УДК 004.896

NAVIGATION IN MOBILE ROBOTS USING TWIN DELAYED DEEP DETERMINISTIC POLICY GRADIAENT

Бархум М (ИТМО)

Научный руководитель – профессор, доктор технических наук Пыркин А.А. (ИТМО)

Introduction. Navigating mobile robots through complex and dynamic environments poses a significant challenge in the field of robotics. Traditional methods often struggle to provide adaptive and efficient solutions, prompting the exploration of advanced techniques. This research delves into addressing the navigation problem by employing the Deep reinforcement learning techniques namely, Twin Delayed Deep Deterministic Policy Gradient algorithm. Which is a very effective reinforcement learning algorithm, that is used to optimize continuous control policies for mobile robot navigation. By leveraging deep neural networks, it enables the learning of robust and adaptive policies making autonomous robots navigate with precise movement actions, while adaptively avoiding obstacles.

Main part. Twin Delayed Deep Deterministic Policy Gradient algorithm offers a lot of advantages in solving the navigation problem, since it generally has stable and easy to tune hyperparameters, not to mention, its ability to mitigate overestimation bias, that can occur in other reinforcement learning algorithms, through the use of twin Q-networks and delayed policy updates.

Taking that into account the algorithm runs through the following stages:

- Random goal is given to the robot as polar coordinates.
- Using sensor data, the robot learns to discern obstacles.
- Obtain a new point of interest.
- Calculating the distance between each candidate point of interest and the goal.
- Choose the point of interest with the minimum distance function.
- Obtain action from the actor network.
- Perform the action and calculate immediate rewards.
- Based on the twin critic networks output choose optimal Q-value.

Conclusion. The work is implemented as a Robotic Operating System (2.0) simulation of a mobile robot trained in an environment. The results will show how the trained robot navigates in the simulation environment while avoiding obstacles and choosing the optimal path to the random goal given at each time step. Showing that the navigation system is capable of exploring and navigating in a previously unknown environment and reliably finding its way to the global goal.

References.

1. Cimurs, R., Suh, I.H. and Lee, J.H., 2021. Goal-driven autonomous exploration through deep reinforcement learning. *IEEE Robotics and Automation Letters*, *7*(2), pp.730-737.
2. Bai, S., Song, S., Liang, S., Wang, J., Li, B. and Neretin, E., 2022. UAV maneuvering decision-making algorithm based on twin delayed deep deterministic policy gradient algorithm. *Journal of Artificial Intelligence and Technology*, *2*(1), pp.16-22.