Research Article

An Improved Niche Chaotic Genetic Algorithm for Low-Energy Clustering Problem in Large-Scale Wireless Sensor Networks

Min Tian, Jie Zhou, and Xin Lv

1College of Information Science and Technology, Shihezi University, Shihezi, Xinjiang 832003, China
2The Key Laboratory of Oasis Ecological Agriculture of Xinjiang Production and Construction Group, Shihezi University, Shihezi, Xinjiang 832003, China

Correspondence should be addressed to Xin Lv; lxshz@126.com

Received 4 October 2017; Revised 13 January 2018; Accepted 6 February 2018; Published 1 April 2018

Academic Editor: Jaime Lloret

Copyright © 2018 Min Tian et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Large-scale wireless sensor networks consist of a large number of tiny sensors that have sensing, computation, wireless communication, and free-infrastructure capabilities. The low-energy clustering scheme is usually designed for large-scale wireless sensor networks to improve the communication energy efficiency. However, the low-energy clustering problem can be formulated as a nonlinear mixed integer combinatorial optimization problem. In this paper, we propose a low-energy clustering approach based on improved niche chaotic genetic algorithm (INCGA) for minimizing the communication energy consumption. We formulate our objective function to minimize the communication energy consumption under multiple constraints. Although suboptimal for LSWSN systems, simulation results show that the proposed INCGA algorithm allows to reduce the communication energy consumption with lower complexity compared to the QEA (quantum evolutionary algorithm) and PSO (particle swarm optimization) approaches.

1. Introduction

The advancement of microsensors, microelectromechanical systems, and communication make economically and technically feasible a network composed of low computational complexity, numerous, vulnerable, and fast-response wireless sensors [1]. Large-scale wireless sensor networks are composed of a large number of sensing units that have limited computing, communication, sensing, and free-infrastructure capabilities [2–4]. Each sensor node consists of four units, including the transmitter unit, the computation unit, the signals sensing unit, the data storage unit, and the battery unit [5]. Large-scale wireless sensor networks (LSWSNs) have been widely studied and usefully employed in many areas such as target detection and tracking, advanced health care delivery, intelligent family, military affairs, multimedia surveillance, and environmental monitoring [6].

2. Problem Statement and Our Contributions

Generally, sensor nodes are small devices with limited monitoring capabilities [7]. Because sensor nodes are equipped with restricted power device, it is very important and necessary to explore new low-energy clustering algorithms [8]. Such a problem is important in enhancing network lifetime [9]. Generating an optimal low-energy clustering for large-scale wireless sensor networks with restricted power device and communication energy consumption constraints is very useful but is an NP-hard problem [10]. While exhaustive search is recognized as one possible solution to low-energy clustering problem [11], its computational complexity is too high to be implemented for practical real-time applications [12].

Due to its computational complexity, many heuristics have been proposed to obtain near-optimal solutions in reasonable time [13–18]. For the low-energy clustering problem,
the authors proposed in [19] a PSO-based low-energy clustering method, which is found to be more efficient than GA methods. It provides a wider search space by randomly selecting low-energy clustering solutions when each solution in population gets updated. However, PSO suffer from either low convergence rate or high computational complexity. In [20], a low-energy clustering technique that enables trade-offs between the computation cost and communication energy consumption is investigated using fuzzy logic. Their work focused on a low-energy clustering optimization with fuzzy logic. The fuzzy logic approach is simple and fast, but it usually yields high computational cost with large-scale WSNs. There are several approaches based on the intelligent clustering protocol, which attempt to solve this issue by using the concept of heuristic [21]. The proposed method uses a distance-aware clustering protocol instead of a traditional clustering method. However, the protocol is complex, and the computation complexity is still high.

Based on a combination of quantum theory and evolution theory, such as rotation gates, superposition of states, and quantum bits, the authors applied the QEA to the low-energy clustering problem [22]. Compared with traditional genetic algorithm, QEA is robust and global in operation. The QEA with Q-bit representation can explore the search space with quicker convergence speed. This enables the solution of low-energy clustering problem and reduces the computational burden.

Recently, the concepts of niche computing and chaos theory have motivated the new development of evolutionary theory. Evolutionary algorithms have been shown to be effective in optimizing multidimensional problems. Hence, an improved niche chaotic genetic algorithm- (INCGA-) based low-energy clustering approach is proposed in this article. To evaluate the algorithm’s performance, we first model the low-energy clustering problem as an integer programming that is proved to be NP complete. Also, in order to enhance the search ability of the proposed genetic algorithm on the low-energy clustering problem, we add the niche operator and chaotic operators to speed up the convergence rate. It combines the merits of niche selection algorithm that considers various features and a chaotic generator that enhances the convergence rate. This can maintain the diversity of the algorithm and improve the algorithm’s global search ability. INCGA could adjust the parameters automatically as well as appropriate partitions of the populations and avoid local optima. INCGA also uses chaotic generator strategy, which aims to avoid local optima. Simulations are conducted for the LSWSN with low-energy clustering methods based on INCGA, QEA, and PSO. Simulation results show that the proposed INCGA method significantly outperforms the QEA and PSO method.

3. System Model

This section describes a model of low-energy clustering in LSWSN with respect to the constraints of cluster head percentage and communication energy consumption. A typical low-energy clustering model in LSWSN is represented in [23]. The proposed model is of max–min type with nonlinear constraints. Later in [24], the authors showed a similar model by a fewer number of sensors and cluster head nodes. We design an optimizing model minimizing the communication energy consumption for LSWSN.

In LSWSN, the transmission radio energy could be represented as

\[ \text{Cost}_\text{send}(x, y) = E_{\text{elec}} \cdot x + \epsilon_{\text{amp}} \cdot x \cdot y^3, \]  

where \( y \) is the distance between two nodes, \( \text{cost}_\text{send}(x, y) \) is the communication radio energy, \( x \) is the size of transmission bits, \( n \) is the path-loss exponent, \( \epsilon_{\text{amp}} \) is the power amplification parameter, and \( E_{\text{elec}} \) is the electronics energy parameter. The received communication energy of \( x \) bits data can be represented as

\[ \text{Cost}_\text{rev}(x) = E_{\text{elec}} \cdot x, \]  

where \( \text{cost}_\text{rev}(x) \) is the receiver dissipated energy for receiving \( x \) bits. In the actual system, \( E_{\text{elec}} \) for sending and receiving are not exactly the same. This thesis uses a simplified model and we assume that the \( E_{\text{elec}} \) for sending and receiving are the same. In this paper, we set \( k = 1 \) Mbit, \( \epsilon_{\text{amp}} = 100 \text{ pJ/bit/m}^2 \), \( E_{\text{elec}} = 50 \text{ nJ/bit} \), and \( n = 3 \).

Consider a LSWSN system with \( M \) sensors in the monitoring area. The gateway node is located in the center of the field. The distance \( y \) represents the distance between two random sensors. Each sensor node sends \( x \) bit data to its cluster head node and then to the gateway node. We need to choose the cluster head according to the predetermined ratio in advance. We number each sensor node and indicate the set of cluster head node. When the cluster head node is determined, the energy consumption of the communication is determined according to (1) and (2). So the input of the system is the set of cluster head nodes, and the output is the total communication energy consumption of the system. The total communication energy consumption can be represented as

\[ E = \sum_{n=1}^{N} \text{cost}_\text{send} + \text{cost}_\text{rev}, \]  

where \( N \) is the total number of sensor nodes in LSWSN and \( n \) is the order of the sensor.

4. INCGA for Low-Energy Clustering Problem in Large-Scale Wireless Sensor Networks

The genetic algorithm (GA) is a stochastic optimization algorithm, which is inspired by the biological principles of evolution theory, which include natural selection and mutation. GA was originally proposed by Holland in 1975 as a computational strategy to artificially model biological evolution. GA is usually utilized by lots of function optimization issues and is confirmed to be wonderful in finding suitable and near-optimal answers. Typically, GA has comprehensive explore skills, and the technique is routinely not successful for large-size complex problems. In most cases, GAs are iterative, and typically, they converge to a local optimum.
Currently, chaos theory- and evolutionary theory-inspired strategies are actually being formed. They turned out to be very effective in researching for effective solutions, aiming at an enhanced convergence and resolving complex optimization issues. Here, an improved niche chaotic genetic algorithm (INCGA) is proposed, that is, an alteration of GA, for minimizing the communication energy consumption in large-scale WSNs. In INCGA, according to the conception and strategies of chaos theory and evolutionary theory, novel niche approaches are being designed, which could have a much better stability between searching and get more attractive results in contrast to GAs. In this mixed strategy, the scan abilities of niche selection algorithm and chaotic generator are incorporated to rapidly get convergence. It undergoes parallel searching in complex search places.

In the time of every iteration, INCGA creates a new population using the last population by utilizing different operations, including mutation, crossover, and selection. In addition to that, the presented niche selection operator is utilizing niche theory, and the provided chaos strategy is using the chaotic generator to help improve the global searching ability of INCGA.

4.1. Encoding. In INCGA, an answer is shown by an individual. We use a vector as a chromosome to represent an answer to the low-energy clustering problem. Every solution might depict a number of arbitrarily chosen cluster head nodes. The low-energy clustering problem is transformed into finding out a cluster head node set in the prospect solution space using the INCGA. In this way, a potential solution will be shown as a string. Therefore, binary encoding is used on every individual. Every single a sensor node within the system corresponds to a gene. The chromosome contains Boolean factors denoting whether the corresponding sensor node is selected as the cluster head node or not. The gene quantity is equal to the quantity of sensor nodes in LSWSN, for instance, assuming a chromosome is \( C_1 = (0, 1, 0, 0, 0, 0, 0, 1) \). There are 8 sensors in the region and No. 2 and No. 8 nodes are selected as cluster head nodes.

Such encoding is adequate and powerful since it contains the whole searching region. Moreover, the length of the string is the number of design variables. As an example, for LSWSN with 100 sensor nodes, the entire string length is 100. Here the lookup region is the searching area of all solutions, which meet the cluster head proportion constraint. That is, the \( i \)th gene matches the \( i \)th sensor node. As stated, the number of all combinations is \( 2^{100} \).

4.2. Generation of Initial Population. INCGA keeps a population of individuals, and the size of the population is called the population size. The INCGA gets started by producing a set of random applicant solutions called the initial population. In every individual, a randomly selected amount bit within the individual is set to random binary numbers with logistic map in (4). According to the proportion and the number of cluster heads, sensors corresponding to largest random number will be selected as the cluster head. For implementing this, the initial set is constructed of randomly designed individuals. Every individual in the set would be assigned randomly. In this way, the initial population is constructed. This procedure will then be employed over and over again until every individual is ready. The initialization operation is kept random to ensure population diversity.

\[
z_{r+1} = 4z_r(1 - z_r)
\]

4.3. Selection. For every single generation, there will be a number of children produced by merging the features of mother or father chromosomes. The algorithm selects an individual inside the latest population with chances relative to its valuation result.

Solutions in the present population are considered dependent on their advantage to survive over the next population. This implies that each solution in a population is connected with a amount of advantage as well as a fitness value. When a couple of parent solutions are to be selected through the current population, foremost, the fitness is assessed by (3).

The selection of parents would be determined by arbitrary and specific strategies. In each generation, mothers and fathers are chosen by function value. High-value chromosomes are often employed to produce children compared to other chromosomes, which likely enhances the population quality. Considering that the purpose is to minimize the communication energy consumption, a binary vector with fairly small fitness must have basically superior valuation, which means the individuals that are better would have a greater probability to be selected. It imitates natural selection process.

The fitness-based selection is put into practice for selecting the individuals. Fitter individuals will get a fairly greater chance to be picked for processing. The idea of fitness-based selection is easy. The chance of every individual is based on the fitness by using the scaling. Making use of the scaling, the chance of picking out the \( i \)th chromosome for processing is proportionate to the fitness.

In every stage, mothers and fathers are chosen foremost in accordance with the fitness-based selection. The procedure is going to be employed continuously 1/2 population size times until each and every solution becomes achievable. With this, in the crossover operation, vectors were separated into 1/2 population size couples, with each and every two producing a couple of new individuals. After that, the components could be merged to develop a new individual (child). The previously mentioned selection operation enables reliable constraint, boosting the quality of solutions by selecting the child solutions from high fitness locations.

4.4. Crossover. While selection decides which reproduce, the crossover operation produces novel patterns to search the pattern area. To produce novel binary chromosomes, as well as layout features, the crossover operation picks chromosomes within the mating pool to generate children. The reproduction procedure permits the genes belonging to the mating pool being forwarded to the next iteration while producing novel structures, which might turn out to be effective. The children as a consequence of crossover obtain characteristics coming from the mating pool and
therefore maintain an enhanced possibility of including an alternative pattern having a improved fitness compared to the past iteration. If the children pattern is bad, alternatively, it is going to be removed over the following selection operation. By doing this, crossover tries to take advantage of the information which was included in the last generation with the development of the parents, as well as discovering unknown parts of the searching area.

Joined together, crossover and selection offer INCGA to be able to effectively identify good regions of the searching area. Generally, there are a number of various types of the crossover operation. Nevertheless, scientific research has revealed that uniform crossover is quite efficient for the detection of novel individual styles by generating various children. Uniform crossover, similar to several crossover strategies, starts with an arbitrary choice of two individuals from the mating pool produced by selection. The couple of the individuals will be mated to generate two offspring, which would consist of a portion of the following iteration. When individuals get mated, they will be removed, and a couple of novel individuals are chosen randomly from the mating pool to create children. The operation is repetitive till a novel children population is full.

The main distinction between different types of the crossover operation is the procedure during which a couple of the individuals generate children. In uniform crossover, individuals are mixed so that every bit spot within the individuals can be a crossover position. Each position of the individuals generate children. In uniform crossover, chromosomes 1 and 2 are chosen from the population that were comprised of the problem. The solutions with the most desirable values of every solution are measured. Every solution needs additional assessment.

4.5. Mutation Operation. The next operation in INCGA is mutation. The mutation operation promotes diversity by marginally transforming the individuals in the new generation. This operation offers novel changes into the genes for a more substantial searching capability. Mutation usually happens at a reasonable possibility, arbitrarily changing a small proportion of the genes in the individuals. For instance, take into account the individual of 1110000. In case the third bit of the individual is mutated, the new individual turns into 1100000.

Mutations could be carried out in a number of strategies, the most basic of which is a bit flipping with a certain possibility. If the choice from the bit flipping would be to mutate, the present bit is going to be transformed to its opposite number of binary. Nevertheless, this technique requires creating an arbitrary binary for each bit in each individual. A reduced complexity strategy is the random mutation operator, which employs a uniform distribution to get the position of the upcoming bit to get mutated. Other sorts of mutation strategies that take advantage of structure of problem may also be applied.

4.6. Recombination. After applying the crossover operator and the mutation operator, there had been 2 * population size chromosomes in the population that were comprised of population size parent strings and population size offspring strings. Distinct from the application in GA, the selection and recombination operations in INCGA are blended thoroughly to develop a more suitable population. Recombination techniques are a type of uniform crossover which makes one offspring for each two strings. Right after recombination, the new offspring fitness will be assessed based on (5).

4.7. Fitness Computation. In the evaluate stage, the fitness values of every solution are measured. Every solution within the population is evaluated according to the fitness of the problem. The solutions with the most desirable fitness are chosen. In low-energy clustering problem, the target is to decrease the communication energy consumption. So the solution fitness will rely on the cluster head set selection. The fitness associated with the cluster head set is then evaluated.

$$\text{Fit}(S) = \sum_{n=1}^{N} \text{cost}_{\text{send}} + \text{cost}_{\text{rev}},$$

where $N$ is the number of sensor nodes in LSWSN.

4.8. Niche and Elite Operation. The objective of niche and elitism is to ensure that the optimal solutions from past iterations are forwarded to upcoming iterations. Therefore, the optimal individual within the population will be better from iteration to iteration. Keeping the formerly found best solution makes sure that the INCGA will never be necessary to
search again for good regions of the space, and in addition, it provides the elite patterns additional mating chances.

Similar to the other operations, various editions of elite maintenance strategy have already been designed, and a variety of the most suitable edition is problem specified. For this application, a particular proportion of the top performing patterns from the past iteration switches the equal proportion of the worst patterns of the present population. The installation of the elite patterns takes place following the fitness assessment of the present population, just prior to selection. This strategy boosts selection stress for the elites, which may increase overall performance but could additionally negatively influence diversity. Nevertheless, uniform crossover may counterbalance the decrease of diversity, which might be caused by this ingredients of elitism.

4.9. Clone Operation. The INCGA gets underway using the creation of an initial population, commonly by distributing arbitrary binary throughout the research area, which means that this fitness is graded in reducing sequence. The best 10% points are then specified for clone. The solutions are then copied. The number of copies is proportional to the fitness. After that, the algorithms carry out the mutation strategy. To save the data, the mutation operation only works on the cloned strings. For implementing this, the effective use of the clonal mutation will permit a neighborhood survey around the original string, while the utilization of the chaotic mutation permits a global lookup around the chromosome. Afterwards, the cloned individuals are examined with the objective valuation, and basically, the top of each clone is allowed to transfer to the next generation, holding precisely

Figure 1: Communication energy consumption with 100 (a), 200 (b), 300 (c), and 400 (d) sensors and 10% of cluster heads.
the same dimensions of the population. Using this type of replacement, the variety is preserved, and new areas of searching space can be probably researched.

5. Simulation Results and Discussion

In this section, we test the performances of INCGA method with the QEA and PSO method for low-energy clustering in LSWSN. With different numbers of sensors and cluster head nodes, we demonstrate the capabilities of the INCGA via simulations. We implemented the simulations in MATLAB on a personal computer with 2.5 GHz dual-core Intel Core i5 processor and 4 GB RAM. We used a fitness function (4) in order to compare the efficiency and convergence speed of the INCGA with respect to other two heuristics. The sensors are deployed in a 100 × 100 rectangle area, and the coordinates of sensors are randomly generated within the area.

In the simulation experiments, we implement INCGA with QEA and PSO. The maximum generation of INCGA, QEA, and PSO is set to 100. In INCGA, the population size is 50 (the same as that in QEA and PSO). In INCGA, the crossover rate is 0.8 and the mutation rate is 0.08. The PSO parameters values are $c_1 = c_2 = 2$, the dynamic range of the particle has been set to 0.4, and the maximum velocity is 4.

Figure 1 illustrates the convergence curve for INCGA, QEA, and PSO with 100 iterations. It compares the communication energy consumption computed by INCGA, QEA, and PSO on the low-energy clustering problem with 100,
Table 1: Communication energy consumption with 10% of cluster head nodes.

<table>
<thead>
<tr>
<th>Number of nodes</th>
<th>INCGA</th>
<th>PSO</th>
<th>QEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>125.53</td>
<td>135.66</td>
<td>143.28</td>
</tr>
<tr>
<td>400</td>
<td>182.98</td>
<td>203.08</td>
<td>232.16</td>
</tr>
<tr>
<td>600</td>
<td>256.27</td>
<td>306.32</td>
<td>413.48</td>
</tr>
<tr>
<td>800</td>
<td>301.54</td>
<td>357.68</td>
<td>518.93</td>
</tr>
<tr>
<td>1000</td>
<td>360.21</td>
<td>467.59</td>
<td>710.20</td>
</tr>
</tbody>
</table>

200, 300, and 400 sensors, respectively. The ratio of cluster heads is 10%. Only the global best solution at each iteration is recorded. In order to show the convergence, the result is the value of just 1 run. The comparison between the given methods reveals that the performance of INCGA method is significantly superior to that of niche computing, evolutionary theory, and chaos theory. From Figure 1(a), at the initial iterations, the initial communication energy consumption found by INCGA is lower than that by QEA and PSO. It can be seen that the proposed INCGA converges after around 50 iterations, which is so fast that it can meet real-time requirements of low-energy clustering. QEA and PSO show stagnation. Actually, it is proved that INCGA can jump out of the local optimal solutions. From 50 to 100 iterations, INCGA has approached close to 67 J, while QEA and PSO are still higher than that. Over all 100 iterations, INCGA provides a lower energy consumption than QEA and PSO, which means that it converges with a faster rate. Similar conclusions can be obtained from Figures 1(b)–1(d).

Figure 2 compares the communication energy consumption computed by INCGA, QEA, and PSO on the low-energy clustering problem with the cluster head proportion of 5%, 10%, 15%, and 20%, respectively. For each experiment, the methods have been executed 100 times, and all the results are averaged over 100 Monte Carlo runs. Table 1 shows the best results of INCGA, QEA, and PSO after 100 iterations (corresponding to Figure 2(b)). In all simulations, the performance of INCGA is better than that of the other two heuristic algorithms. For example, as shown in Table 1, the proposed INCGA achieves a communication energy consumption at 360.21 J with 1000 sensor nodes. Meanwhile, the communication energy consumption obtained by the PSO and QEA are 467.59 and 710.20, respectively. Similar conclusions can be obtained from Figures 2(a), 2(c), and 2(d).

According to the simulation results, when the number of sensors in wireless sensor networks is close to or less than 200, the performance of PSO and QEA is close to INCGA. The scale of LSWSN is generally larger than 400 nodes. When the number of sensors is greater than 400, the INCGA has obvious advantages. Therefore, the algorithm should be selected according to the network size. The rules are as follows: when the number of sensors is close to or less than 200, the performance of all algorithms is close, and all three algorithms can be selected. When the number of sensors is greater than 400, INCGA should be selected.

6. Conclusion

In this paper we propose an improved niche chaotic genetic algorithm (INCGA) for low-energy clustering problem applied in LSWSN. To evaluate the algorithm’s performance, we first model the low-energy clustering problem as an integer programming that is proved to be NP complete. Simulations are conducted to show the performance of our proposed INCGA scheme against QEA and PSO. Simulation results show that the proposed INCGA scheme outperforms the conventional QEA and PSO schemes with less communication energy consumption.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References


