

## **Децентрализованный алгоритм управления конвейерной системой с использованием методов мультиагентного обучения с подкреплением**

Мухутдинов Д. В., Университет ИТМО, г. Санкт-Петербург  
Научный руководитель – Фильченков А.А., к.ф-м.н., доцент ФИТиП Университета ИТМО

### **Введение**

Конвейерные системы широко применяются для автоматизированной транспортировки объектов и материалов. Они являются неотъемлемой частью комплексов промышленного оборудования, пунктов сортировки грузов, багажных систем в аэропортах и т. д. Оптимизация работы таких системы имеет высокую практическую и экономическую значимость, что мотивирует поиск эффективных методов управления конвейерными системами.

При рассмотрении штучных конвейеров, перемещающих отдельные объекты, как в случае систем для перемещения багажа в аэропортах, задача управления конвейерной системой по большей части сводится к задаче пакетной маршрутизации. Большинство существующих алгоритмов пакетной маршрутизации разработаны специально для применения в компьютерных сетях. Однако, задача маршрутизации в компьютерных сетях качественно отличается от задачи маршрутизации в конвейерных сетях, прежде всего, потому что при маршрутизации в компьютерных сетях ограничения на скорость работы алгоритма и количество сообщений, пересылаемых узлами сети друг другу, существенно строже. Таким образом, алгоритмы маршрутизации, применяемые в компьютерных сетях, как правило, не являются оптимальными для маршрутизации в конвейерных сетях. Кроме того, эти алгоритмы, как правило, не учитывают специфические дополнительные критерии качества работы системы, такие как низкое энергопотребление.

### **Цель работы**

Целью данной работы является разработка децентрализованного алгоритма маршрутизации в конвейерных сетях для штучных объектов на примере систем управления багажом, способного оптимизировать работу системы по нескольким критериям (таким как среднее время объекта в пути и уровень энергопотребления системы).

### **Базовые положения исследования**

Основной идеей является формулирование задачи маршрутизации как задачи мультиагентного обучения с подкреплением. Такой подход позволит применить к задаче современные методы обучения с подкреплением, в том числе глубокого обучения с подкреплением (с использованием нейронных сетей).

### **Предварительные результаты**

На данный момент разработана имитационная модель конвейерной системы для проведения экспериментальных исследований. Разработан алгоритм DQN-routing, основанный на методе глубокого Q-обучения. Алгоритм DQN-routing является гибридным алгоритмом, объединяющим в себе идеи алгоритма Q-routing [1] и алгоритма link-state [2]. В алгоритме DQN-routing нейросетевой агент получает на вход данные о текущем состоянии системы (номер узла назначения текущего пакета, матрицу смежности сети, полученную с помощью link-state протокола, и т. д.), и выдает в ответ Q-функцию — оценки стоимости

дальнейшего пути пакета при условии пересылки пакета в одного из соседей текущего узла. Обучение нейросетевого агента осуществляется аналогично алгоритму DQN [3].

В ходе экспериментальных исследований алгоритм был сравнен с алгоритмами link-state и Q-routing. Было показано, что, в отличие от алгоритма link-state, алгоритм способен одновременно оптимизировать и энергопотребление, и скорость доставки пакетов до пунктов назначения. Также было показано, что алгоритм адаптируется к изменениям в топологии сети эффективнее, чем Q-routing.

В ходе дальнейших исследований планируется исследовать эффективность применения других техник обучения с подкреплением к данной задаче, а именно методов градиента стратегии (актора-критика) и методов, основанных на моделировании среды (model-based).

#### Список литературы

1. Boyan J. A., Littman M. L. Packet routing in dynamically changing networks: a reinforcement learning approach // *Advances in Neural Information Processing Systems*. — 1994. — No. 6. — P. 671–678.
2. McQuillan J. M., Richer I., Rosen E. C. The New Routing Algorithm for the ARPANet // *IEEE Trans. on Comm.* — 1980. — Vol.28,no.5. — P.711 – 719.
3. Human-level control through deep reinforcement learning / V. Mnih [et al.] // *Nature*. — 2015. — No. 518. — P. 529–533. 11