PID Optimization Method Based on Heuristics Algorithm And Q-Learning

Yuang Li

School of Automobile Engineering, Harbin Institute of Technology (Weihai), Weihai, China

* Corresponding Author Email: 2022210152@stu.hit.edu.cn

Abstract. Proportional-integral-derivative (PID) control is a feedback control algorithm that adjusts the output signal through a combination of three control terms, proportional, integral, and differential, and is used to stabilize the actual output of the system close to the desired value. Heuristic algorithms are a class of intelligent problem-solving methods that optimize the solution using rules and trial-and-error learning; Q-learning is a reinforcement learning algorithm that focuses on decision-making in the Markovian decision-making process and achieves the optimal strategy by continuously learning and updating the value function. As people's requirements for the accuracy of various control systems increase, the requirements for the stability and robustness of PID control algorithms also increase. The traditional Ziegler-Nichols method to find PID parameters can no longer meet people's needs well, and people want to find more stable PID parameters faster. This article introduces different heuristic algorithms, such as particle swarm optimization algorithm and Q-learning algorithm for optimizing PID parameters in different systems. The effect of optimization of different algorithms is also analyzed.

Keywords: PID control; heuristic algorithms; Q-learning.

1. Introduction

Proportional-integral-derivative (PID) control stands out as the predominant control algorithm in industry and enjoys widespread acceptance in industrial control applications. The prevalence of PID controllers can be ascribed, in part, to their resilient performance across a broad spectrum of operating conditions and, in part, to their functional simplicity, enabling engineers to employ them in a straightforward and uncomplicated manner. P is primarily responsible for correcting the current error and bringing the system quickly closer to the setpoint. Increasing the proportional term will increase the response speed of the system but may cause the system to overshoot. I am mainly responsible for correcting the accumulated error of the system and eliminating the steady-state error. Increasing the integral term reduces the steady state error but may cause the system to respond slower and increase the overshoot. d is mainly responsible for suppressing oscillations in the system and reducing the overshoot. Increasing the derivative term improves the stability of the system but may lead to excessive sensitivity to rapidly changing errors.

In order to quickly seek the optimal parameters of PID, people began to try to combine heuristic algorithms with PID. To make the aircraft safer, the Harris Hawks optimization (HHO) algorithm is employed for regulating the aircraft's pitch angle [1]. The combination of Particle Swarm Optimization (PSO) algorithm, Ant Colony Optimization (ACO) and PID to find the optimal parameters has also been used for adjusting underwater vehicles [2]. At the same time, Rabiatuladawiah A. Hanifah et al. combine POS algorithm and ACO algorithm with PID to control the auxiliary steering of electric vehicles, making the battery utilization efficiency of electric vehicles higher [3]. Afterward, the whale optimization algorithm was proposed, and then Uzair Khaleeq uz Zaman et al. combined PID with it to find the optimal parameters of PID. The results showed that the effectiveness of this algorithm was much greater than that of POS algorithm and ant colony optimization algorithm [4].

Q-learning is a reinforcement learning algorithm designed to address problems based on Markov Decision Processes (MDP). The algorithm guides the agent's decision-making in the environment by learning a value function called Q, which represents the long-term expected cumulative reward for taking a specific action in a given state. During the training process, the agent continually updates Q values through interactions with the environment to maximize cumulative rewards. The core idea of Q-learning is based on the Bellman equation, iterating to optimize Q values, enabling the agent to make better decisions in future actions and ultimately achieve the goal of learning an optimal policy. This algorithm excels in model-free, sample-based learning and finds widespread applications in the fields of machine learning and artificial intelligence, particularly demonstrating significant achievements in control problems and gaming scenarios.

In order to further ensure the control of self-driving vehicles and make the system robust and adaptive, Yongqiang Yao et al. designed a variable domain fuzzy PID intelligent control method based on Q-learning, which can dynamically change the size of the domain to achieve the purpose [5]. Ignacio Carlucho et al. proposed an incremental Q-learning strategy for adaptive PID control and implementing this method on an actual mobile robot showcases its suitability for simultaneously tuning multiple adaptive PID controllers in real-time. The approach is demonstrated to be effective for real systems functioning under dynamic conditions within authentic environments [6]. Vo Nhu Thanh et al. developed a mathematical model about a linear following mobile robot system for a restaurant service robot and the service robot is controlled by an adaptive PID controller using Q-learning algorithm for control [7]. To ensure better voltage regulation, Jayita Saha et al. tried to design an adaptive PID controller with Q-learning algorithm to achieve the satisfactory performance of an automatic voltage regulator (AVR) [8]. To have better control for ships, Li Shijie et al. invented a Q-learning based PID control parameter tuning method, system, and storage medium for ships [9]. They proposed a mobile robot trajectory tracking control algorithm based on reinforcement learning and PID technology [10].

2. Analysis and presentation of optimization methods

2.1. PID parameter optimization based on heuristics algorithm

Many industrial systems are controlled with PID. In order to achieve good stability and adaptability, different optimization methods are used to adapt PID to different systems.

With more and more people traveling by air, aircraft safety is becoming increasingly important, and the need for a stable control system is very significant. The aircraft's take-off and landing processes are significantly influenced by its pitch angle, making the control of this parameter a crucial component of the aircraft control system. Numerous heuristic algorithms have been applied to manage the pitch angle effectively. Davut Izci et al. suggested employing a PID controller to regulate the aircraft's pitch angle and optimizing the PID parameters using the HHO algorithm [1]. The performance of this algorithm was evaluated by using statistical analysis and time and frequency domain analysis. The comparison reveals that the HHO algorithm outperforms the salp swarm algorithm (SSA) and atom search optimization (ASO) algorithm in controlling the pitch angle of the aircraft. The ASO algorithm, the SSA algorithm, and the HHO algorithm were run 20 times and the objective function values obtained from Tammen were compared as shown in Table 1. The results indicate that the HHO algorithm exhibits superior statistical performance in minimizing the objective function. The conclusive findings demonstrate that the approach based on HHO contributes to enhancing the stability performance of the aircraft pitch control system.

Table 1. Statistical Outcomes for Objective Function [1]

Algorithm	Best	Mean ± Std	Rank
HHO	0.0060	0.0065 ± 3.5298E-04	1
ASO	0.0071	0.0077 ± 3.4295E-04	2
SSA	0.0092	0.0096 ± 2.4566E-04	3

In order to better observe the ocean and develop marine resources, it is necessary to study and control the Autonomous Underwater Vehicle. The Autonomous Underwater Vehicle has factors such as fluctuation, swaying, lifting position, rolling, pitch, and yaw angle that need to be controlled. PID control is the first choice for controlling these factors, so how to find the right PID parameters has become the direction of many people's discussion. It has been engineered to adjust the proportional gain, integral gain, and derivative gain using proportional gain and the Ziegler-Nichols technique. T Herlambang et al. will use PSO and ACO to find the optimal parameters of PID in the control of Autonomous Underwater Vehicle [2]. PSO is composed of stochastic possible solutions initialized as a particle swarm, with each particle assigned a starting position and speed. As particles discover a direction leading to a food source, others trail behind. In contrast, the ACO algorithm is employed in the early stages when ants embark on their journey from home to the food source, selecting paths randomly. Prior to returning home, ants deposit pheromones on the paths they have previously traversed. Upon reaching their nest, the pheromone information undergoes updates influenced by the evaporation rate. Throughout the process of optimization, pheromone updates persist until all ants converge on similar paths, ultimately identifying the optimal route.

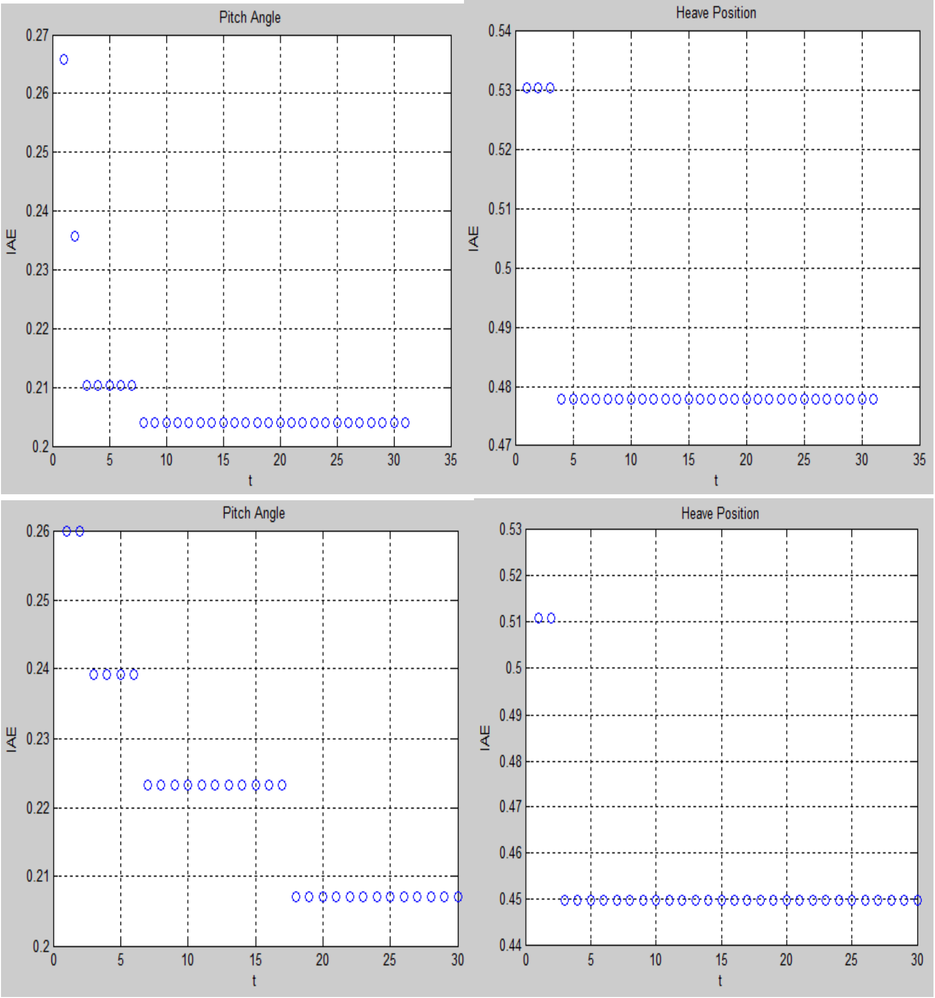

Fig. 1 Optimization results of pitch angle and heave position with PSO and ACO [2]

Two heuristics, PSO and ACO, were employed in their simulations. Using the POS algorithm, set the population number to 5 and iterate 31 times. The results are shown in Fig. 1. Fig. 1(a) shows the optimization process of the pitch angle by POS. The obtained PID parameters are Kp=1.896, Ki=1.733, Kd=0.637. And Fig. 1 (b) shows the POS's optimization process for the heave position, the three parameters are -0.603, -1.561, -1.339. In the ACO simulation, the specified parameters include a population number of 100, a set of 5 ants, a pheromone attenuation factor of 0.01, and a maximum iteration limit set at 30 as the stopping criterion. The ACO optimization procedure is illustrated in Fig. 1 (c) and Fig. 1 (d). Fig. 1 (c) shows the optimization process of ACO on the pitch angle. The PID parameters are 1.918, 1.872, 0.970. Fig. 1(d) shows the optimization process of ACO on the lifting position. The PID parameters are -0.181, -0.214, -1.323. The simulation results show that PSO and ACO can optimize the PID parameters and resultant response curves.

In order to reduce greenhouse gas emissions and the impact of fluctuating oil prices on people's lives, electric vehicle technology has gained growing interest. In electric vehicles, the battery is the most important part, especially the energy saving of the battery has a great impact on the life and performance of the electric vehicle. As one of the load appliances in electric vehicles, the electric-assisted steering system can be controlled with less current, where the PID controller parameters need to be tuned with optimal performance settings to reduce the current required for its operation, trying to use less current, to achieve optimal results, which results in less battery energy consumption. Rabiatuladawiah A. Hanifah et al. also used Particle Swarm Optimization and Ant Colony Optimization to find the optimal parameters of the PID to minimize the operating current within the electric power-assisted steering system [3]. Both algorithms were employed to derive optimal parameters for the PID controller and assess the viability of reducing auxiliary current in a Simulink model. Under the traditional PID controller, as depicted in Fig. 2(a), the RMS current value is 8.66A, with a maximum positive cycle value of 14.24A, an average value of 13.39A, and a maximum negative cycle assist current of 17.90A, and an average of -17.43A. Exploring the assist currents under the PID-PSO regulation controller, Fig. 2(b) illustrates an RMS current value of 8.47A. The maximum positive cycle assist current is 13.57A, and the negative cycle records a maximum of 17.98A, with average assist currents of 12.52A and -17.27A during the positive and negative cycles, respectively. Operating the system with a PID-ACO regulating controller, as seen in Fig. 2(c), results in a current of 8.38A, with a maximum assisted motor current of 13.20A and a maximum negative half-cycle value of -17.01A. Notably, the standard assisted current in the positive half-cycle is 12.22A, and during the negative half-cycle, it is -16.69A. Concerning the effective value of the assisted currents, the PID-ACO regulating controller outperforms the PID-PSO regulating controller in minimizing the supporting current. Furthermore, the results indicate that the PID-ACO controller yields a smaller supporting current in contrast to the traditional PID controller.

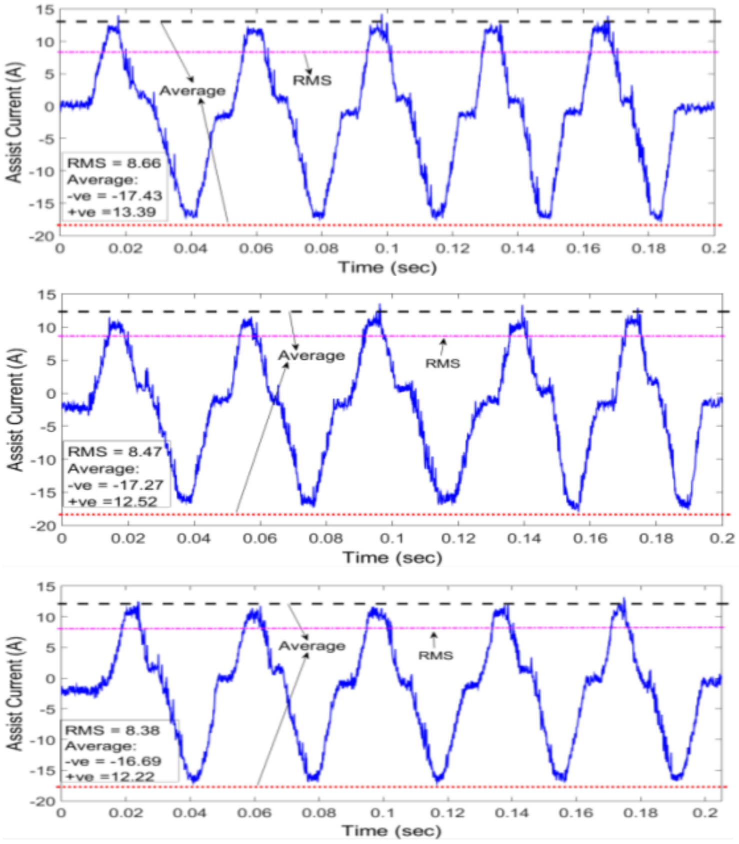

Fig. 2. Assist motor current [3]. (a) using traditional PID. (b) employing a controller tuned with PID-PSO. (c) employing a controller tuned with PID-ACO

In addition to particle swarm optimization and ant colony optimization, Uzair Khaleeq uz Zaman et al. investigated the tuning of PID parameters for different systems using the whale optimization algorithm [4]. The Whale Optimization Algorithm is a modern metaheuristic algorithm that mimics the foraging behavior of humpback whales to determine the optimal solution. The researchers applied this optimization method to fine-tune the parameters of the PID controller, incorporating three distinct transfer functions with diverse performance criteria. The objective of the investigation is to juxtapose the outcomes derived from the Whale Optimization Algorithm (WOA) with those from traditional tuning methods and various heuristic algorithms, aiming to assess the robustness of the obtained results. The PID curves resulting from different methods are depicted in Figure 3 and Table 2. By comparing the obtained results with Ziegler-Nichols (Z-N), Genetic Algorithm (GA), and Particle Swarm Optimization (PSO), it was found that the performance of the results from the Whale Optimization Algorithm is higher than that of the results from the above methods.

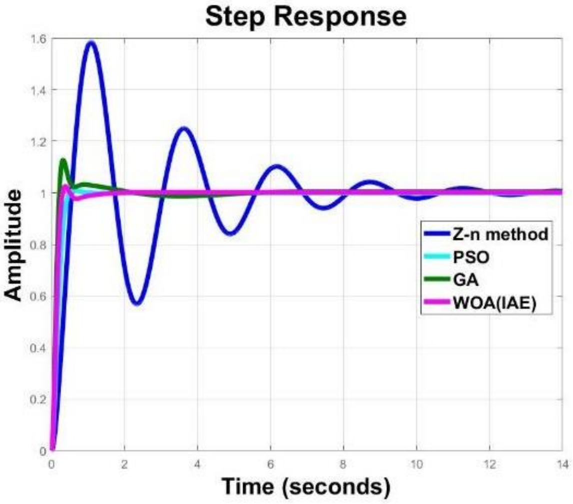

Fig. 3. Step response of TF-2 for different tuning approaches [4]

Table 2. Comparison of characteristic values for TF-2 [4]

No	Method	Rise Time(sec)	Overshoot (%)	Settling time (sec)	Steady State Error (%)
1	Z-N	0.386	58.3616	10.171	0.095975
2	PSO	0.279	0.90689	0.4331	0.82617
3	GA	0.143	12.6574	1.4783	0.1196
4	WOA	0.190	2.4651	0.7908	0.0053362

2.2. PID parameter optimization based on Q-learning

With the development of artificial intelligence, reinforcement learning, especially Q-learning, began to be used in various aspects. Q-learning finds the best strategy by using a certain reward or punishment and its formula to find the maximum value. Thus, it began to be hoped that Q-learning could be used to find the optimal parameters of PID controllers. In automatic driving, PID control is widely used because of its stability and simplicity, but in complex road environments, vehicle control must be precise. Yongqiang Yao et al. formulated a Q-learning-based fuzzy PID intelligent control approach. To enhance the vehicle's control, a variable domain is implemented, ensuring system robustness and adaptability. The dynamic adjustment of the domain size further solidifies control objectives [5]. By employing the Q-learning method, the researchers achieved online learning of the proportionality factor. The variable domain fuzzy PID algorithm, rooted in Q-learning, utilizes the variable error rate and error rate as inputs for real-time PID parameter tuning. It is found through simulation that the results are accurately enhanced by a 15% margin when contrasted with the conventional fuzzy PID, reflecting this method's effectiveness.

In addition to autonomous driving, Q-learning's approach to PID optimization is also used in the direction of mobile robotics. Numerous intelligent systems, such as those overseeing technological systems, are currently in development. PID control is extensively applied across diverse control tasks. However, in scenarios where the system is not fully comprehended or operating conditions are variable and unknown in advance, conventional PID tuning becomes impractical. Ignacio Carlucho et al. introduced an incremental q-learning algorithm for the online learning of mobile robots and adaptive PID controller parameters [6]. The incorporation of a memory concept in the learning process facilitates the management of the specialization process. This process strategically defines the learning space for states and actions based on system behavior, aiming to mitigate computational challenges and address the well-known curse of dimensionality in the realm of reinforcement learning problems. By experimenting with a mobile robot, this result greatly improved the robustness of the control.

Wheeled mobile robots are extensively employed among mobile robot types, which also include legged robots. The prevalence of wheeled mobile robots in industrial and consumer applications can be attributed to their uncomplicated structure, ease of modeling and control, motion flexibility, and economic operation and manufacturing. The performance of wheeled mobile robots is heavily reliant on the fine-tuning of PID control parameters and the robustness of the mechanical system. To accommodate changes in system parameters and operating conditions, analog controllers are being replaced by digital counterparts. This shift allows for the automatic adjustment of PID control parameters, enabling optimal tuning to address uncertainties and system variability in different operating conditions. Vo Nhu Thanh et al. developed a linear-following mobile robot, that serves as a mathematical model for a restaurant service robot [7]. The service robot is governed by an adaptive PID controller using a Q-learning algorithm. The algorithm is trained with a fixed set of initial settings, enabling the robot to move smoothly along a path to a predetermined position at a set speed. Experiments were conducted on a 10-meter fixed path, involving two 90-degree turns between 4-5 meters and 6-7 meters. The restaurant mobile robot was operated sequentially using both the PID controller and the Q-learning PID controller. In Figure 4, the offset values of the restaurant mobile robot concerning its intended path are depicted for runs involving the PID controller and the Q-learning PID controller. Simulation outcomes suggest that the Q-learning adaptive PID controller exhibits clear advantages over both the conventional PID controller and the Q-learning algorithm controller.

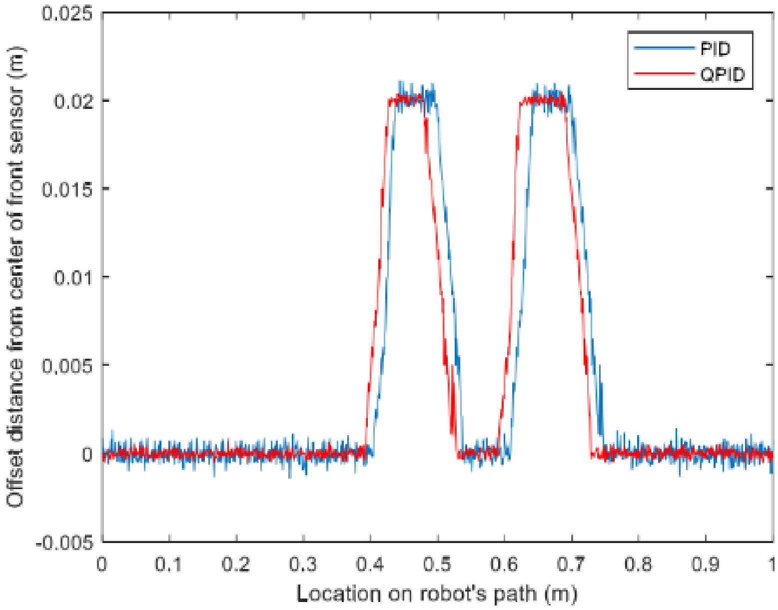

Fig 4. Performance of PID and QPID controller [7]

In power control, the same is required with a stable PID controller. Voltage and reactive power control are important aspects of achieving smooth operation of power systems. Every electrically interconnected device is progressing toward a designated voltage level, referred to as the nameplate or rated voltage. To ensure a consistent voltage level and minimize losses in the transmission line, controlling the flow of both reactive and real power becomes paramount. Consequently, Automatic Voltage Regulators (AVRs) are extensively employed to manage field excitation, ensuring the maintenance of the specified voltage at the generator terminals to address these scenarios. Due to load fluctuations and the high inductance of the generator field windings, maintaining a stable and fast regulator response is a challenge. Therefore, it is crucial to improve the AVR performance to ensure an effective and stable response to transient changes in terminal voltage, so most people use PID controllers to control the AVR. Adjusting the gain of the PID controllers can be difficult to avoid long-term oscillations and unwanted overshoots. Jayita Saha et al. proposed a Q-Learning (QL)-based PID controller for AVR control [8]. combined with a genetic algorithm to determine the effectiveness of the proposed method. Then, the simulation results were analyzed in the same scenario and after comparison. Figure 5 is the curve after optimizing PID parameters using different methods. It was found that the QL strategy resulted in faster response and less overshooting. In addition, the stability and sensitivity of the system were analyzed and determined to be satisfactory. The examination of the outcomes verifies that the newly devised adaptive PID controller takes into account the speed of response, peak overshoot, and damping of voltage oscillations.

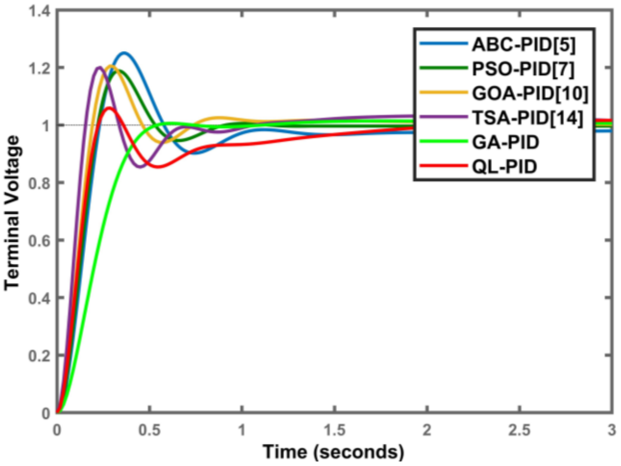

Fig. 5. Comparative terminal voltage profile of AVR with various controllers.[8]

In the field of ship control technology, PID control is also widely used. Still, in order to use PID control system to control the ship's heading, it is often necessary to manually adjust the three parameters Kp, Ki, and Kd in a cumbersome manner. The adjustment of the work usually relies on engineering experience, which depends on human experience. In the process of manually adjusting the adjustment parameter, randomness is large. Manual adjustment of the adjustment parameter will probably lead to the ship control system not being able to operate stably. In order to avoid the instability of the ship control system caused by the process of manually adjusting the parameters, and to improve the precision of ship control, Li Shijie et al. invented a ship PID control parameter adjustment method, system, and storage medium based on Q-learning [9]. They first determine a ship state; determine an action of a PID control parameter, and then use a Q-learning algorithm to rectify the ship PID control parameter according to said ship state and said action of the PID control parameter. This method accurately adjusts the PID control parameters employing the ship state, the action of the PID control parameters, and the Q learning algorithm to improve the control accuracy of the ship, thereby avoiding the tedious and repetitive operation of manually piecing together the control parameters by relying on the engineering experience and reducing the occurrence of the instability of the ship system.

Q-learning and PID control can enhance the trajectory tracking algorithm for mobile robots. The widespread utilization of mobile robots in road sign detection, service occupation, and outer space exploration has highlighted the significance of refining their tracking capabilities. While numerous studies focus on the path-tracking control of mobile robots, achieving precise positioning or tracking of reference trajectories in space remains challenging. The advent of reinforcement learning offers innovative solutions to address this issue. Shuti Wang et al. introduced a trajectory-tracking control algorithm for mobile robots based on a combination of reinforcement learning and PID techniques [10]. This approach utilizes Q-learning and PID methods to track the desired trajectory of the mobile robot. To enhance tracking accuracy, the method reduces the computational complexity of the reward function in Q-learning. While Q-learning excels in learning and achieving optimal results through trial and error, it faces challenges in complex conditions, is susceptible to dimensional catastrophes, and requires improvement in its control capabilities. On the other hand, PID control proves effective and robust but lacks learning capabilities. The experimental simulation results show that Q-learning-PID can guarantee that the value function converges to zero, which indicates that the path-tracking effect of Q-learning-PID is better than that of single Q-learning or PID. This new type of controller can be applied to more engineering fields, and it is worth further study.

3. Conclusion

In this paper, different optimization algorithms are presented for optimizing PID parameters in different systems. Compared with the traditional methods, heuristic algorithms, and Q-learning algorithms can find more stable PID parameters faster, which leads to better stability and robustness of the system. For different control systems, the best optimization algorithm may be different. PID control is widely used in various fields, and future development may include a smarter and self-adaptive PID control system, which is able to adjust the parameters automatically according to the real-time environment and the system change that improves system performance. With the development of IoT and cloud computing, using them to implement distributed PID control systems can make control decisions smarter, more flexible, and able to learn and optimize from large amounts of real-time data. Heuristic algorithms and Q-learning can be used to automatically optimize the parameters of PID controllers, adjusting them in real time according to system feedback to improve performance. With the development of reinforcement learning algorithms, the PID parameters can also be learned and adapted online at runtime to adapt to changes in the system over a long period. At the same time, these algorithms can play an important role in coping with complex, nonlinear systems, helping PID control to better adapt to changing industrial environments. Therefore, PID control will be more useful in the future, and the future with the appropriate algorithms to optimize PID control the development of PID control is an indispensable link

References

[1] Izci D, Ekinci S, Demirören A, et al. HHO algorithm based PID controller design for aircraft pitch angle control system[C], in 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 2020, 1-6.

[2] Herlambang T, Rahmalia D, Yulianto T. Particle swarm optimization (PSO) and ant colony optimization (ACO) for optimizing pid parameters on autonomous underwater vehicle (AUV) control system[C], in Journal of Physics: Conference Series, IOP Publishing, 2019, 1211(1): 012039.

[3] Hanifah R A, Toha S F, Ahmad S, et al. Swarm-intelligence tuned current reduction for power-assisted steering control in electric vehicles. IEEE Transactions on Industrial Electronics, 2017, 65(9): 7202-7210.

[4] Zaman U K, Naveed K, Kumar A A. Tuning of PID controller using whale optimization algorithm for different systems[C], in 2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2), Islamabad, Pakistan, 2021, 1-5.

[5] Yao Y, Ma N, Wang C, et al. Research and implementation of variable-domain fuzzy PID intelligent control method based on Q-Learning for self-driving in complex scenarios. Mathematical Biosciences and Engineering, 2023, 20(3): 6016-6029.

[6] Carlucho I, De Paula M, Villar S A, et al. Incremental Q-learning strategy for adaptive PID control of mobile robots. Expert Systems with Applications, 2017, 80: 183-199.

[7] Thanh V N, Vinh D P, Nam L H, et al. Reinforcement Q-learning PID Controller for a Restaurant Mobile Robot with Double Line-Sensors[C], in Proceedings of the 4th International Conference on Machine Learning and Soft Computing, Haiphong City, Vietnam, 2020, 164-167.

[8] Saha J, Guha D, Jha S K. Voltage Control of Power System employing Q-Learning based PID Controller[C], in 2022 IEEE 19th India Council International Conference (INDICON), Rajkot, India, 2022, 1-6.

[9] LI S, ZHANG P, LIU J, WANG L, XIE L. Setting ship PID control parameter based on Q-learning [P], China, CN 112947054 A, 2021.

[10] Wang S, Yin X, Li P, et al. Trajectory tracking control for mobile robots using reinforcement learning and PID. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 2020, 44: 1059-1068.