithm (GA) [5]. To reveal the power availability of the sensor gadgets which assist to gain the lifetime of the community, the advanced routing by GA described powerful overall performance. Furthermore, the entire upkeep of the first-rate of WSN services maintained as a multi-objective task as the stability among each tradeoff inclusive of coverage energy, latency, efficiency and network lifetime are treated continuously [7].
A. GA based WSN optimization

Several applications associated with environmental sensing counted within the method of WSN design, which are used widely in recent times in internet of things (IoT) structures. In systems, a crucial function is performed by means of sensor networks, those systems address regular environmental situations. Numerous sensing strategies of sensing that help to data gathering, along with remote sensing thru sensors of airborne and satellites, embedded, networked systems and mobile systems which include Wireless Sensor Networks. There are numerous complicated clustering methodologies [8] inside WSNs’ literatures in terms of power save. Literatures tackled the power save established via the working modes optimizing for sensors, such reason necessitate easy approach of clustering sensors is followed for the formation of clusters in the community as ordinary working modes with their closest cluster head (CH) sensor. Wherein, CH refers to the Access Point (AP), moving the signal among sensors (HSR and LSR) is its main function [9]. The optimization is handled either by minimizing the parameters that related to the power or maximizing the factors related to uniformity sensing. The most two important parameters for energy are: operating energy (OE) and communication energy (CE), at the same time as the most two interesting sensing factor uniformity are: sensors for each cluster head error (SCE) and sensors out-of-range error (SORE). Accordingly, OE and CE may be used as fitness functions for attaining most effective WSN shape which its overall performance is evaluated by the use of SCE and SORE. OE points to the energy consuming of the sensor at operational time. It essentially relies upon sensor operational mode, i.e., it works as a LSR, HSR, or CH, or inactive. The OE consumption parameter is given as [10]:

\[
OE = \left(20 \times \frac{N_{CH}}{N}\right) + \left(2 \times \frac{N_{HSR}}{N}\right) + \left(\frac{N_{LSR}}{N}\right)
\]

(1)

in which, NCH, NHSR, and NLSR are the quantity of LSR, HSR, or CH within WSN, and N equal to the summation of all clusters of types NLSR, NHSR and NCH together.

CE is consuming power related to the communication among sensors in ordinary cluster heads and operating modes. It particularly relies upon distance among CH and their sensors, which referred as CE and depicted as follows [11]:

\[
CE = \sum_{i=1}^{N_{CH}} \sum_{j=1}^{n_i} \mu \cdot d_{ij}^k
\]

(2)

Where, \(n_i\) is No. of \(i^{th}\) cluster, \(d_{ij}\) is distance from cluster head to \(j^{th}\) sensor while \(\mu\) and \(k\) are constants represent the topological characteristic of the WSN. For a determined application accuracy of monitoring of the open field, these factors are \(\mu=1\) and \(k=3\).

The WSN performance evaluation is usually related to the existence of network connection and satisfying all needed constraints, which includes both SCE and SORE. SCE is used to investigate that CH do not have a number of sensors is greater than the maximum allowed one in regular operating mode that is given as follows [12]:

\[
SCE = \begin{cases} 
\frac{\sum_{i=1}^{N_{full}} n_i}{N_{full}} & \text{if } N_{full} > 0 \\
0 & \text{otherwise}
\end{cases}
\]

(3)

In which, \(N_{full}\) is the cluster-heads (or clusters) with more than 15 energetic sensors of their clusters where \(n_i\) is the wide variety of sensors within the cluster i, (15 energetic sensors is the default number). Consequently, SCE is the physical verbal exchange capabilities of the sensors as well as their facts management talents in terms of quantity of statistics that may be processed by way of a cluster-head sensor. Even as, you could use the SORE to make certain that each sensor is able to connect with its CH that depends at the sensor’s signal range capability. Through supposing the HSR sensors cowl a curved region that has radius of 10 duration units, LSR sensors work on a 5 period gadgets radius circular area, then the number of energetic not-communicated sensors with their CH (Nout) relative to N indicates the overall quantity of energetic sensors within the network [11]:

\[
SORE = \frac{N_{out}}{N}
\]

(4)

The relations in Eq. s (1–4) to make appropriate objective characteristic during the assessment of the fitness feature of GA enable to give excessive pleasant for each possible answer of the optimization hassle [10, 11]. GA is an artificial intelligence seek approach based totally on Darwin’s concept of Evolution, which inspired with the aid of the idea that the one that is great adapted to the environment will continue to exist. usually, GA results in a populace of people, or a hard and fast of feasible answers to the hassle, subjecting it to random actions the usage of biological evolution of genetic mutations and recombination. Also, it helps to make the selection according to the criterion, based on which it is decided to be the most adapted individuals that survive through evolution. The most important concepts in GA are coding scheme and fitness Function, coding scheme is a possible solution to the represented problem, the most common encoding of solutions is through binary strings, real numbers, letters, or characters to represent chromosomes. Fitness function is the base of GA that computed for each individual as surviving score based on how good solutions they are to the problem [13]. Parameters that control the performance of GA are the probability of crossing, probability of mutation, population size, number of generations, etc. The following steps show the generic structure of the GA that depicted in Fig. 2 [14,15]:

1. An initial population of individuals usually is generated randomly.
2. Selection phase: The best individuals are selected.
3. Reproduction phase: The selected individuals are crossed by means of the crossing function, giving rise to a new generation that will replace the previous one.
4. Mutation Phase: Mutations are introduced into certain individuals in the new population at random. Mutation allows the introduction of the new chromosomal material into the population, just like what happens in their biological equivalents.

5. New generation is produced usually with better solutions than the previous one.
6. Genetic algorithms end either when reaching a specific number of generations or when they fulfil a stop condition.

B. Problem Statement

The uniform distribution of sensors in a uniform geographic area (square) may contain some problems when the half distance between each two of CHs is greater than the range of LSR and HSR. Thus, there are areas that cannot be covered and the sensors in those areas will be unconnected to the nearest CH. To overcome this problem, there are two solutions: The first lies in increasing the number of the CHs and spreading it in all areas where there are likely to be sensors, but the presence of a large number of CHs leads to the consumption of more energy, and this waste of energy is not intended. The second solution involves finding the optimum distribution of CH locations using optimization algorithms to be in locations that completely sufficient to cover all the sensors in the sensing area, which is the method that adopted in the present work.

C. Literature Review

Many literatures dedicated to WSN, they range in many elements which includes hypothesis, technique, or even the utility boundaries. there may be a brilliant deal of attention become granted to WSN [16]. several techniques were advanced for reaching more green techniques that serving the programs field of hobby. The maximum good-sized literatures are noted with information in the following:

1. An energy-efficient clustering routing protocol relying on the premise of a modified genetic algorithm worked on in [17] utilizing back propagation neural network for underwater wireless sensor networks. The three main stages of the proposed protocol are: the cluster head node selection, the cluster arrangement, and the data transmission. GA used to present multi-hop transmission paths, each made with security ideas to further develop transmission efficiency and improving network resources [17].
2. consumed strength changed into superior to in addition increase the life span of WSN via communiqué set of rules. The clustering of nodes improves the energy use of devices, wherein nodes are allowed to move as indicated to to be had energy device nodes popularity. GA turned into used for finding the cluster middle in the wi-fi community. assurance of cluster middle became executed based on node strength and the space from the sink node. identifying packets of information became transferred to the sink node straightforwardly from each cluster center [18].
3. Proposed a focal effective clustering protocol primarily based on GA to increase the three-D WSN stability length for better dependability and maximum excessive network lifestyles time. New genetic set of rules fitness function is applied to enhance the quantity of clusters in the network, which basically contains statistics approximately the type

![Fig. 2. General sequential stages of GA](image)
of connections that the member nodes lay out with cluster heads in a cluster and cluster heads establish with the sink within the community [19].

4. The dynamic task distribution of WSNs with considering the weakness of the wireless communication and the location of the nodes is proposed. This incorporates a real location algorithm that can recover the network from potential failures. GA is popularly utilized for optimizing the underlying solutions using mapper algorithm. An adaptive solution envelope (ASE) is produced to comply with the time constraint while extending the network lifetime and ensuring the security requirements. For a real scenario, the heterogeneous nodes were considered too [20].

5. Applying AI machine-learning (ML) algorithms on the sensor-statistics streams which can be given by way of the WSNs and designed to be installed in the pilot web sites producing occupant-conduct profiles that mimic the inhabitants’ movements and add to the accuracy of strength performance assessment, utilizing ML-empowered occupant conduct strategies [21]. The applying, technology, topology, and fee of sturdy WSNs are addressed to collect and ship actual-time information. The proposed real WSN considers different necessities within the field, constraints, and selections on the pilot destinations, and hence, it units the lines for similar institutions predicted by using the development enterprise for amassing dynamic data required for smart systems [22].

6. more desirable Multi-objective Non-dominated Sorting Genetic Routing set of rules (ENSGRA) is a new algorithm was proposed for positive modifications in allocating reference points and crossover operation. This upgrade can be performed with the aid of improving 3 objectives in deployment and routing degree: range of dynamic sensor nodes, electricity consumption, and community inclusion. ENSGRA became working in a hit dynamic WSNs scheme whilst evaluating with distinctive works in as some distance as calculation time and optimized route [23].

7. Chain routing algorithm is a clustering of sensors (CS) primarily based became proposed, WSN is separated into M routes, wherein the electricity productiveness and the reconstruction mistakes are capabilities of the quantity of paths. The energy performance and the variety of routes also depend on the transmission scopes of the nodes (R), whilst the reconstruction-errors likewise relies upon the satisfactory of the sensing matrix and the range of routes. To get the premier wide variety of measurements (Mopt), the surest sending range of the nodes (Ropt), and the most beneficial sensing matrix as mutual coherence, that raise the strength-performance and bounds the reconstruction errors, numerous goal genetic algorithms were used [24].

D. Contribution

Our contribution addressed in this research is considering that each WSN model is an individual consisting of one gene that contains codes consisting of a specific number of CH, LSR, and HSR with their locations in that model. The optimization lies in the process of mating these individuals by mixing half the number of this gene from one parent with the other half gene belonging to the other parent. Since the number of CH and its connected sensors in half of the first gene is not equal to their number in the second gene, the number of CH in the resulting individuals will be slightly more or less than their initial default number. Such increase in the number of the CH provides better connectivity and does not waste a lot of energy, and its objective function will often be better compared to the previous generation, and it will also be a candidate for mating in the next generation. Thus, the WSN model improved generation after generation with a few changes in the number of CH till achieving the best generations.

II. Method

The proposed technique consists of establish a most desirable WSN gadget such as two tiers: the primary degree is used to create a WSN model consistent with certain assumed conditions that may be used as a character inside the 2nd stage. the second one stage involves using the concept of GA to locate the quality distribution of sensors that consumed the least operating energy, the following subsections provide an explanation for greater information about the proposed GA based totally WSN optimization technique according to the sequential stages of proposed method shown in figure (3).

The proposed model includes non-uniform redistributing the CHs depending on the presence of a sufficient number of connection sensors of the two types: HSR and LSR to ensure that all these peripheral sensors are connected to the available CH (close and with a connection capacity). This is done by using a genetic algorithm to distribute CHs among the sensors, ensuring that they are all linked in a specific geographical area. Thus, the number of the CHs depends mainly on the capacity of the extension of this geographical area to ensure comprehensive coverage, and therefore if it is assumed that the capacity of the extension of area is 30x30 (i.e., Width=Height=30 length unit), then the assumed number of the CH (NCH) is 10% of the capacity of that area (i.e., NCH=90). Since each CH has a specific communication capacity with the sensors that distributed surrounding it, the number of HSR is NHSR=2×NCH =180, while LSR (NLSR) equals double the number of HSR (i.e., NLSR =2×NHCSR), such determinations of these parameters credit achieving a maximum number of sensors in the sensing area, where the CH are redistributed to obtain comprehensive coverage with lower cost. The maximum connection capacity of each CH is set to contain only 10 sensors of HSR and 10 of LSR, which is satisfied with the conditions of Equation (3). In this model, the CHs are distributed in random chosen locations within the sensing area, with a distance of LSR from the edges of the area, and with an interstitial distance greater than twice the LSR. Then a number of NLSR of LSR are distributed in randomly chosen locations between the CHs on all places in the sensing area, and then connecting each LSR to the closest CH. The same manner is applied on HSR that are distributed on the remaining empty locations of the sensing area, which are directly connected with closest CH. In the event that there is an LSR or HSR sensors located in a place far from the CH by a distance exceeds its range, it is considered to be not connected. The empty locations in between CH, LSR, and
HSR in the sensing area are considered to be contained inactive sensors.

A. WSN Model Optimization

The genetic algorithm was adopted in the process of optimizing the WSN model, where it was considered that each WSN model is an individual with one chromosome containing four genes: 1, 2, and 3 are representing the symbols of the types of CH, HSR, and LSR sensors respectively, while the gene 0 indicates there is no sensor in that location [25-35]. The parameters of GA were set to be as given in Table 1, while the implementation of the GA (Fig. 1 & Algorithm 1) is carried out according to the following work steps:

<table>
<thead>
<tr>
<th>TABLE 1. GA PARAMETERS VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Population size</td>
</tr>
<tr>
<td>Mutation rate</td>
</tr>
<tr>
<td>Crossover type</td>
</tr>
<tr>
<td>Generation size</td>
</tr>
</tbody>
</table>

Random Generation: A number of 100 individuals are generated to form the population that representing the first generation. Each individual is generated in the form of the proposed WSN model. This model consists of a number of CH, HSR, and LSR sensors with a number of NCH, NHSR, and NLSR in random locations of length 30×30 sensing area.

Selection: Selecting a number of chromosomes with amount 50% of the population size of the current generation to make pairs between the [36-45] The selected chromosomes are the ones with the best fitness function, which are 50 chromosomes (equivalent to 25 pairs) used to produce the next generation. The used fitness function is:

\[ F_k = w_1 \times J_1 + w_2 \times J_2 + w_3 \times J_3 + w_4 \times J_4 \]  

(5)

Where, \( w_1, w_2, w_3, \) and \( w_4 \) are contribution weights of fraction values between (0-1), \( J_1, J_2, J_3, \) and \( J_4 \) are objective functions representing the objective parameters OE, CE, SCE, and SORE given in equations (1-4) mentioned before [46-60].

Crossover: It is the process of mating between any pair chosen randomly from among the selected chromosomes in order to generate newly two individuals by single replacement as that shown in Fig. 4. There is no prohibition for any individual chosen at random to mating more than once, and this process continues until completing the number of individuals of the new generation. Thus, the crossover is exchanging the lower half area between each pair of individuals to produce other individuals belong to the next generation, this process repeated 50 times to produce 100 individuals of the next generation.

<table>
<thead>
<tr>
<th>Fig. 3. Block diagram shows the sequential stages of the proposed WSN optimization method</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
</tr>
<tr>
<td>Population</td>
</tr>
<tr>
<td>Selection</td>
</tr>
<tr>
<td>Crossover</td>
</tr>
<tr>
<td>Mutation</td>
</tr>
<tr>
<td>End</td>
</tr>
</tbody>
</table>

Fig. 4. Producing 4 children in the next generation by crossover the individual pairs

Mutation: The genetic mutation takes place in the same chromosome with occurrence probability of 0.001%, in which about 5% of the sensors are inverted into another type. Thus, about 45 genes (sensors) are randomly chosen to be inverted into another type except the type it was. Fig. 5 shows few genes of one type of CH, HSR, or LSR are chosen to be inverted due to mutation, they marked with a white circle.

| Fig. 5. Producing new individual in the next generation by mutation |
Regeneration: The process of the regeneration occurs when the optimal state of the individuals of the last generation is not reached, the new generations continue to be generated with developing the genes until meeting the stop condition of the GA. The stop condition is either to reach the state where there are a few changes occurred in the fitness function of about 0.01 (or no change at all), or reaching to 1000 generation as shown in Algorithm (1).

### III. RESULTS AND DISCUSSION

The two stages of WSN modeling and optimization were implemented according to the setting mentioned previously [26, 27], the achieved results are given in the following subsections. The proposed modeling method was implemented and WSNs were obtained with an area of 30x30 length unit containing 90 CHs were randomly distributed in the sensing region, each CH was associated with a number of HSR and LSR nearby. Fig. 6 shows the results of four WSN models obtained by the proposed method. It is shown that there few sensors are occupied places far from any CH, so they were not connected to them, this is due to their random distribution in the sensor area. Table 2 presents the amount of the computed objective functions for the four resulting WSN models. The results show that although the distribution was random, it is considered regular because most of the sensors were connected to the CH close to it, and this is what makes these WSN models acceptable and can be used in the subsequent optimization process.

![Resulted four WSN models using proposed WSN modeling method](image)

**Fig. 6.** Resulted four WSN models using proposed WSN modeling method

The values of the contribution weights for each objective function were adjusted according to their subject matter importance as: w1=0.4, w2=0.3, w3=0.2, w4=0.1. By observing the behavior of F, it is found that its values were higher for the first generation relative to their values for the last generation, and this indicates that the process of producing generations was directed towards improving the genes to be the best possible at the last generation.

Also, it is observed that the behavior of F in Fig. 7 contained more fluctuations in the first generation, while this behavior had less fluctuation in the behavior of the last generation, and this indicates the stability occurred in the genes of the last generation. To search more on this subject (gene stability with more generations), the fitness function F in Fig. 7 has been analyzed into its four components: J1, J2, J3, J4 shown in Fig. 8 and Fig. 9.

### B. WSN Optimization Results

The genetic algorithm was implemented based on the parameters setting given in Table (1), where there 100 individuals were generated from a WSN model whose average objective function was shown in Fig. 6, this was for a number of generations of 1000. Then, the last best generation was obtained. The optimization is determined by the fitness function shown in Fig. 6 relative to that of the first generation.
It is noticeable that the reason of fluctuations in the $F$ function in each generation is due to the behavior of the fourth objective function ($J_4$) of the higher values, which represents the amount of inactive sensor in each individual, which appears different according to the different distribution of the sensors in each individual. Also, the fluctuation in the first objective function ($J_1$) appeared effective due to different number of CH were shown in individuals of last generations. The NCH begins with constant value (90 CH), whereas it is noticeable that this value was varied with more generation due to the single crossover. The rate of their behavior will appear more random among individuals belonging to one generation, and then the function improves slightly with the progress of generations and its fluctuations are appeared as few as possible in the last generation as shown in Fig. 10. This change in NCH is an important result worth interesting. Fig. 11 shows the four randomly selected WSN models of the new generation, it is shown that all the sensors in them have been connected to the nearest CH that occupied location that guarantee the connection for all sensors of HSR and LSR type with the least amount of objective functions: less power, less cost, and less error. This indicates the extent of benefiting from GA to develop the distribution of sensors in the WSN model, and this confirms the validity of the application of GA and the accuracy of the obtained results.
IV. CONCLUSIONS

It concludes that the proposed two-dimensional model for representing WSN models was successful in showing a WSN model with a specific capacity of sensors connected to the cluster head. Also, the genetic algorithm was successful in the process of differentiating models, and its application to the proposed model has given distinct results, represented by generations with high quality characteristics. It is hoped that the work will be developed to apply the clustering operations such as k-means on the WSN models resulted from the genetic algorithm. Also, the proposed method can be applied to mobile sensors.

ACKNOWLEDGEMENTS

The authors are grateful to the Iraqi Ministry of Higher Education and Scientific Research (MOHESR) for technically supporting the current research.

REFERENCES


