

Robust Multi-Agent Reinforcement Learning for Autonomous Vehicle in Noisy Highway Environments Supplementary Material

Lilan Lin
Xiaotong Nie*
Jian Hou

LINLILANLLL@163.COM
NIXIAOTONG@ZSTU.EDU.CN
CHANGELEAP@163.COM

School of Computer Science and Technology, Zhejiang Sci-Tech University, Zhejiang, China

Editors: Vu Nguyen and Hsuan-Tien Lin

Appendix A. Hyperparameters for Experiments

Table 1: Hyperparameter setting

Name of the parameters	Value
Total training episodes	20000
Testing interval	200
Testing rounds	3
Step length per round	100
Learning rate	5e-4
Discount factor	0.99
Activation functions	Softmax
Memory capacity	100
Critic loss	Huber
GAE lambda	0.97
Optimizer	Adam

Appendix B. The Configuration of CA-MAPPO

The factors influencing the effectiveness of CA-MAPPO are further explored. Both from the training results and the test results, the performance of CA-MAPPO has the most obvious advantage over the other algorithms when the noise intensity K is 5. Therefore we select the initial value of the noise intensity $K_0 = 1, 2, 3$ and set the step size ΔK to 0.5 and 1 respectively in the noisy environment when $K = 5$. Fig. 1 shows how the K and reward alter throughout the training process.

From Fig. 1, it is evident that when $K_0 = 1$, curriculum learning proceeds effectively, with K reaching the target value of 5 and achieving an average reward of approximately 20. As K_0 increases, we observe that K fails to reach the target value during curriculum learning, accompanied by an obvious decline in the average reward. At $K_0 = 3$, the value

* Corresponding author

of K is fixed and the average reward is below 0, curriculum learning fails. We attribute this to the excessively large K_0 making the initial task too challenging. Additionally, at $K_0 = 1$, a ΔK of 1 accelerates the curriculum learning process compared to a ΔK of 0.5, yet their final average rewards are comparable. Similarly, at other values of K_0 , the average rewards also exhibit similar trends.

The experimental results show that K_0 is a critical factor influencing the process of curriculum-based adversarial learning. Therefore, selecting the appropriate K_0 is crucial for the effectiveness of CA-MAPPO.

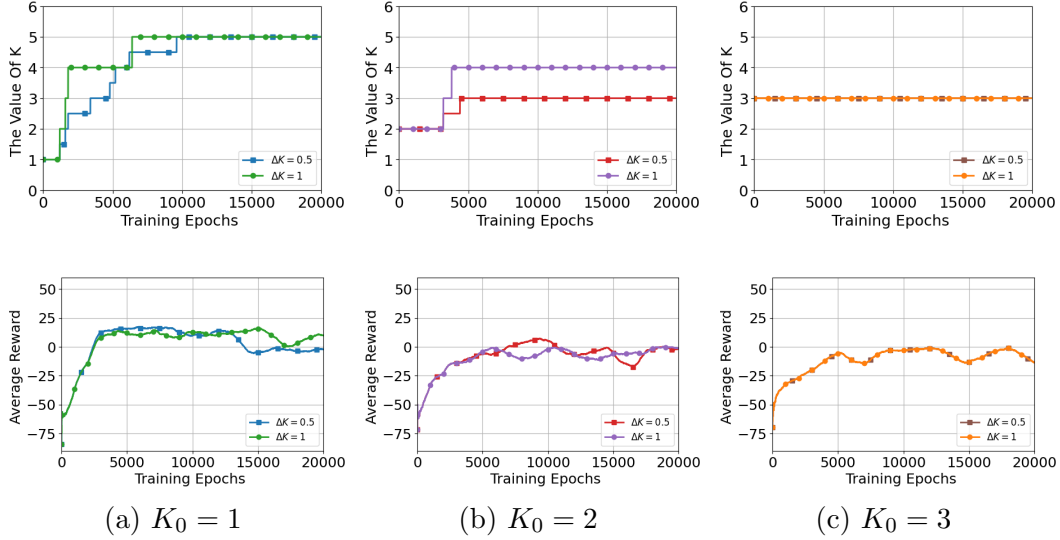


Figure 1: The influence of K_0 and ΔK on CA-MAPPO under $K = 5$