Research on the Multi-Armed Bandit Algorithm in Path Planning for Autonomous Vehicles

Jingyu Li*

Department of Computer Science, Southwest University, 400715, Chongqing, China

Abstract. In the technological revolution of the 21st century, autonomous driving technology is rapidly changing transportation modes, and path planning, as a key component, relies heavily on advanced algorithm optimization. The Multi-Arm Bandit (MAB) algorithm may become an efficient decision optimization tool in autonomous driving path planning. Because it can continuously experiment, learn, and quickly determine the optimal strategy to maximize profits under resource constraints. When applied to autonomous driving, the MAB algorithm may be able to demonstrate its advantages. In complex traffic environments, it dynamically adjusts strategies to adapt to constantly changing road conditions, plans safe and efficient driving paths, and quickly responds to unexpected situations to ensure driving safety. Compared with other algorithms, the learning and adaptability of MAB algorithm makes it particularly suitable for the dynamics and unpredictability of real-world driving scenarios. However, the practical application of MAB algorithm in autonomous driving faces challenges, including accurately evaluating path efficiency, efficiently processing large amounts of traffic data, and ensuring the stability and reliability of the algorithm. Further in-depth research and exploration are crucial for fully utilizing the advantages of MAB algorithm in path planning and promoting the sustainable development and enhancement of autonomous driving technology.

1 Introduction

In the wave of technology development in the 21st century, autonomous driving technology for automobiles, as an important component of intelligent transportation systems, is gradually moving from laboratories to practical applications, leading to a profound change in future modes of transportation. In recent years, autonomous driving technology has undergone a rapid development stage from assisted driving to highly automated driving and even fully automated driving [1]. This technological revolution not only greatly improves the convenience and safety of driving, but also provides possibilities for achieving intelligent transportation management, optimizing urban spatial layout, and promoting green and low-carbon travel.

The importance of autonomous driving is self-evident. Path planning is one of the core technologies of autonomous vehicles, which involves how to select the optimal driving

* Corresponding author: 222023321102092@email.swu.edu.cn

route according to the current location, target location, traffic rules, and other factors in the complex and changing traffic environment. At present, automatic driving path planning technology has been widely used in L2 and above autonomous vehicles. These vehicles are able to perceive their surroundings in real time, plan their routes based on traffic conditions, and achieve autonomous navigation and control. Data shows that the penetration rate of L2 level new passenger vehicles in China will reach 47.3% in 2023, and has exceeded 50% from January to May 2024. This indicates that autonomous driving path planning technology is rapidly becoming popular and gradually becoming a standard feature for mid to high-end car models. However, the core problem of automatic driving path planning is how to select an appropriate path for autonomous vehicles on the road network according to the current state and goal. This requires path planning algorithms to comprehensively consider various factors such as safety, efficiency, and comfort, while processing large amounts of real-time traffic data to ensure stable and efficient driving of vehicles in complex environments [2]. It also requires path planning algorithms not only to be able to perceive and process massive traffic information in real time but also to accurately predict changes in future traffic conditions to make optimal driving decisions [3]. The study plans to explore the application of multi-arm computation in autonomous driving route planning in this study.

2 Related work

Although traditional path planning algorithms, such as the Dijkstra algorithm and A * algorithm, can meet the basic needs of autonomous vehicles to a certain extent, their flexibility, robustness, and efficiency still need to be improved in the face of complex traffic environment, dynamic obstacles and real-time traffic information [4]. These algorithms are often based on static traffic network models for planning, making it difficult to adapt to real-time changing traffic conditions. In view of the core problems and some limitations of the above-automated driving path planning, the multi-arm algorithm may be used as an effective decision-making optimization method to provide new ideas in the field of autonomous vehicle path planning. The multi-arm algorithm, also known as the MAB Problem, is a classic stochastic optimization problem. Its core idea is to maximize cumulative returns through continuous exploration and utilization under limited resource constraints. The core of the multi-arm algorithm lies in finding a balance between exploring new options and utilizing known best options. Exploration means trying out options that are not fully understood or may have higher potential returns, in order to gather more information about these options. Utilization is based on the currently known information, selecting the options that perform the best or have the highest expected returns. In practical situations, pure exploration may lead to missing out on known high-yield options, while pure exploitation may miss out on potential high-yield options due to a lack of understanding of new options. Therefore, the multi-arm algorithm needs to make a trade-off between the two to achieve maximum long-term benefits [5]. Additionally, the application fields of multi-arm algorithms are very wide. Such as online advertising, recommendation systems, portfolio optimization, and so on. In these applications, algorithms need to make quick decisions based on real-time data and environmental changes to maximize benefits or minimize regrets. This is the MAB algorithm, Although the MAB algorithm was not originally designed for autonomous driving path planning, its decision optimization properties can be cleverly applied to path planning problems in autonomous driving, especially in dealing with uncertainty, dynamic environments, and real-time decision-making. In my opinion, it may play a crucial role in some aspects of it.

MAB algorithm can dynamically adjust the driving strategy according to real-time traffic information, vehicle status, user preferences, and other factors in an uncertain traffic environment to achieve optimal path selection [6]. Specifically, the application of the dobby algorithm in path planning of autonomous vehicles can be reflected in the following aspects: First, it can realize real-time evaluation and selection of different paths by building a path selection model based on the dobby algorithm; second, it can utilize the exploration and utilization mechanism of the multi-arm algorithm to balance the relationship between exploring new paths and utilizing known optimal paths in path planning[7]; third, it can also combine advanced technologies such as deep learning and reinforcement learning to enhance the adaptive ability and decision-making efficiency of multi-arm algorithms in complex traffic environments. Through the application of the dobby algorithm, autonomous vehicles may be able to respond more intelligently to the complex and changing traffic environment and improve the safety and efficiency of driving [8].

3 Methodology

The path planning of autonomous driving is similar to that of some ordinary machinery, such as industrial robots, automated warehousing systems, etc. Both require planning one or more optimal or suboptimal paths from the starting point to the endpoint in a complex and ever-changing environment, based on specific goals and constraints. In the field of autonomous driving, path planning needs to consider factors such as vehicle speed, steering angle, road conditions, traffic rules, and coordination with other vehicles. This is similar to the joint angles, motion trajectories, obstacle avoidance, and other factors that ordinary machinery needs to consider when performing tasks. The idea of multi-arm computation is to allocate and schedule multiple tasks or objectives reasonably to optimize overall performance, which can also be applied to path planning in autonomous driving. For example, in the case of a large number of autonomous vehicles, multi-arm computation can be used to coordinate the driving paths of each vehicle, avoid congestion and collisions, and improve overall traffic efficiency. Therefore, the study will combine these partial datasets of ordinary machinery to complete the path planning for autonomous driving.

3.1 Data processing

This study applied the artificial field potential method, which was first proposed by Khatib and can be used for obstacle avoidance with robots. During the process of autonomous driving, the obstacle avoidance of vehicles can also be achieved using this method to achieve real-time obstacle avoidance capabilities, enabling safe and efficient navigation and trajectory planning in complex environments [9]. Using the improved repulsion field function, the distance p (x, xg) between the vehicle and the obstacle, and the adjustment coefficient n of the relative distance, where k represents the positive proportional gain factor of the repulsion potential field, X represents the position coordinate of the autonomous vehicle, X0 represents the position coordinate of the obstacle, Xg represents the position coordinate of the target point, ρrepresents the maximum influence distance of the repulsion potential field of the obstacle, andρ(X, Xo), ρ(x, xg) represent the relative distance between the vehicle and the obstacle and the target point, respectively. When the vehicle approaches an obstacle, due to some new adjustment factors [10], the corrected repulsion will approach 0, so the target point remains the global minimum. According to its requirements, the simultaneous repulsive potential field function is expressed as Formula 1.

$$\begin{cases} 0 \quad , \quad \rho(X, \ X_0) > \rho_0 \\ \frac{1}{2} K_{rep} \left(\frac{1}{\rho(X, \ X_0)} - \frac{1}{\rho_0} \right)^2 \ \rho^n(X, \ X_0), \ \rho(X, \ X_0) \le \rho_0 \end{cases} \tag{1}$$

In addition, the corrected repulsive force field should also be considered, which will gradually decrease with the distance p (X, Xg) between the autonomous vehicle and the target. However, in the previously cited literature on robots, this mainly has a significant impact on the robotic arm, and the impact on automobiles may not be considered. However, the Formula2 will still be expressed here.

$$F_{rep}\ (X) = \begin{cases} 0, & \rho(X, X_0) > \rho_0 \\ F_{rep1} + F_{rep2}, & \rho(X, X_0) \le \rho_0 \end{cases} \tag{2}$$

The path planning of autonomous driving can also be combined with other algorithms such as reinforcement learning, deep learning, etc. to further improve the accuracy and robustness of the planning. These algorithms can be continuously learned and optimized to make the auto drive system more adaptable to the complex and changing environment and task requirements.

3.2 Obstacle avoidance path planning

Obstacle avoidance path planning requires deep consideration of environmental perception, and obtaining obstacle information through sensors and modelling. Meanwhile, safety and stability are crucial to ensure a collision-free path and smooth driving. Algorithm selection needs to be based on environmental characteristics. Here, the artificial potential field method is used to optimize the path, shorten the path length, and reduce energy consumption. In addition, the real-time and adaptability of path planning cannot be ignored. It should be able to dynamically adjust the path according to environmental changes to ensure that the mobile subject safely and efficiently reaches the destination. Fig.1 actually represents the conclusion of the entire 9-step process.

Step 1: Data acquisition and pre-processing. The first step involves obtain surrounding environmental information through sensors and import the data into the machine for pre-processing. The car will obtain surrounding maps and obstacle location information, determine the starting point coordinates and target point coordinates of the planned path, use them to construct an environmental model, and generate a spatial model map.

Step 2: Function planning and path determination. After determining the map, use the artificial potential field method for global path planning, and conduct a search in the overall space to obtain an obstacle-free and collision-free optimal path, which is the initial path; Next, it is necessary to optimize it using a random function to ensure that the motion path of the actuator is the optimized path.

Step 3: Inflection point and path planning. Search for all inflection points on the moving path, convert the inflection point coordinates into local coordinates, and use them for the next step of local path planning.

Step 4: Start the experiment. The car travels along the original route, taking each turning point as a sub-target on the road, and using the optimized artificial potential field method to finely plan the local path of the car.

Step 5: Calculation of formulas. Formula correction and solution is very important, sequentially solve the position information of nodes on the path. Perform kinematic correction and inverse-solve the path nodes sequentially. Select the optimal solution combination based on the principle of minimum rotation angle, and perform forward kinematics to obtain the joint space coordinates. If the forward kinematics solution fails, it indicates that the path node exceeds the working range and needs to return to Step 2 to re-plan the path.

Step 6: Check if the sub-target point has been reached. If not, continue to use the artificial potential field method for path planning; If the sub-target point has been reached, adjust the car's actions to complete the path planning.

Step 7: Determine the current inflection point coordinates as the final destination. If so, complete the algorithm and output it as the optimal route; On this basis, an improved artificial potential field method is adopted to locally plan the next route.

Step 8: Detection. After obtaining the spatial position of the car, perform detection of the car's position and obstacles. If the car collides with an obstacle, return to step four and try to choose another solution. If all solutions in step four cannot meet the obstacle avoidance requirements, re-mark the obstacle point and set the value of the obstacle point's corresponding position in the environmental map to 1. Then return to Step 2 to re-plan the path until the obstacle avoidance requirements are met.

Step 9: Determine whether the target position has been reached. If it has reached the end of the algorithm, return to Step 4. The process of system collaborative obstacle avoidance path planning is shown in Fig.1.

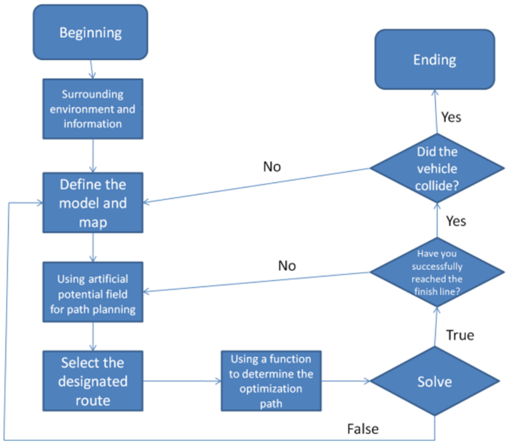

Fig. 1. The process of system collaborative obstacle avoidance path planning (Photo/Picture credit: Original).

3.3 Evaluation indicator setting

The following metrics are used to evaluate the performance of algorithms in autonomous vehicles, assessing factors such as efficiency, adaptability, and real-time response:

(1) Path Length. This measures whether the path planned by the algorithm is the shortest or near-optimal, which directly impacts driving efficiency and energy consumption in autonomous vehicles. Once the starting and ending points are determined, the algorithm plans a path from the start to the destination, and the path length can be calculated accordingly.

(2) Energy Consumption. This reflects the energy efficiency of the algorithm in practical applications and is crucial for evaluating both the economic and environmental impact. Professional energy measurement tools or methods are used to record energy consumption values, which are then used to perform comprehensive calculations.

(3) Real-time Performance. This metric examines the response speed of the algorithm when calculating the path. Real-time performance is critical for autonomous vehicles as they must respond rapidly to dynamic environmental changes. Time measurement tools or performance analysis tools can be employed to track response times.

(4) Adaptability Score. This evaluates the algorithm's ability to adjust and stabilize in response to environmental changes. A robust algorithm should be capable of flexibly adapting to various complex environments, with the adaptability score serving as a comprehensive indicator of this capability.

(5) Obstacle Avoidance Success Rate. This metric is vital for assessing the adaptability and reliability of the algorithm in real-world conditions. It highlights the stability and effectiveness of the algorithm in avoiding obstacles. Data from all experiments, including the number of successful obstacle avoidance attempts and the total number of attempts, are collected and statistically analysed to calculate the success rate. The success rate S_{oa} is calculated by dividing the number of successful obstacle avoidance attempts N_{sa} by the total number of experiments N_{te}, then multiplying by 100% to express the result as a percentage. This measure quantifies the effectiveness of the algorithm in successfully avoiding obstacles across all trials, offering insight into the algorithm's reliability and adaptability in dynamic environments.

4 Experiment and result analysis

The experimental reference data is presented in Table 1. Based on the data, to make the comparison more intuitive and concise, the information has been further organized into statistical charts. These charts highlight the advantages of the two algorithms Artificial Potential Field (APF) and MAB. However, before reviewing the charts, it is important to first consider the test settings:

Table 1. Experimental Parameter Table.

algorithm	Length(meter)	Energy consumption (kWh)	Real time performance (seconds)	Adaptability rating（1-0）	Success rate of obstacle avoidance（%）
APF	85.3	12.5	0.25	9.5	98
MAB	92.7	14.2	0.35	8.0	95

(1) Scenes Setting. In order to comprehensively evaluate the performance of the algorithm, a simulated urban road environment was set up in the experiment. In theory, it should include static obstacles (such as trees and buildings) and dynamic obstacles (such as pedestrians and other vehicles), but due to limited experimental conditions, only static obstacles are considered here, which is also the key to obstacle avoidance testing.

(2) Algorithm implementation. The artificial potential field method and dobby algorithm are respectively implemented in the path planning system of autonomous vehicles for direct comparison.

In this experiment, the performance of the artificial potential field method and multi-arm algorithm was mainly compared in several key indicators. From the data, the artificial potential field method has shown advantages in multiple aspects.

Firstly, according to the Fig.2, It shows that in terms of path length, the average path length of the artificial potential field method is 85.3 meters, while the average path length of the multi-arm algorithm is 92.7 meters. This means that the artificial potential field method is more effective in finding shorter paths, which is an important advantage for scenarios that require time and resource savings.

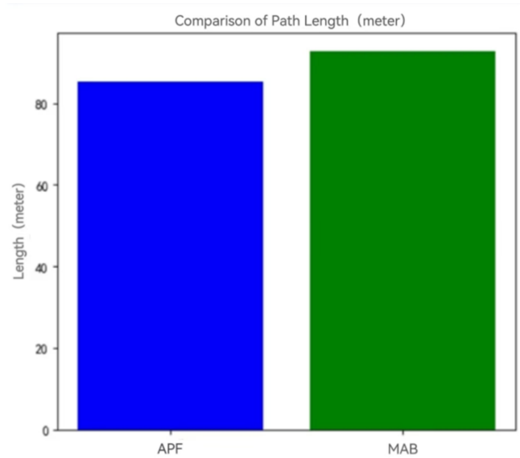

Fig. 2. Comparison of path length (Photo/Picture credit: Original).

Secondly, in terms of real-time performance, the artificial potential field method has a real-time performance of 0.25 seconds, while the multi-arm algorithm has a real-time performance of 0.35 seconds. It's what the Fig.3 wants to reveal. The artificial potential field method has a faster response speed, which is a key advantage for application scenarios that require rapid decision-making and response.

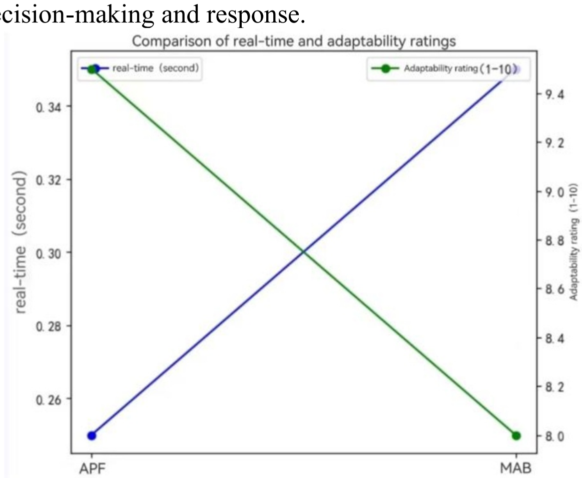

Fig. 3. Success rate of obstacle avoidance (Photo/Picture credit: Original).

Thirdly, in terms of adaptability score, the artificial potential field method scored 9.5 out of 10, while the multi-arm algorithm scored 8.0. This indicates that the artificial potential field method performs better in adapting to different environments and conditions, which is an important consideration for applications that need to operate in diverse environments.

Finally, in terms of obstacle avoidance success rate, the artificial potential field method has a success rate of 98%, while the multi-arm algorithm has a success rate of 95%. That information could be understood by Fig.4. The artificial potential field method performs better in obstacle avoidance, which is an important advantage for applications that require high safety and reliability. The artificial potential field method is superior to the multi-arm algorithm in terms of path length, energy consumption, real-time performance, adaptability, and obstacle avoidance success rate. This indicates that the artificial potential field method may be more suitable for application scenarios that require efficient, fast, and reliable path planning. However, the specific choice of algorithm still needs to be considered based on the actual application requirements and limitations.

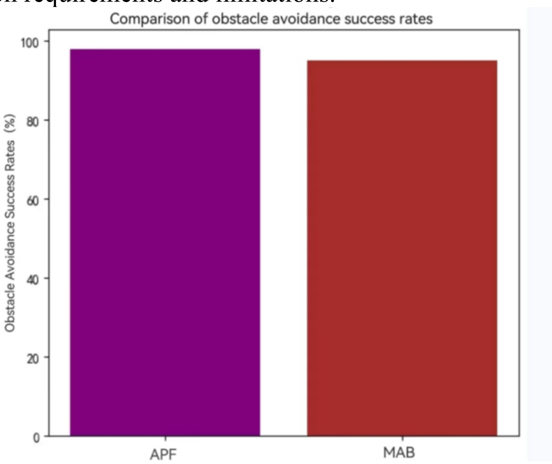

Fig. 4. Success rate of obstacle avoidance (Photo/Picture credit: Original).

5 Conclusion

In autonomous driving path planning, algorithms such as multi-arm algorithms and artificial potential field methods have their characteristics. Multi-arm algorithms, as a general method, have achieved significant results in fields such as online advertising but face challenges in autonomous driving. Its universality makes it difficult to fully meet the complex requirements of path planning, such as obstacle avoidance, compliance with traffic rules, and smooth driving. In contrast, the artificial potential field method has unique advantages in path planning. This method draws on the concept of a potential field in physics, treating the target point as a gravitational source and obstacles as repulsive sources, to construct a virtual potential field environment. In this environment, vehicles move in the direction of the potential field like particles, subject to both gravitational and repulsive forces. Therefore, the artificial potential field method can handle complex factors in path planning more finely and demonstrate better performance. This indicates that in the field of autonomous driving path planning, professional algorithms such as the artificial potential field method often have more advantages than general algorithms such as multi-arm algorithms. Looking ahead, future research and development in autonomous driving path planning should focus on enhancing adaptability and robustness of the artificial potential field method in dynamic environments, integrating advanced sensors and Artificial Intelligence technologies for improved accuracy and efficiency, developing real-time optimization techniques and ensuring scalability for widespread adoption, and exploring hybrid algorithms that combine the artificial potential field method with other approaches. These efforts will evolve the field, enhancing safety, efficiency, and comfort of autonomous vehicles.

References

1. T. Liu, Y.X. Li, Z.Y. Que, The Development Trend and Challenges of Autonomous Driving Technology with 5G and AI as Its Core, Guangxi Communication Technology, 2023(3),34

2. Y.F. Wang, Z.F. Zhao, Research on the Spatial Transformation of Existing Residential Streets Under the Influence of Autonomous Driving, China Academic Journal Electronic Publishing House, **3**, 2-4 (2023)
3. J.H. Chen, Life Decisions in Autonomous Driving, Journal of East China University of Political Science and Law, 5 (2020)
4. Q.J. Dong, W. Yang, G.P. Dong, Research on Path Planning for Autonomous Driving Based on A Algorithm*, Heilongjiang Science, **15**, Steptember 75 (2024)
5. Z. Liu, Multi-Armed Bandit Algorithm Based on Improved Collaborative Filtering and Its Application in Personalized Learning System[J], China Science and Technology Information, 16 (2023)
6. X.Y. Zhou, Research on End-to-End Autonomous Driving Technology Based on Deep Reinforcement Learning, (2023)
7. X.H. Yang, W.H. Zhou, Y.F. Li, X. Zhang, A Review of the Current Research Status of Path Planning Algorithms for Harvesting Robotic Arms, Journal of Chinese Agricultural Mechanization, **44**, 162 (2023)
8. T. Gao, Y.H. Wu, X. Zhang, C.F. Yue, Path Planning of Multi-Arms in Orbit Service Spacecraft Based on Double-Layer Game Theory, Chin. J. Space Sci., **42**,1231-1232 (2022)
9. Q.Z. Wang, A Disordered Relationship Optimization Method for Cooperative Obstacle Avoidance Path Planning in Multi-Mechanical Arm Systems, Manufacturing Automation, **46**, 71-73 (2024)
10. H.F. Wang, G.F. Xiang, Y.Q. Zhu, Obstacle Avoidance of Robotic Arm Based on PS-RRT Algorithm Path Planning, Application Research of Computers, **40**, 42-45 (2023)