INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 19, Issue: 3, Month: June, Year: 2024
Article Number: 5382, https://doi.org/10.15837/ijccc.2024.3.5382

CCC Publications

AGORA UNIVERSITY PRESS

An improved ant colony algorithm based on Q-Learning for route planning of autonomous vehicle

Liping Zhao*, Feng Li*, Dongye Sun, Zihan Zhao

Liping Zhao*
Systems Engineering Institute, Academy of Military Sciences, People's Liberation Army
No 2, Fengti South Road, Fengtai District, Beijing, China
825889797@qq.com

Feng Li*
Systems Engineering Institute, Academy of Military Sciences, People's Liberation Army
No 2, Fengti South Road, Fengtai District, Beijing, China
li_bopr@126.com

Dongye Sun*
National Engineering Research Center for Transportation Safety and Emergency informatics, China Transport Telecommunications & Information Center
No.1 Anwai Waiguan Houshen, Beijing, China
14114218@bjtu.edu.cn

Zihan Zhao
Systems Engineering Institute, Academy of Military Sciences, People's Liberation Army
No 2, Fengti South Road, Fengtai District, Beijing, China
zzh_track@163.com

Abstract

In view of the problems existing in the path planning algorithms of unmanned vehicles, such as low search efficiency, slow convergence speed and easy to fall into the local optimal. Based on the characteristics of route planning for unmanned vehicles, this paper introduces Q-Learning into the traditional ant colony algorithm to enhance the learning ability of the algorithm in dynamic environment, so as to improve the overall efficiency of route search. By mapping pheromones into Q values in Q-learning, rapid search in complex environments is realized, and a collection-free path satisfying constraints is quickly found. The results of case analysis show that compared with the traditional ant colony algorithm and the improved ant colony algorithm considering reward and punishment factors, the improved ant colony algorithm based on Q-Learning can effectively reduce the number of iterations, shorten the path optimization time and path length and other performance indicators, and has many advantages in jumping out of the local optimal, improving the global search ability and improving the convergence speed, and has good adaptability and robustness in complex environments. It ensures the safety and stability of unmanned vehicles in complex environments.

Keywords: Autonomous vehicle, Path planning, Q-Learning, Improved ant colony algorithm.

1 Introduction

Key issues in autonomous vehicle research include positioning, perception, planning, decision-making, and control. On the basis of map information, environment information and intelligent vehicle state information, the planning mainly uses artificial intelligence technology, probability theory method and game theory method to complete global path planning and local trajectory planning. Path planning is the bridge between information perception and intelligent control of unmanned vehicles, and it is one of the decisive technologies in the realization of unmanned vehicles. Its task is to search for a collision-free and optimal path to reach the target position safely according to certain evaluation criteria in the environment with obstacles according to certain path planning algorithm [1, 2]. It mainly involves two aspects: one is to model the road environment; The second is to select a suitable algorithm for global path planning.

Road environment modeling mainly adopts several kinds of methods: grid method, viewable method, free space method. The grid method divides the workspace into regular and uniform grids containing binary information. The grid method divides the workspace into regular and uniform grids containing binary information. Li Tiancheng et al. established a fan grid map model based on polar coordinate system, which can effectively solve the problem of circular wave propagation algorithm[3]. Guo Lijin et al. proposed an adaptive raster map creation algorithm based on quadtree to solve the problem that different raster sizes greatly affect the accuracy and effect[4]. Under the condition of the same accuracy, the algorithm can effectively reduce the amount of data storage and computation, and improve the real-time performance. Compared with the viewable method, the free space method and the raster method, they are more flexible and consume fewer computing resources. However, the disadvantages of this method are also obvious. The number of obstacles is proportional to the complexity of the algorithm. If the algorithm is too complex, its reliability will be reduced.

Path planning algorithms can be classified into two categories, one is the conventional method, the other is the reinforcement learning method. Conventional methods can be divided into graph search method (such as Dijkstra, and A*), random sampling method (such as probabilistic road map, and fast search random tree RRT), optimization-based method (such as model prediction planning method), artificial potential field method and intelligent bionic algorithm (such as particle swarm optimization algorithm, ant colony algorithm, genetic algorithm, simulated annealing algorithm, and artificial neural network)[5–11]. Bacha et al. adopted Dijkstra algorithm to plan the global path of intelligent vehicles or the local path with fewer scene nodes, but the planning efficiency of the algorithm was low[12]. Some scholars use the improved A* algorithm considering the time factor to solve the problem of collision-free path planning of automatic guided vehicles, which can effectively reduce the turning problem during driving, but its spatial complexity is exponential and not suitable for large-scale complex environment[13]. At the same time, some scholars use Rapidly exploring Random Trees (RRT*) algorithm to search two-dimensional AUV paths online, but the randomness is strong, so it is often difficult to get the global optimal path[14].

Jalalmaab et al. used the model-based prediction method to complete the overtaking trajectory planning of intelligent vehicles under highway conditions and maintain the vehicle's driving stability[15]. However, since the optimization-based method does not have the ability to search, it is difficult to apply this method in complex scenes with many pedestrians or vehicles. Since artificial potential field method cannot deal with vehicle dynamics constraints, this method is often used in conjunction with optimization methods such as model prediction method[16].

In recent years, some scholars have effectively solved some technical problems in the field of transportation by using intelligent methods such as machine learning and reinforcement learning[22–24]. Miao et al. uses ant colony algorithm to carry out multi-objective optimization, comprehensively considering path degree, safety and fuel consumption, and uses adaptive improved ant colony algorithm to obtain an optimal multi-objective path[18]. Xu Ling et al. developed a new way to extend the ant search method to 16 directions and 24 adjacent domains to expand the search scope and improve the path optimization effect and search efficiency of ant colony algorithm[19]. Li Tao et al. adopted the ant colony algorithm based on Darwin's theory of evolution to solve the problem that the traditional ant colony algorithm blindly searches in the blank grid, speeding up the iteration speed of the algorithm and shortening the running time[20].Ant colony algorithm is first proposed by Italian scholar Maro Dorigo, is a distributed bionic algorithm, itself has strong robustness, but there is also slow convergence, easy to fall into the local optimal problems[17].Q learning algorithm is a common reinforcement learning method in path planning problems[25]. In the general Q learning algorithm, it is easy for the agent to repeatedly explore the suboptimal path in the iterative process, and fall into the local optimal, resulting in slow convergence of the algorithm. The traditional path planning algorithm and single algorithm have some limitations in practical application, so the traditional algorithm can be combined with reinforcement learning method to optimize. There are a lot of improvement methods for ant colony algorithm, but the ant colony algorithm itself is easy to fall into local optimal and deadlock two problems have not been completely solved because of these improved methods, Q learning algorithm does not exist deadlock problem, local optimal problem is lighter than ant colony algorithm, so this paper chooses Q learning as the research object[26].

LI et al. introduced a heuristic search strategy into the improved Q-Learning algorithm to speed up the learning process and narrow the search space by limiting the variation range of direction angles[21].MEERZA et al. proposed a path planning algorithm based on Q-learning and particle swarm optimization. PSO was used to improve the iteration of Q table, which had better performance in speed and precision than using these two algorithms alone[27]. YAO et al. based on Q-Learning algorithm, combined with artificial potential field method, took black hole potential field as the environment, so that the robot could jump out of the local optimal solution without prior knowledge[28]. LIU et al. combined RRT with Q-Learning algorithm and proposed a partition heuristic RRT algorithm based on Q-Learning. Q-Learning was used to improve the reward function and obtain the global optimal path, which could obtain smoother results and improve the ability of searching and avoiding obstacles[29]. SHI et al. fused Q-learning algorithm with pheromone mechanism in ant colony algorithm, and exchanged information between robots through pheromones to solve the problem of information sharing in multi-agent path planning. Under the effect of Q value, robots made status updates and decision choices[30].

For the complex working environment of the unmanned vehicle, considering that not only fixed obstacles but also moving obstacles should be taken into account when the unmanned vehicle is driving on the road, multiple turns of the traditional ant colony algorithm will increase the driving time of the unmanned vehicle, and increase the amount of calculation and storage, thus reducing the operating efficiency of the system. Q-Learning algorithm is easy to fall into local optimization, resulting in slow algorithm convergence. The problem to be solved in this paper is to search for an optimal or suboptimal path in this abstract model, that is, to search for a series of continuous non-collision points from the starting point to the end point. We call this process path search. According to the shortcomings of improved ant colony algorithm in path planning application in the above analysis, combined with the problems faced in path planning of unmanned vehicles, this paper proposes Q-learning ant colony algorithm, which maps pheromone into Q value in Q-learning, strengthens the Learning ability of the algorithm in dynamic environment, and improves the overall efficiency of path search. It can realize the fast search for the non-collision path that meets the constraint conditions in complex environment.

2 The method principle of path optimization

The basic principle of ant colony algorithm is an optimization algorithm with positive feedback mechanism. According to the definition of unmanned vehicle path planning and the characteristics of ant colony algorithm, the principle of unmanned vehicle path planning based on ant colony algorithm can be clearly defined. First of all, M ants are placed in the initial position, and with each ant as the node center, the next node is determined according to the amount of information on each path and the heuristic information on the path. Secondly, if ant i reaches the destination position first, it is proved that ant's path is the best in the optimization process of the epicycle. Therefore, global pheromone update is carried out for ant i's path. Then, starting from the end position, the starting position is taken as the target to continue searching for optimization. If the obtained new path is better than the current optimal path, it will be replaced and the global pheromone will be updated; otherwise, the original optimal path will be maintained. Finally, the rule is repeated until a set constraint and a specified number of iterations are met.

3 Modeling of unmanned vehicle path planning problem

In this paper, the path planning of unmanned vehicle is defined as global static path planning based on environment model. Therefore, path planning is mainly divided into environmental model building and path searching.

3.1 Road environment modeling of unmanned vehicles

According to the characteristics of the path planning problem of unmanned vehicle, the position structure information of all obstacles in the global path planning space can be obtained through detection in advance. Moreover, different types of obstacles on the road surface are almost on the same plane as the road surface, making the planning space approximate to a horizontal two-dimensional plane. At the same time, considering the simple realization of grid method, it can express regular obstacles clearly, especially easy to combine with other intelligent algorithms. Therefore, this paper adopts the grid method to establish the environment model.

3.1.1 Create environment array

Setting up environment array, first use cartesian coordinate method to identify the planning space. Take the upper left corner of the grid model as the origin, set the horizontal to the right as the horizontal axis forward, and the vertical to the vertical axis forward. Each grid position corresponds to a unit length on the coordinate axis. After this processing is completed, each grid has a unique corresponding rectangular coordinate (x, y).

When an unmanned vehicle meets an obstacle during path planning, necessary measures must be taken, that is, to avoid the obstacle in the horizontal direction. Therefore, the obstacle area is defined in this paper as the area where unmanned vehicles must avoid in the horizontal direction. It is assumed that the path planning space of unmanned vehicle is a grid plane $W = H2$, that is, a grid of $20 * 20$. In the planning space W, all grids are divided into two categories: free zones that unmanned vehicles can pass through smoothly and obstacle zones that cannot pass through. These grids are represented by the following formula:

$$f : W \to \{0, 1\} \qquad \begin{cases} f(x) = 0, x \in free\ zone \\ f(x) = 1, x \in obstacle\ zone \end{cases} \tag{1}$$

According to the above formula, the whole planning space W is divided into small areas marked 0 or 1. Then W can be represented by a two-dimensional array, called $G[\ ][\ ]$. $G[w][h]$ represents the grid with row coordinate W and column coordinate h in the planning space w. If $G[w][h]$ is 0, it is a free region, represented by white; If $G[w][h]$ is 1, it is the obstacle area, shown in black.

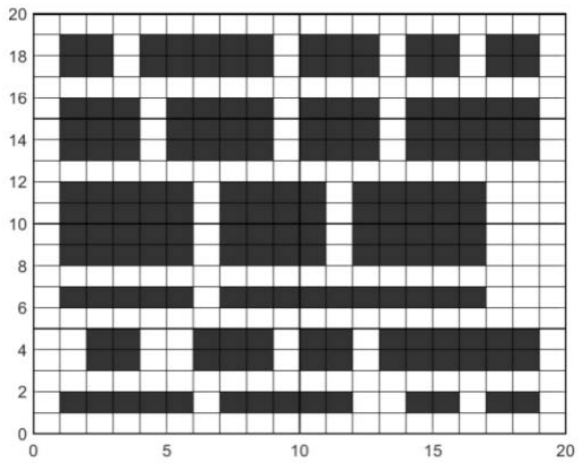

Figure 1: Raster model of unmanned vehicle path planning map

3.1.2 Obstacle and boundary processing in environmental model

After the grid size is determined, connect the beginning and end of the path planning. Make the line rotate 45 degrees (counterclockwise) as the horizontal axis, and the line perpendicular to the horizontal axis is defined as the vertical axis. Then, according to the determined grid size, grid discretization is carried out on the planning space. In the discretization of the planning space, there are several aspects that need to be paid attention to when dealing with obstacles. Firstly, the boundary of the environmental model is treated as an obstacle. Secondly, when the size of the obstacle is not enough for a grid, it will be filled into a grid. Thirdly, if there is a concave part of the obstacle, this concave is also treated as an obstacle to prevent local dead zone in path search.

3.2 Application of improved ant colony algorithm in route search of unmanned vehicle

3.2.1 Path planning steps of unmanned vehicle based on traditional ant colony algorithm

Step 1: Parameter initialization.

The raster environment information of unmanned vehicle is represented by matrix composed of 0 and 1. Where, 0 means that no obstacles can pass through, 1 means that the grid is occupied by obstacles and cannot pass through. Suppose the number of ants is M, the pheromone importance factor a, the importance factor b of the heuristic function, and the number of cycles N_c.

Step 2: Transition probability calculation.

At the initial time, the amount of information on each path is equal, denoted as $\tau_{ij}(0) = C$. In the process of moving, ant k determines the next node to be transferred according to the path information and path heuristic information. $P_{ij}^k(t)$ represents the transfer probability of ant k from grid i to adjacent grid j at time t. The specific formula is as follows:

$$P_{ij}^k(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^a \cdot [\eta_{ij}(t)]^b}{\sum_{s \in A_k} [\tau_{is}(t)]^a \cdot [\eta_{is}(t)]^b} & , j \in A_k \\ 0 & , j \notin A_k \end{cases} \tag{2}$$

$$\eta_{ij}(t) = \frac{1}{d_{ij}} \tag{3}$$

Where, $A_k = \{0, 1, 2, \ldots\ldots, n-1\} - Ban_k$ represents the grid that ant k is allowed to access next. Different from the actual ant colony, the ant colony system in the system has memory function. The set Ban_k is the tabu set, which is used to record the grid nodes that ant k currently walks through, and the Ban_k is constantly adjusted along with the search process of ants. η_{ij} is a heuristic function, representing the expected degree of ant k moving from position i to position j. d_{ij} is the geometric distance between node i and node j.

Step 3: Update the ant tabu table set Bank.

With each step taken by ant k, node j is added to the tabu set Bank.

Step 4: Calculated path length.

Repeat steps 2 and 3 until all ants within the loop reach the end point and calculate the length of the path for each ant to reach the end point.

Step 5: Pheromone renewal.

Over time, the pheromones left on the path gradually disappear, and the remaining pheromones are treated after each ant takes a step or at the end of a cycle. The $t+n$ pheromone update rules are as follows:

$$\tau_{ij}(t+n) = \rho \cdot \tau_{ij}(t) + \Delta\tau_{ij}(t) \tag{4}$$

$$\Delta\tau_{ij}(t) = \sum_{m} \tau_{ij}^{k}(t) \tag{5}$$

Where, ρ is the pheromone residue factor. To avoid the infinite accumulation of pheromones, its value range $\rho \in (0,1)$; $\Delta\tau_{ij}(t)$ is the pheromone increment on path (i,j) in this loop, The initial time $\Delta\tau_{ij}(t) = 0$; $\tau_{ij}^{k}(t)$ is the pheromone increment of the ant k on the path of this cycle (i,j).

$$\Delta\tau_{ij}(t) = \begin{cases} \dfrac{Q}{L_k}, & Ant\ K\ passes\ through\ the\ path\ (i,j) \\ 0, & Ant\ K\ doesn't\ go\ through\ the\ path\ (i,j) \end{cases} \tag{6}$$

Where, Q is the pheromone intensity, which mainly affects the convergence rate of the algorithm; L_k is the total length of the path taken by the ant k in this cycle.

Step 6: Ends the loop and prints the result.

If the end condition is met, the loop is terminated and the program calculation results are output; otherwise, all tabu tables are cleared and go to step 2.

3.2.2 Improved ant colony algorithm based on Q-Learning

In order to improve the search efficiency and path planning quality of ant colony algorithm in unmanned vehicle path planning, conventional ant colony algorithm was improved in this study[31–33].

(1) Improved state transition rules

Although the existing state transition rules can effectively ensure the diversity of paths, the convergence rate of the algorithm is slow because the difference between the weighted product of the pheromone and the inspired information at each feasible point is not that large. Therefore, the state transition rule is improved, as shown in Equation 7:

$$\tau_j(t) = \begin{cases} argmax_{s\in A_k}\{\tau_s(t)^a \cdot \eta(\tau_{is}(t))^b\} & ,\ q \le q_0 \\ J & ,\ q > q_0 \end{cases} \tag{7}$$

Where, $\tau_j(t)$ represents the next accessed waypoint; J represents the next waypoint selected by probability according to Formula (2); q_0 is a given threshold between $(0,1)$ to adjust the algorithm's ability to explore new paths.

After the improvement, the algorithm has a greater chance to choose the point with the largest weighted product of pheromone and heuristic information as the next path point, which speeds up the convergence rate of the algorithm. Through experimental verification, under the threshold value set in this paper, the algorithm's ability to explore new paths is not hindered, the diversity of paths is still guaranteed, and the algorithm will not easily fall into local optimal.

(2) The global pheromone update rule is modified

Q-Learning is a method for solving reinforcement learning problems using time series difference without state transformation model[34, 35]. In this paper, Q-Learning is introduced to strengthen the learning ability of the algorithm in dynamic environment. Q-Learning behavior value function update formula is as follows:

$$Q_{t+1}(s,a) = (1-\chi)Q_t(s,a) + \chi(\gamma_{t+1} + r \cdot maxQ(s_{t+1},a)) \tag{8}$$

Where, use $Q(s,a)$ as the cumulative return value of action a; χ is the learning rate, $\chi \in (0,1)$; r is the attenuation coefficient; γ_{t+1} is the income obtained by selecting action a at time $t+1$ according to the environment state s.

The pheromone concentration of ant colony algorithm was mapped to Q value of Q-learning. The pheromone updating rules in the ant colony algorithm are as follows:

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \rho(\Delta\tau_{ij}(t) + r \cdot max(\tau_{ij}(t))) \tag{9}$$

Where, $max(\tau_{ij}(t))$ is the maximum pheromone concentration at time t;$\Delta\tau_{ij}(t)$ is the global pheromone increment to ant from time t to time $t+1$.

(3) Algorithm step

The improved ant colony algorithm based on Q-Learning proposed in this paper is shown in Figure 2. The specific steps are as follows:

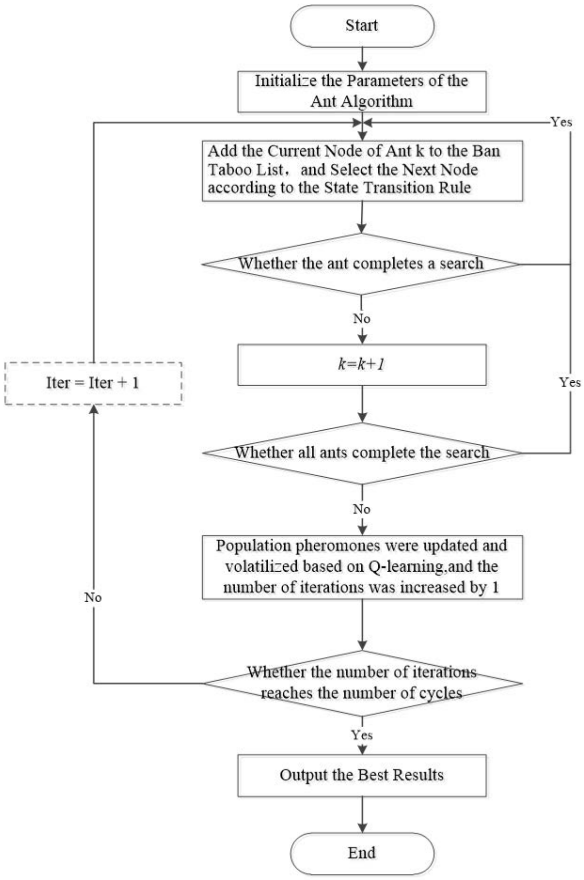

Figure 2: Flow chart of improved ant colony algorithm based on Q-Learning

Firstly, the raster map environment is modeled, and the ant colony algorithm parameters such as the number of ants and the number of algorithm iterations are initialized.

Secondly, set and initialize the tabu table of each forward ant. The tabu table of each forward ant is used to store the nodes that have been visited so that the forward ant mark $k = k + 1$.

Then, within the range of nodes accessible to each forward ant, the next hop neighbor node j is selected according to the improved state transition rule calculation formula. And the loop iterates until all forward ants reach the destination node and convert to backward ants.

Finally, the backward ant k is updated according to the improved global pheromone updating rule, which strengthens the global search ability of the algorithm, and repeats until the end of the cycle.

4 Experimental simulation and result analysis

The effectiveness of the improved ant colony algorithm is verified by simulation analysis. Firstly, the environment is modeled by using raster method, and two kinds of raster maps are designed, which are normal environment and complex environment respectively. In these two scenarios, the effects of traditional ant colony algorithm, improved ant colony algorithm and improved ant colony algorithm based on Q-Learning are compared and analyzed. The running environment of the simulation experiment was built by MatlabR2021a, and the processor was Intel(R) Core(TM) i7-4710MQ CPU @2.50GHz.

4.1 Analysis of simulation results under normal environment

The path planning method of the improved ant colony algorithm is verified by MATLAB. Using raster method to conduct modeling in a 20×20 environment, simulation parameters are as follows: the number of ants is 30, the number of cycles is 200, the pheromone importance factor a is 1, the importance factor b of the heuristic function b is 7, the pheromone residue factor ρ is 0.5. In an improved ant colony algorithm based on Q-Learning, a given threshold q_0 is 0.1, the learning rate χ is 0.9, the attenuation coefficient r is 1.

It is difficult to show the advantage of improved ant colony algorithm based on Q-Learning by using multi-index algorithm because of the limitations of raster maps. Therefore, the length of the optimal path is used as the fitness function index for path planning, and by comparison, we verify whether the algorithm proposed in this paper has improved the global search ability, convergence time, iteration times and jumping out of the local optimal solution. Through simulation, different algorithms all found an optimal path from the starting point to the end point in the same grid environment, specific path planning is shown in Figure 3 to Figure 5. Figure 3 shows the simulation path using traditional ant colony algorithm, the path falls into local optimal, and the path length is 48.1424. Compared with the traditional ant colony algorithm, the improved ant colony algorithm adaptively selects the parameters, the ability to jump out of the local optimal is enhanced, but the global search ability is still slightly insufficient, and the path length is 37.3137. Finally, the improved ant colony algorithm based on Q-Learning has a great improvement in jumping out of local optimal and global search, the path length is 32.7279.

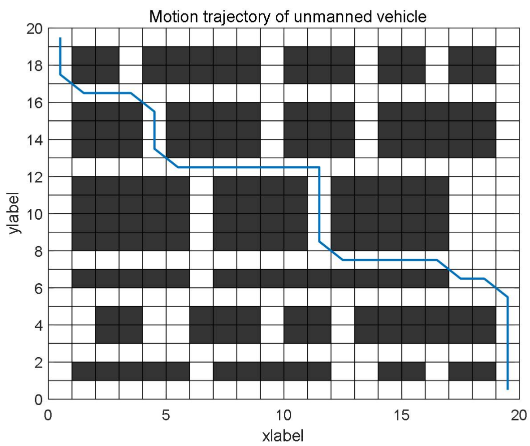

Figure 3: the path using traditional ant colony algorithm

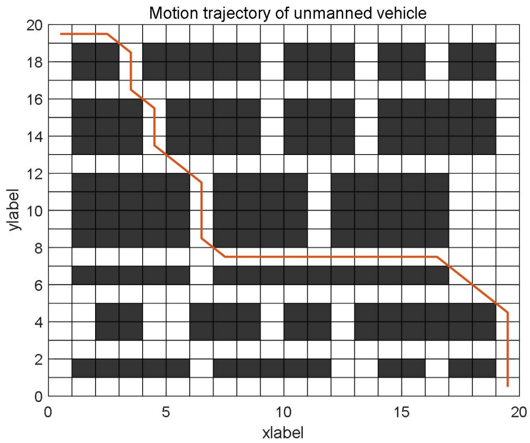

Figure 4: the path using improved ant colony algorithm

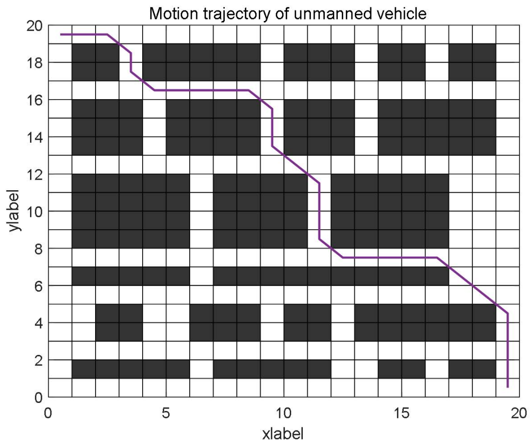

Figure 5: the path using improved ant colony algorithm based on Q learning

Figure 6 shows the iteration curves of the three algorithms, it can be seen that the early curve of the improved ant colony algorithm based on Q-Learning can quickly approximate the optimal solution, and it is significantly better than the previous two algorithms. This is because the algorithm proposed in this paper maps pheromone concentration to Q value of Q-learning in the process of global pheromone update, the search times and time in the early stage are greatly reduced, and then its iteration curve can approach the optimal solution at a very fast speed without fluctuation. The minimum path length and system running time under this algorithm are significantly reduced compared with the other two algorithms.

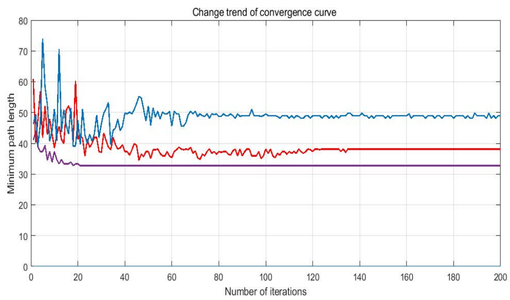

Figure 6: Convergence curve trend diagram of different algorithms

Table 1: Performance comparison of different algorithms in 20 × 20 grid environment

Parameters	Minimum path length/m	System uptime/s	Optimal number of iterations/times
traditional ant colony algorithm	48.1424	25.1632	119
improved ant colony algorithm	37.3137	10.5483	71
improved ant colony algorithm based on Q learning	32.7279	7.2526	18

4.2 Analysis of simulation results under complex environment

To verify the applicability of the proposed algorithm in different environments, raster maps are complicated and 40 × 40 raster maps are established. To ensure the effectiveness and reliability of the algorithm, the number of ants is changed 50. Through simulation, different algorithms all find an optimal path from the starting point to the end point in the same grid environment, the specific path is shown in Figure 7 to Figure 9.

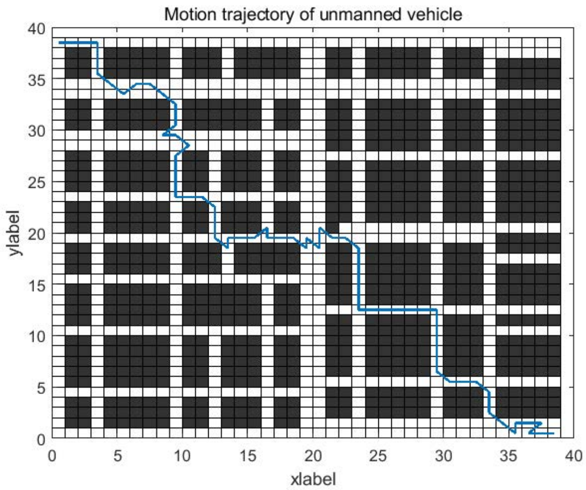

Figure 7: the path using traditional ant colony algorithm

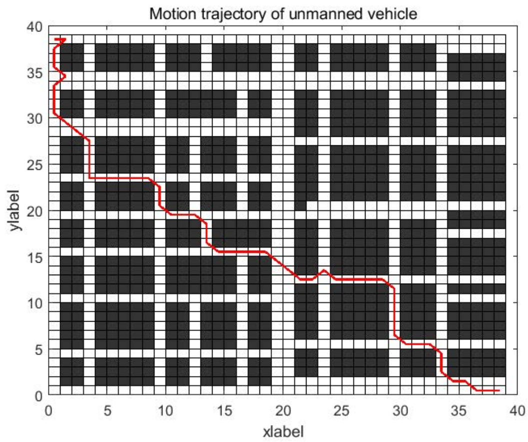

Figure 8: the path using improved ant colony algorithm

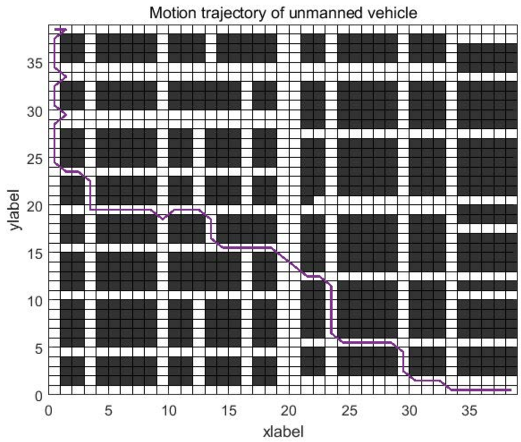

Figure 9: the path using improved ant colony algorithm based on Q learning

The path length of traditional ant colony algorithm is 91.3553m, the path length of improved ant colony algorithm is 87.1127 m, and the path length of the proposed algorithm is 75.6985m. As can be seen from Figure 10,

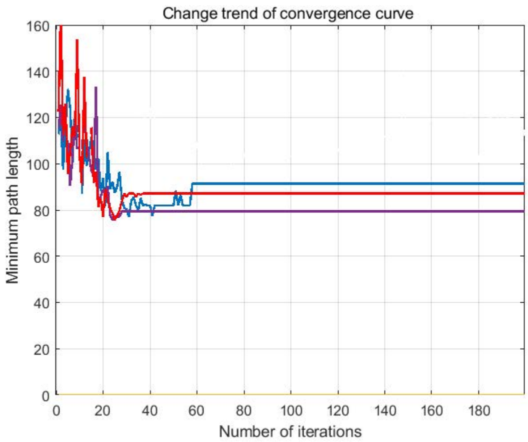

Figure 10: Convergence curve trend diagram of different algorithms

the number of iterations of the algorithm in this paper is only 41.38% of that of the traditional ant colony algorithm, and 61.54% of that of the improved ant colony algorithm. Compared with the traditional ant colony algorithm, the path length and system running time are reduced by 17.14% and 61.81% respectively.

In the case of complex map, the number of iterations of the traditional ant colony algorithm is 58, the convergence speed is too slow and the jumping ability is weak after falling into the local optimal. The minimum convergence times of the improved ant colony algorithm is 39 times. It can be seen that the algorithm has a great improvement in accelerating the convergence rate, but it needs to be improved for jumping out of the local optimal case. The algorithm in this paper converges at least 24 times. Compared with the above algorithm, although the path is the shortest and the ability to jump out of the local optimal is greatly improved, it still needs to be further strengthened to accelerate the convergence speed.

Table 2: Performance comparison of different algorithms in 20×20 grid environment

Parameters	Minimum path length/m	System uptime/s	Optimal number of iterations/times
traditional ant colony algorithm	91.3553	125.4772	58
improved ant colony algorithm	87.1127	90.3726	39
improved ant colony algorithm based on Q learning	75.6985	47.9216	24

5 Conclusion and future work

According to the specific characteristics of the path planning of unmanned vehicle, the grid method is adopted to conduct environment modeling for the planning space, and the ant colony algorithm is applied to the environment model to set the path selection probability of ants, pheromone updating rules, etc. And the convergence rate of the algorithm is accelerated by mapping pheromone to Q value. Finally, MATLAB is used to verify the effectiveness and feasibility of the method in multiple simulation environments with different scales and characteristics. The method in this paper can effectively reduce the number of iterations of the algorithm and find the optimal solution, but it cannot guarantee the efficiency of the operation. At present, the unmanned path planning of static obstacles in known environments is relatively mature, while the path planning of dynamic obstacles in unknown environments is the focus of research. In the future, with the development of artificial intelligence, Internet of Things and other technologies, model-free adaptive dynamic planning based on data and event-driven will be the focus of future research.

6 Funding

This research was funded by National Natural Science Foundation of China (52302423).

References

[1] Katrakazas C; Quddus M; Chen W H; et al. (2015). Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions, *Transportation Research Part C*, 60, 416–442, 2015.

[2] Patle B; Pandey A; Parhi D; et al. (2019). A review: On path planning strategies for navigation of mobile robot, *Defence Technology*, 15(4), 582–606, 2019.

[3] LI T C; SUN S D; GAO Y. (2010). Fan-shaped Grid Based Global Path Planning for Mobile Robot, *ROBOT*, 32(4), 547–552, 2010.

[4] GUO L J; SHI W X; LI Y; LI F X; et al. (2011). Mapping algorithm using adaptive size of occupancy grids based on quadtree, *Control and Decision*, 26(11), 1690–1694, 2011.

[5] Y Yang; K He; Y P Wang; Z Z Yuan; Y H Yin; M Z Guo. (2022). Identification of dynamic traffic crash risk for cross-area freeways based on statistical and machine learning methods, *Physica A: Statistical Mechanics and its Applications*, 595(2022), 127083, 2022.

[6] Azim E; Chaoxian W; Chuanyang S. (2021). Research Advances and Challenges of Autonomous and Connected Ground Vehicles, *IEEE Transactions on Intelligent Transportation Systems*, 22(2), 683–711, 2021.

[7] LI D L, WANG P, DU L. (2019). Path planning technologies for autonomous underwater vehicles-a review, *IEEE Access*, 7, 9745–9768, 2019.

[8] Özgur C; Sarikovanlik V. (2022). Forecasting BIST100 and NASDAQ Indices with Single and Hybrid Machine Learning Algorithms, *Economic Computation And Economic Cybernetics Studies And Research*, DOI: 10.24818/18423264/56.3.22.15, 56(3), 235–250, 2022.

[9] Y. Yang; N. Tian; Y. Wang; Z. Yuan. (2022). A Parallel FP-Growth Mining Algorithm with Load Balancing Constraints for Traffic Crash Data, *International Journal of Computers Communications & Control*, 17(4), 4806, 2022.

[10] Liu J.-Y.; Liu S.-F.; Gong D.-Q. (2021). Electric Vehicle Charging Station Layout Based on Particle Swarm Simulation, *Int. Journal of Simulation Modelling*, 20(4), 754–765, 2021.

[11] Yang Y; Yuan Z; Meng R. (2022). Exploring Traffic Crash Occurrence Mechanism toward Cross-Area Freeways via an Improved Data Mining Approach, *Journal of Transportation Engineering Part A Systems*, 148(9), 04022052, 2022.

[12] Bacha A; Bauman C; Faruque R; et al. (2008). Odin: Team Victor Tango's entry in the DARPA Urban Challenge, *Journal of Field Robotics*, 25(8), 467–92, 2008.

[13] Zhang X Y; Zou Y S. (2021). Collision-free path planning for automated guided vehicles based on improved A* algorithm, *Systems Engineering-Theory & Practice*, 41(1), 240–246, 2021.

[14] Carreras M; Hernandez J D; Vidal E; et al. (2016). Online motion planning for underwater inspection, *Autonomous Underwater Vehicles. IEEE*, 336–341, 2016.

[15] Jalalmaab M; Fidan B; Jeon S; et al. (2015). Model predictive path planning with time-varying safety constraints for highway autonomous driving, *International Conference on Advanced Robotics (ICAR)*, 213–217, 2015.

[16] Receveur J-B; Victor S; Melchior P. (2020). Autonomous car decision making and trajectory tracking based on genetic algorithms and fractional potential fields, *Intelligent Service Robotics*, 13(2), 315–330, 2020.

[17] Afify, H.M.; Mohammed, K.K.; Hassanien, A.E. (2020). Multi-Images Recognition of Breast Cancer Histopathological via Probabilistic Neural Network Approach, *Journal of System and Management Sciences*, DOI: https://doi.org/10.33168/JSMS.2020.0204, 10(2), 53–68, 2020.

[18] Miao C W; Chen G Z; Yan C L; et al. (2021). Path planning optimization of indoor mobile robot based on adaptive ant colony algorithm, *Computers & Industrial Engineering*, 156(1), 1–12, 2021.

[19] XU L; FU W H; JIANG W H; LI Z T. (2021). mobile robots path planning based on 16-directions 24-neighborhoods improved ant colony algorithm, *Control and Decision*, 36(05), 1137–1146, 2021.

[20] LI T; ZHAO H S. (2022). Path optimization for mobile robot based on evolutionary ant colony algorithm, *Control and Decision*, DOI:10.13195/j.kzyjc.2021.1324, 1–9, 2022.

[21] LI S D; XU X; ZUO L. (2015). Dynamic path planning of a mobile robot with improved Q-learning algorithm, In *Proceedings of 2015 IEEE International Conference on Information and Automation*, 409–414, 2015.

[22] Yang Y; Yuan Z; Chen J; Guo M. (2017). Assessment of osculating value method based on entropy weight to transportation energy conservation and emission reduction, *Environmental Engineering & Management Journal*, 16(10), 2413–2424, 2017.

[23] Yang Y; Yang B; Yuan Z; et al. (2023). Modeling and Comparing Two Modes of Sharing Parking Spots at Residential Area: Real-time and Fixed-time Allocation, *IET Intelligent Transport Systems*, 2023.

[24] Yuan Z; Yuan X; Yang Y; et al. (2023). Greenhouse Gas Emission Analysis and Measurement for Urban Rail Transit: A Review of Research Progress and Prospects, *Digital Transportation and Safety*, 1(1), 37–52, 2023.

[25] Tan B; Peng Y Y; Lin J G. (2021). A local path planning method based on q-learning, In *International Conference on Signal Processing and Machine Learning*, 80–84, 2021.

[26] TIAN X H; HUO X; ZHOU D L; ZHAO H. (2022). Ant colony pheromone aided Q-learning path planning algorithm, *Control and Decision*, DOI: https://doi.org/10.13195/j.kzyjc.2022.0476, 2022.

[27] MEERZA S I A; ISLAM M; UZZAL M M. (2019). Q-learning based particle swarm optimization algorithm for optimal path planning of swarm of mobile robots, *Proceedings of 2019 International Conference on Advances in Science, Engineering and Robotics Technology*, 1–5, 2019.

[28] YAO Q F; ZHENG Z Y; QI L; et al. (2020). Path planning method with improved artificial potential field- a reinforcement learning perspective, *IEEE Access*, 8, 135513-135523, 2020.

[29] LIU Z Y; LAN F; YANG H B. (2019). Partition heuristic RRT algorithm of path planning based on Q-learning, *Proceedings of 2019 Advanced Information Technology, Electronic and Automation Control Conference*, 386–392, 2019.

[30] SHI Z G; TU J; ZHANG Q; et al. (2013). The improved Q-Learning algorithm based on pheromone mechanism for swarm robot system, *Proceedings of the 32nd Chinese Control Conference*, 6033-6038, 2013.

[31] Zhu J Y; GAO M T. (2021). AUV Path Planning Based on Particle Swarm Optimization and Improved Ant Colony Optimization, *Computer Engineering and Applications*, 57(06), 267–273, 2021.

[32] HU C Y; JIANG P; ZHOU G R. (2020). Application of improved ant colony algorithm in AGV path planning, *Computer Engineering and Applications*, 56(8), 270–278, 2020.

[33] MA Y N; GONG Y J; XIAO C F; et al. (2019). Path planning for autonomous underwater vehicles: an ant colony algorithm incorporating alarm pheromone, *IEEE Transactions on Vehicular Technology*, 68(1), 141–154, 2019.

[34] HE X L; JIANG H; SONG Y; et al. (2019). Routing selection with reinforcement learning for energy harvesting multi-hop CRN, *IEEE Access*, 7, 54435–54448, 2019.

[35] ARUNITA K; LOBIYAL D K. (2021). Q-learning based routing protocol to enhance network lifetime in WSNs, *International Journal of Computer Networks & Communications*, 13(2), 67–80, 2021.

Journal's webpage: http://univagora.ro/jour/index.php/ijccc/

C O P E

Member since 2012
JM08090

This journal is a member of, and subscribes to the principles of,
the Committee on Publication Ethics (COPE).
https://publicationethics.org/members/international-journal-computers-communications-and-control

Cite this paper as:

Liping Zhao.; Feng Li.; Dongye Sun.; Zihan Zhao. (2024). An improved ant colony algorithm based on Q-Learning for route planning of autonomous vehicle, *International Journal of Computers Communications* & *Control*, 19(3), 5382, 2024.

https://doi.org/10.15837/ijccc.2024.3.5382