

Abstract

The emergence of unmanned aerial vehicles (UAVs) has provided convenience for people's lives and can also help people working in cities with tasks such as traffic monitoring, multi-target tracking, express delivery, and so on. If the UAV loses signal at work, it will get out of control. Therefore, in order to ensure the working quality of the UAV, optimizing the trajectory of the UAV to ensure the connection quality has become a feasible way.

In recent years, various techniques have been continuously used for trajectory optimization problems, among which graph, dynamic programming, reinforcement learning (RL), Deep-Q-network (DQN), etc. are widely used. However, when these technologies optimize the UAV trajectory, the actions of the UAV need to be preset and fixed. If the action range is larger, the granularity of the trajectory will be larger, and some better flight points may be missed. If the range of motion is small or there are many actions, it will cause computational performance problems. A feasible way is to use the continuous output Advantage Actor Critic (A2C) algorithm, which means that the UAV has a richer action space. It is improved from the Actor-Critic (AC) algorithm and solves the problem of excessive variance of the AC algorithm. We created a simulation environment and compared the performance of the DQN and A2C algorithms on the connectivity-aware UAV path planning problem. The results show that the A2C algorithm is slightly better than the DQN algorithm.