Optimization of wireless sensor networks routing protocol based on genetic ant colony algorithm

Sun Zeyu

Department of Computer and information Engineering, Luoyang Institute of Science and Technology, Luoyang, China

_____

ABSTRACT

The combination of traditional ant colony algorithm in solving the optimization process to consume a large amount of time, easily falling into local optimal solution and convergence is slow and other disadvantages, while also generating a lot of useless redundant iterative code, operation efficiency is low. Therefore, ant colony optimization algorithm is proposed. The algorithm based on genetic algorithm has the ability to search the global ant colony algorithm also has a parallel and positive feedback mechanisms. Changes in the use of genetic algorithm selection operator, crossover operator and mutation operator action to determine the distribution of pheromone on the path, the ant colony algorithm for feature selection using support vector machine classifiers for evaluating the performance characteristics of the feedback sub-Variorum And by changing the pheromone iteration, parameter selection and increase the local pheromone update feature nodes guided the re-combination. The algorithm uses probability expectation values are obtained to meet under the conditions with minimal sensor nodes, and gives the optimal coverage and connectivity probability models and reasoning. The experimental results show that, the algorithm can not only use the least nodes complete the effective target area to be covered, and in reducing the network energy consumption is also greatly improved, simultaneously reduces the cyber source configuration, improve the network life cycle.

Key words: Wireless sensor networks; Genetic ant colony algorithm; Network lifetime; Network model

_____

INTRODUCTION

Ant colony algorithm (ant colony algorithm) was put forward by Dorigo, Maniezzo and Colorni in 1991, which is the original of the ACO algorithm.The algorithm is applied to the Traveling salesman Problem early ('ll Salesm and Problem, TSP) scheduling Problem, assignment Problem, Problem solving, etc.Feature selection problem is a NP problem, selection methods are divided into two categories: the Filter and Wrappe. Its main idea is to search a close to the approximate optimal solution.Process is characteristic of the Filter class mainly adopts specific independent inductive learning algorithm, and its main characteristic is to the whole path search speed faster, but the algorithm in the application process.

A precocious phenomenon, make the low accuracy;Wrapper class also has the characteristic of inductive learning algorithm and its performance as the characteristics of the evaluation and selection standard, its main characteristic is algorithm accuracy is higher, but each to a subset of the evaluation to the restructuring of the calculation, the computational complexity is high, not suitable for large-scale data search and calculation.Combined with the Filter and Wrapper is two complementary mode, a combination of both for the study of ant colony algorithm provides a larger value [1].Literature [2] is will be introduced to the weights of the pheromone updating process, improve the global search ability by repeated iteration to obtain the shortest path.Reference [3] is used by the transition probability and pheromone distribution estimation algorithm combining produce new individuals to improve the convergence speed and global search ability to inhibit earliness and defects[4] proposes a genetic ant colony algorithm, the idea is to use genetic algorithm to search the global first, and then use ant colony algorithm for the

solution of optimization is to achieve the objective of the optimal solution.

Improved genetic ant colony algorithm, this paper USES the Filter and Wrapper class characteristics, together with mutual information as the heuristic function, the use of ant colony algorithm to dynamic search feature subset, and combining classifier classification effect as evaluation criteria, based on the characteristics of the node classification precision guidance pheromone update, based on the genetic algorithm to produce a more optimal solution of dynamic change crossover operator and mutation operator at the same time, the solution of the optimal path pheromone intensity, and then according to the ant colony algorithm after completion of a traversal of path pheromone update and local pheromone iteration, global update strategy to enhance the ability of global optimization, speed up the convergence speed, the system improves the searching efficiency and avoid the premature phenomena in the process of iteration, restrain redundant code, makes more accurate precision effectively inhibit the occurrence of the phenomenon of premature convergence, improves the genetic ant colony algorithm in the limitations of global convergence and improves the convergence speed, the experimental results show that the algorithm has higher efficiency.

In addition, with the increase of the sensor nodes and coverage region changes, the complexity of the algorithm will become larger, and thus computational efficiency will reduces [5][6][7]. The coverage configuration protocol proposed by reference [8][9][10] is to judge the distributive node functions qualification using the local node location information, which is when the communication radius is greater than or equal to 2 times the sensing radius and the network k-cover the given convex region in the region[11][12], the network is k-connected, thereby promoting its agreement to meet the multi-coverage requirements. LEACH protocol has been improved in reference [13][14][15]. The main idea is: clustering routing algorithm uses periodic random selection of cluster head node to balance the node energy consumption and to achieve the purpose to extend the network lifetime [16][17][18]. Under the premise of maintaining a certain coverage and connectivity rate, in the operation of each data retrieval collected by the sensor nodes [19], a subset of a sensor node is selected to collect data and to realize nodes subsets dynamic scheduling conversion mechanism [20][21], and thus achieve the optimal choice for each subset of data acquisition, thereby optimizing the number of nodes, improving network stability and extending the lifetime of wireless sensor networks [22][23]. Finally, the number relations of nodes needed to be deployed in coverage and connectivity is given under the condition whether there is influence of edge node.

PROBLEM DESCRIPTION
1. SBM model
The coverage and connectivity this paper studiesis based on local positioning algorithm and has the following basic assumptions:

Assumption 1: The covering radius and communication radius of each node showed a disc-shaped.

Assumption 2: the specific location information of its own node is obtained through some localization algorithm, such as: the RSSI location algorithm, the Euclidean positioning algorithm, without the participation of large-scale physical communications equipment.

Assumption 3: All sensor nodes are randomly and uniformly deployed in square in side length $l$monitoring region $\Omega$with consideration of the boundary of deployment region.

Assumption 4: The perception of each node is much smaller than the entire network coverage region, and all nodes are isomorphic and initial energies are equal.

Definition 1: Set up a two-dimensional plane E, the coordinates of the nodes $s_i(x_i, y_i)$, the coordinates of the target $t_k(x_k, y_k)$, the Euler distance between the nodes s_i and t_k is less than s_i, perception radiusr_s : that is $ES_i = \left\{ t_k \middle| t_k \in E^2, D(s_i, t_k) \le r_s \right\}$, said point t_k is covered by s_i.

Definition 2: for a given target set $T=\{t_1, t_2, t_3...t_k\}$ and the sensor node set $V=\{v_1, v_2, v_3...v_n\}$, in a time slot, if the target set T in any goal t_k is covered by at least one node within node set V , then the target set T is full coverage.

Definition 3: suppose that nodes s_i, s_j, the target regions they cover are C_i and C_j, and $C_i \cap C_j \neq \varnothing$, so nodes s_i and s_j are coverage connected. Suppose that T is am nodes set randomly distributed in the target region, E is the set of edges of the network diagram, indicating that the $e_{ij} = 1$ positional relationship in e_{ij}; e_{ij} represents the positional relationship of node s_i and the target node t_j. When $e_i = 1$ when and only when the Euclidean distance of target node t_j and the node s_i is less than or equal to the perception radius r_i, otherwise $e_i = 0$. $W = \{w_1, w_2 \cdots w_n\}$ is the initial energy

_____

set of sensor nodes; W is in normal distribution; the w_i represents initial energy of sensor node s_i; w_i is the maximum energy in the process of node work.

Definition 4: In the monitoring region any point at least at the same time located in the the perception range of K nodes is known as the K re-cover. That is $A_k \in \bigcap\limits_{i=c}^{c+k} As_i$, where $1 \leq c \leq n$.

Definition 5: coverage rate of a point s_i in coverage region is as follows:

$$p(s_i, s_j) = \begin{cases} 0 & \text{if } R_s \leq d(s_i, s_j) \\ e^{-\varepsilon d} & \text{if } (R_s - R_e) < d(s_i, s_j) < R_s \\ 1 & \text{if } d(s_i, s_j) \leq (R_s - R_e) \end{cases} \quad (1)$$

Of which: $d = d(s_i, s_j) - (R_s - R_e)$, ε is physical parameters of the sensor node; R_e represents the monitoring dynamic parameters of the sensor nodes; $d(s_i, s_j)$ represents the Euclidean distance between the sensor nodes s_i, s_j ; When the coverage $p(s_i, s_j)$ of a node s_i meets $d(s_i, s_j) \leq (R_s - R_e)$, then the node is detected.

In the balance accelerate convergence and reduce the stagnation is one of the difficulties of the current research of ant colony algorithm.The algorithm as the search space as large as possible, to find the possible solution of the optimal solution space;Current at the same time, should make full use of community with the effective information, adapt algorithm focuses on those who may have a higher value of individual space, with large probability convergence to the global optimal solution.With dynamic weighting method, based on the ant solution quality is different, give different weights.According to different weight, the characteristics of ants by pheromone adaptive update.Therefore, to determine ant colony is gathered to one or a few solutions or scattered among the search space.Can through the calculation of the average of the iterative solution of all the characteristics of the subset classification error rate and the iterative solution of the minimum feature subset classification error rate to determine the differences between degrees of polymerization of ant colony.Can be concluded if the whole ant colony is aggregation, then the difference is smaller than relatively decentralized ant colony.So can be the difference between the two as a measure of whether ant colony convergence.

$$E_{avg} = \frac{1}{n} \sum_{k=1}^{na} E_k \quad (2)$$

Is ants in the iteration of the structure of the solution of the feature subset classification error rate, when updating pheromone, for less than ants gives a relatively large weight, the less you give weight, the greater the for more than or equal to the ant to give a relatively small fixed weights.So the ants in the current iteration of the selected feature on the release of pheromones as follows:

$$\Delta\tau(f_i) = \begin{cases} Q\lambda_k / E_{avg} \\ 0 \end{cases} \quad (3)$$

One type is obtained under the weight of:

$$\lambda_k = \begin{cases} (E_{avg} - E_k) / (E_{avg} - E_{min}) \\ 0 \qquad E_k \geq E_{avg} \end{cases} \quad (4)$$

Can see the closer, the weights of lambda approximates to 1.According to the formula (4) the structure characteristics of the calculation of each iteration the ant, again according to individual ants tectonic characteristics of solution and the differences between the corresponding pheromone for change.So you can increase the diversity of the solution space, improve the global search ability of ant colony, to avoid local convergence too fast, at the same time also won't reduce the search speed of ant colony.

In the ant colony algorithm, the ant principle of selecting the next node structure solution and pheromone update

rules of ant colony can search to the optimal solution play a decisive role.

Usually, coverage directly reflects the extent that the objectives are concerned, the concerned target node region having higher coverage, taking into account the functional relationship between sensor node p in the region Ⅱ expectations and coverage region, shown in Fig.1.
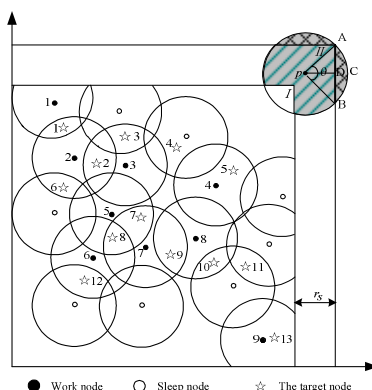

Fig.1 Diagram of connective coverage of target region and node *p* in Ⅱ region

Now take the analysis of Fig.1 as the example. The square region l is divided into Ⅰ and Ⅱ two parts. Randomly deploy the nodes in the monitoring region to construct a finite set S with the coverage region of each node being E(C), so that the coverage probability of each node is E(C)/Ω. When the nodes set is empty, the network coverage of the n deployed nodes will be P(S)=(1-E(C)/Ω)n Thus the network nodes probability value has been obtained in the situation that collective S is not an empty set.

$$P(S)=1-\left(1-E(C)/\Omega\right)^{n} \quad (5)$$

When the number of nodes $n \to \infty$, $\lim_{n \to \infty} E(P(S)) = 1$, which indicates that the number of nodes is large enough, this coverage region is completely covered[10]. Considering the boundary effect in solving the node coverage region and expectations, because the square region is divided into region I and II, based on the concept of expectation value in probability, expectation value of the coverage of network nodes can be obtained:

$$E(C) = P(\Omega_I)E(C_{\Omega_I}) + P(\Omega_{II})E(C_{\Omega_{II}}) \quad (6)$$

Among them, $P(\Omega_I)$ and $P(\Omega_{II})$, respectively represents the probability values of the randomly deployed nodes in region Ⅰ and Ⅱ. $E(C_{\Omega_I})$ and $E(C_{\Omega_{II}})$ denote the corresponding expectations of coverage. From the even distribution function of the random deployment of sensor nodes, then gets:

$$P(\Omega_I)=(l-r_s)^2/l^2, P(\Omega_{II})=2r_s(l-r_s)/l^2 \quad (7)$$

Assume that node p is in the region Ⅰ, its coverage being completely contained, so the coverage expectation is:

$$E(C_{\Omega_I}) = \pi r_s^2 \quad (8)$$

When a node p is in region II, the region should equal to its sensing region of the circumference subtracting bow region SACBD. A and B is the intersection of the sensing circle of node p and network boundaries, whose angle θ is the central angle formed by the node p and A, B. i.e. ∠ ApB=θ, $\theta = 2\arccos y/r_s$ so

$$E(C_{\Omega_{II}}) = 2\int_0^{l-r_s} \frac{1}{2}r_s^2(2\pi - \theta + \sin\theta)dx \int_0^{r_s} \frac{1}{2(l-r_s)r_s}dy$$

$$= \frac{r_s}{2(l-r_s)} \int_0^{l-r_s}(2\pi - \theta + \sin\theta)dx \int_0^{r_s} dy \quad (9)$$

HYBRID GENETIC ALGORITHM

Ants in the process of search are always hoped to meet the first select the evaluation function of the node.When the initial, all nodes of the pheromone is consistent, need more inspiration function to guide the ant.In basic TSP problem, the heuristic function is defined as the reciprocal of the distance between the two cities, in the search for characteristics of the node is not applicable.Due to the Filter class feature selection method, the mutual information between features and categories can better reflect the characteristics of the importance of inspiration function for that we will each feature nodes is defined as the characteristic and the category of mutual information, as shown below:

$$\eta_{fi} = I\left(f_i : C\right) \qquad (10)$$

Initial each ant randomly placed on the characteristics of a node, starting from the starting point a feature based on state transition rules of the selection.After the ants construct the solution, for all the characteristics of the solution of training classification respectively, get the classification error rate, and assess its advantages and disadvantages. Ant select features take according to the formula (10).

$$s = \begin{cases} \arg\max\{\tau(u)^{\alpha} * \eta(u)^{\beta}\}, & if\,(q < q_0),\ u \in allowed_k \\ p(s), \text{ otherwise, } & s \in allowed_k \end{cases} \qquad (11)$$

$$p(s) = \frac{[\tau(s)]^{\alpha} * [\eta(s)]^{\beta}}{\sum\limits_{u \in allowed_k} [\tau(u)]^{\alpha} * [\eta(u)]^{\beta}} \qquad (12)$$

Among them: alpha, beta, said the relative importance of the pheromone and the heuristic information;Refers to the current ant search all optional feature set;Said the ant the currently selected features;With the features of adjacency;Ants said the next step allows you to select all of the nodes, the list of records through the current ant nodes, when all nodes are written to the *tabu* table, ant to complete a cycle, the route is a solution of the problem by.Is evenly distributed in the interval [0, 1] is a random variable, the rules used to determine the relative importance of two parts, namely the ants selection can make the greatest feature of the pheromone and the heuristic information of the product and the most ants choose to have the characteristics of the transition probability of the two parts.To ensure ants use past information also have the ability to develop new search space at the same time, to prevent too fast convergence to local optimum.

After moving from node to node on the path of the pheromone update according to the formula (13)

$$\tau_{ij}\left(t+1\right) = \left(1-\xi\right)\tau_{ij}\left(t\right) + \xi\Delta\tau_{ij}\left(t\right) \qquad (13)$$

Among them, is the initial value of information?So can effectively avoid falling into local optimum, reduced has chosen a path to the selected probability again, so as to effectively avoid the ants converge to the same path, and improve the global search ability of the algorithm.

Ants search focuses in the current cycle so far to find the shortest path field area, global update rule is in after all the ants have completed their path to perform [13], update the pheromone formula is

$$\tau_{ij}\left(t+1\right) = \left(1-\xi\Box\tau_{ij}\left(t\right)\right)\Box\tau_{ij}\left(t\right) + \xi\Box\tau_{ij}\left(t\right)\Delta\tau_{ij}\left(t\right) \qquad (14)$$

$$\Delta\tau_{ij}^{k} = \begin{cases} \dfrac{Q}{L_{\min}} & 边\,(i,j) \in L_{\min} \\[2mm] -\dfrac{Q}{L_{\max}} & 边\,(i,j) \in L_{\max} \\[4mm] 0 & \text{otherwise} \end{cases} \qquad (15)$$

SIMULATED

Table.1 Network nodes

NO.	x-coordinate/m	Y-coordinate/m	Number	Time/h	Completed/h
1	13	26	13	(0.78,0.90,1.08)	2.5
2	24	59	20	(0.78,0.90,1.08)	3
3	11	58	3	(0.78,0.90,1.08)	4
4	17	46	5	(0.78,0.90,1.08)	2.5
5	22	60	25	(0.78,0.90,1.08)	4.5
6	78	81	9	(0.78,0.90,1.08)	3.5
7	60	89	14	(0.78,0.90,1.08)	4
8	30	92	19	(0.78,0.90,1.08)	4.5
9	12	8	20	(0.78,0.90,1.08)	3
10	68	5	18	(0.78,0.90,1.08)	4.5
11	78	17	7	(0.48,0.60,0.78)	1.5
12	15	6	11	(0.48,0.60,0.78)	2.5
13	39	25	12	(0.48,0.60,0.78)	1.5
14	45	33	21	(0.48,0.60,0.78)	3
15	49	58	24	(0.48,0.60,0.78)	4.5
16	56	44	15	(0.48,0.60,0.78)	6
17	59	30	14	(0.48,0.60,0.78)	6.5
18	95	61	22	(0.48,0.60,0.78)	2.5
19	92	67	21	(0.48,0.60,0.78)	3.5
20	83	68	4	(0.48,0.60,0.78)	4.5

The algorithm parameters are as follows: $num = 1000$; $MAXGEN = 500$; $\delta_1 = 0.5$; $\delta_2 = 1.6$; $\alpha = 0.7$; $\beta = 0.9$; $\omega = 0.6$; $q = 15$; **Crossover rate is 0.9; mutation rate is 0.05.**
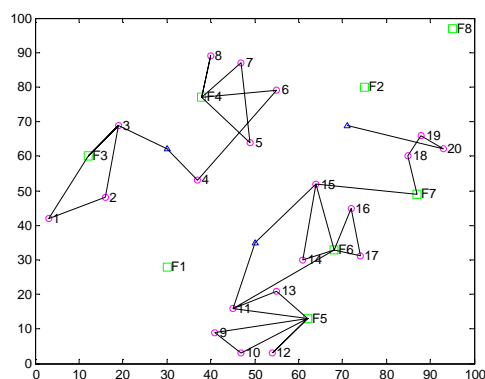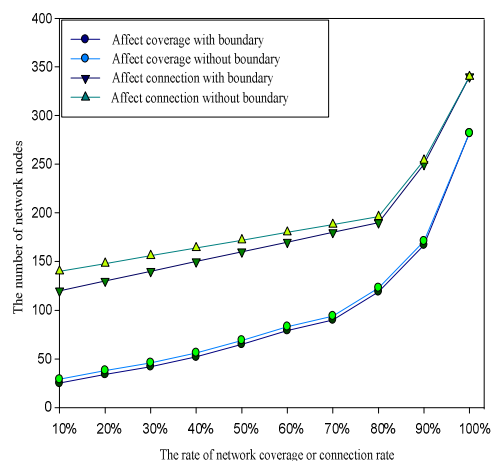

Fig. 2 Calculation Results

Fig.3 Curves of network coverage rate/ connectivity rate with/ without boundary influence

Is a pheromone volatilization coefficient, such as to do some change, can prevent the increase of the pheromone concentration on the path of the infinite, can reduce the possibility of trapped in local optimum, and thus accelerate the convergence speed.Ants are through the shortest and the longest path.In the early stages of the algorithm, the ant walk through every path has a chance to get the amount of information updates, ant chooses the path of the corresponding probability can be improved.With certain path pheromone accumulation, reach a certain number in

cycles, reduced the optimal path with the worst side of the path between the amounts of information gap, and so that the ant search behavior has focused on the optimal path.

Fig.6reflects the number of nodes required to be deployed to achieve different coverage and connectivity rates without the boundary influence. Compared with the boundary influence, the number of nodes deployed increases slightly, and with the increase of nodes, node density will become larger, so the boundary influence becomes lower.

CONCLUSION

Improvement, based on the heuristic function, information update strategy, crossover, and selection strategy to optimize the solution value, to a certain extent, inhibit the phenomenon of premature convergence and improve the global searching ability, to avoid falling into local optimal solution effectively, accelerate the convergence speed,In viewofthe network coverage and connectivity of the random deployment of nodes for wireless sensor network, the relation model of the sensor node and the target node is provided. Considering the boundary influence, coverage and connectivity probability model is provided. It can simplify the computational complexity of network coverage and connectivity to improve the efficiency of the algorithm execution, and thus solve the number of nodes required to deploy more accuratelymeeting certain network coverage and connectivity rate requirements.Finally, the correctness of the theoretical solution and validity of the algorithm are verified by the simulation results. The main work thereafter is the study ofthe effective coverage of the irregular region and the entire network energy savingsbased on sensor network.

REFERENCES

[1] CaiGuoqiang, JIA Li min, XING Zongyi. *Fuzzy Systems and Mathematics*, **2008**, 4(22):87-98.
[2] Feng Zuhong, XU Zong ben. *Journal of Engineering Mathematics*, **2002**, 4(19):35-39.
[3] Dong D, Liao X. *IEEE Transactons on Computers*, **2011**, 61(10):1417-1428.
[4] Liu Yin, MA Liang. Fuzzy Ant Colony algorithm and its Application in TSP. Mathematics in Practice and Theory,**2011**,6(41):150-153.
[5] Ke W, Liu B. *IEEE Transactions on Wireless Communications*, **2011**, 10(4):1154-1164.
[6] Slijepcevic S, Potkonjak M. Power efficient organization of wireless sensor networks. *IEEE Conference on Communication,* **2001**, 6(2):472-476.
[7] Carbunar B, Grama A, Vitek J. *ACM Transactions on Sensor Network,* **2006**, 2(1):94-128.
[8] Mhatre V, Rosenberg C. *IEEE Transaction on Mobile Computing*, **2004**, 3(3):27-31.
[9] Ruixin N, Pramod K, Varshney Q. *Information Fusion*, **2006**, 7(4):380-384.
[10] Jin Y, Jo J. *Computer Communications*,**2008**,31(10):2398-2407.
[11] Jinhua Z, Zhongzhi S, Yong XA. *Journal of Computer Research and Development*, **2004**,41(7):1081-1087.
[12] Sun Z, Yu Y. *International Journal of Applied Mathematics and Statistics*,**2013**,50(20):467-75.
[13] Ammari H, Das S. *IEEE Transactions on Computers*, **2012**, 61(1):118-133.
[14] L. Liang and M. Huadong, *IEEE Transactions on Parallel and Distributed Systems*,**2012**,2(8):118-124.
[15] Y. Yonrim and K. Y. Hyuk, *IEEE Transactions on Cybernetics*, **2013**, 10(5):1473-1779.
[16] L. Liang, Z. Xi and M. Huang, , *IEEE Sensors Journal*, **2013**, 12(10):3625-3631.
[17] H. Chihfan and L. Ningyan, *IEEE Transactions on Wireless Communications*, **2006**, 5(11):3182-3188.
[18] B. Jacques and M. Abdallah, M. Ahmed, *Computer and Communications*, **2008**, 7(5):770-776.
[19] M. Habib and K. Sajal, *IEEE Transactions on Parallel and Distributed Systems*, **2009**, 11(6):872-881.
[20] D. Tian and N. Georganas, *Journal of Wireless Communications and Mobile Computing*,**2003**,2(2):271-278.
[21] C. Hang, Y. Tseng and H. Wu, *ACM Transactions on Sensor network*,**2007**,1(1):137-145.
[22] S. Shakkottai, R. Srikant and N. Shroff, Unreliable sensor grids: Coverage, Connectivity and Diameter, Ad Hoc Network, **2005**, 7(4):702-716.
[23] P. Zhou, X. Cui and S. Wang, *Journal of system simulation*,**2009**,9(10):1416-1422.