

УДК 004.8

**АЛГОРИТМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ НА ОСНОВЕ РРО ДЛЯ
МУЛЬТИАГЕНТНОЙ МАРШРУТИЗАЦИИ.**

Жогов А.А. (МФТИ)

Научный руководитель – кандидат физико-математических наук

Фильченков А.А. (Университет ИТМО)

Введение. В настоящее время задача маршрутизации может применяться почти во всех сферах повседневной жизни от доставки товаров до механизмов управления большими предприятиями. Отдельной темой данной проблемы можно выделить мультиагентную маршрутизацию, примером которой является например конвейерная система. Обычно такие задачи решаются с помощью классических методов [1], но последнее время все больше внимание уделяют использованию машинного обучения в этой области, а конкретно обучению с подкреплением [2]. Для таких систем возможно применение различных алгоритмов обучения с подкреплением, в данном случае мы пытаемся превзойти результаты, полученные для более простых алгоритмов обучения с подкреплением, применяя РРО [3].

Основная часть. Основной задачей является применить идеи РРО, в контексте мультиагентного обучения, где на пути маршрутизируемого объекта встречается множество принимающих решение агентов. Работа проходит в контексте новой симуляционной среды, поэтому целями работы являются:

- Перенос предыдущих алгоритмов обучения с подкреплением на новую симуляционную среду
- Добавления алгоритма, основанного на РРО
- Сравнение предыдущих алгоритмов обучения с подкреплением для задачи мультиагентной маршрутизации с РРО

Выводы. На новую среду перенесены предыдущие алгоритмы и добавлен РРО, проведен сравнительный анализ.

Список использованных источников:

1. *Sørensen R. A., Nielsen M., Karstoft H.* Routing in congested baggage handling systems using deep reinforcement learning // *Integrated Computer-Aided Engineering*. — 2020. — Vol. 27, no. 2. — P. 139–152.
2. *Boyan J. A., Littman M. L.* Packet routing in dynamically changing networks: A reinforcement learning approach // *Advances in neural information processing systems*. — Citeseer. 1994. — P. 671–678.
3. *John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov.* Proximal Policy Optimization Algorithms. arXiv, 2017.

Жогов А.А. (автор)

Подпись

Фильченков А.А. (научный руководитель)

Подпись