Elevator Group Control Systems via Deep Reinforcement Learning

David Li, Keith Wang, Michael Yuan, Zhihao Shan

Supervisor: Mathieu Laurière

Problem Statement

- Elevator Group Control Systems (EGCS)
- Deep Reinforcement Learning
- Multi-Agent Deep Reinforcement Learning

Related Work

- Elevator Group Control System (EGCS)
- Reinforcement Learning (mainly Policy-based)
- Reinforcement Learning in EGCS
    - "Elevator Group Control Using Multiple Reinforcement Learning Agents" Crites &. Barto, 1998

Environment

- **States:** direction buttons (hall call), floor buttons (car call), destinations of passengers…
- **Observation Space:** car call (nxk), hall call (2n), elevator position (nxk). 2n(k+1) in total
- **Actions:** stay on the current floor and do nothing, go up one floor, go down one floor, load passengers intended to go up, load passengers intended to go down, unload passengers

Environment

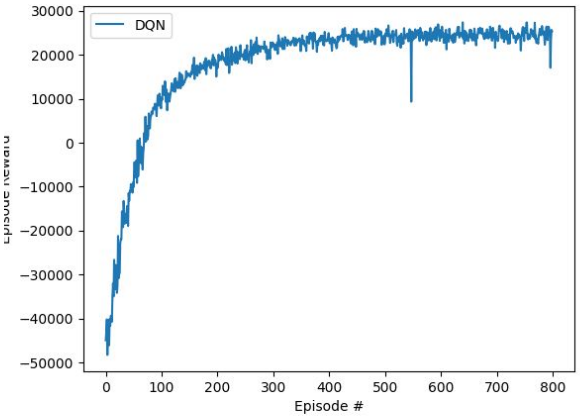
- **State Transition Probability:** At each the time step, motion of elevators is deterministic based on the actions taken. Arrival of new passengers follows poison distributions based on floor number. Destinations of new passengers are randomly decided.
- **Reward Function:** b: number of passengers reaching destinations, c: number of passengers going toward destinations, d: number of passengers trying to join a full queue but are denied, e: number of passengers reaching maximum waiting time and leaving the queue, f: number of passengers going away from destination, g: number of passengers in the elevator, h: number of passengers waiting in the queue

$$r_t = 30b_t + 5c_t - 10d_t - 10e_t - 3f_t - g_t - h_t$$

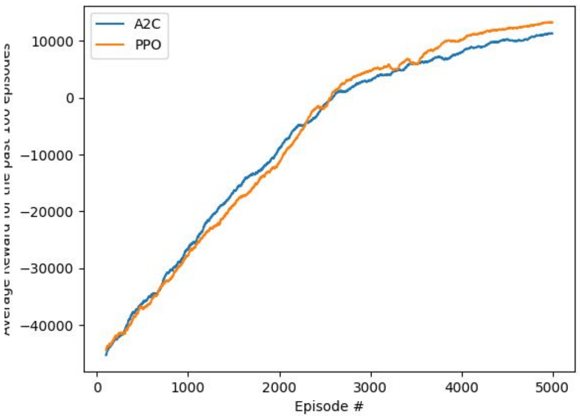
Algorithm

- DQN (twisted for multi-discrete action space)
- A2C
- PPO

Results



DQN
30000
20000
10000
0
−10000
−20000
−30000
−40000
−50000
0
100
200
300
400
500
600
700
800
Episode #



A2C
PPO
10000
0
−10000
−20000
−30000
−40000
0
1000
2000
3000
4000
5000
Episode #

Results

- Parameters
  - **Environment**: termination steps=1000, num_elevator=3, num_floor=6, max_queue=20, capacity=10.
  - **DQN**: lr=1e-3, gamma=0.999, num_epi=1000, buffer size=1e5, batch size=256, update every=4, clone every=1e3.
  - **A2C**: lr=3e-4, n_steps = 1024, gamma = 0.999, gae_lambda = 0.98, ent_coef = 0.01, total timesteps=8e6.
  - **PPO**: lr=3e-4, n_steps = 1024, gamma = 0.999, gae_lambda = 0.98, ent_coef = 0.01, total timesteps=8e6, batch_size = 256, n_epochs = 4.

Conclusion

- Implemented DQN, A2C, and PPO in a OpenAI gym-based environment.
- Tried the application in a continuous time, real-life multi-agent scenario.
- The challenge is the intricate and partially observable state for agents and higher dimension of actions in real-world EGCS.

References

- M. Burke, “Rlevator: A farama gymnasium environment for elevator control,” 2023. [Online]. Available: https://github.com/mwburke/RLevator
- M.-L. Siikonen, “Elevator traffic simulation,” SIMULATION, vol. 61, no. 4, pp. 257–267, 1993.
- M.-L. Siikonen, “Elevator group control with artificial intelligence,” Citeseer, 1997. [Online]. Available: https://api.semanticscholar.org/CorpusID:14627167
- V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, “Asynchronous methods for deep reinforcement learning,” CoRR, vol. abs/1602.01783, 2016. [Online]. Available: http://arxiv.org/abs/1602.01783
- J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” 2017.
- A. Y. Ng, D. Harada, and S. Russell, “Policy invariance under reward transformations: Theory and application to reward shaping,” in Icml, vol. 99.Citeseer, 1999, pp. 278–287.
- J. K. Gupta, M. Egorov, and M. Kochenderfer, “Cooperative multi-agent control using deep reinforcement learning,” In Autonomous Agents and Multiagent Systems: AAMAS 2017 Workshops, Best Papers, São Paulo, Brazil, May 8-12, 2017, Revised Selected Papers 16 .Springer, 2017, pp. 66–83.

References

- R. Crites and A. Barto, “Improving elevator performance using reinforcement learning,” Advances in neural information processing systems, vol. 8, 1995.
- R. H. Crites and A. G. Barto, “Elevator group control using multiple reinforcement learning agents,” Machine learning, vol. 33, pp. 235–262, 1998.
- Q. Wei, L. Wang, Y. Liu, and M. M. Polycarpou, “Optimal elevator group control via deep asynchronous actor-critic learning,” IEEE transactions on neural networks and learning systems, vol. 31, no. 12, pp. 5245–5256, 2020.
- Z. Cao, R. Guo, C. M. Tuguinay, M. Pock, J. Gao, and Z. Wang, “Application of deep q learning with simulation results for elevator optimization,” 2022.
- C. Cassandra, Discrete Event Systems: Modeling and Performance Analysis. Homewood, IL: Aksen Associates, 1993.