

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

Мухутдинов Дмитрий, группа М4239

Научный руководитель: Фильченков А. А., к.ф-м.н., доцент ФИТиП

Консультант: Вяткин В.В., д.т.н., профессор ФИТиП

Факультет Информационных Технологий и Программирования
Мегафакультет Трансляционных Информационных Технологий
Университет ИТМО, Санкт-Петербург

17 апреля 2019 г.

Применения:

- Промышленность
- Сортировка грузов
- Распределение багажа
- ...

Будем рассматривать случай с багажом.

- Статические стратегии управления¹
 - Разрабатываются под конкретную топологию системы
- Model predictive control (MPC)²³⁴
 - Решает глобальную оптимизационную задачу в форме LP/QP
 - Вся система контролируется централизованно
 - Не может обрабатывать изменения в системе, не заложенные в модель (поломки)
- Маршрутизация по аналогии с компьютерными сетями⁵
 - Децентрализованное вычисление, устойчивость к поломкам
 - Оптимизирует только скорость доставки чемоданов

¹De Neufville.

²Cataldo, Scattolini.

³Zeinaly, De Schutter, Hellendoorn.

⁴Luo, Huang, Zhang.

⁵Yan, Vyatkin.

Разработать алгоритм управления конвейерной системой со следующими свойствами:

- Децентрализованность
- Многокритериальная оптимизация (время доставки багажа + энергопотребление)
- Устойчивость к разнородным изменениям в условиях среды
 - Изменения характеристик багажного потока
 - Поломки конвейеров

- Сосредоточимся на обобщенной задаче маршрутизации в ориентированном графе
- Обучение с подкреплением
- Нейросети в качестве обучающихся агентов
- Q-routing⁶
 - Не получил широкого распространения в компьютерных сетях (использует слишком много служебных пакетов)
 - В других задачах (трафик, конвейеры) это не является проблемой.

⁶Boyan, Littman.

Постановка задачи в терминах RL

- Рассмотрим *пакет* в сети как обучающегося агента, взаимодействующего с сетью как со средой
- Полное состояние среды неизвестно, состояние текущего роутера — *наблюдение* пакета
- Действие — переход к одному из соседей
- Q-learning:

$$Q(o_t, a_t) \leftarrow r_t + \gamma \cdot \max_{a \in \mathcal{A}_{o_{t+1}}} Q(o_{t+1}, a)$$

- Принцип аналогичный Q-routing:

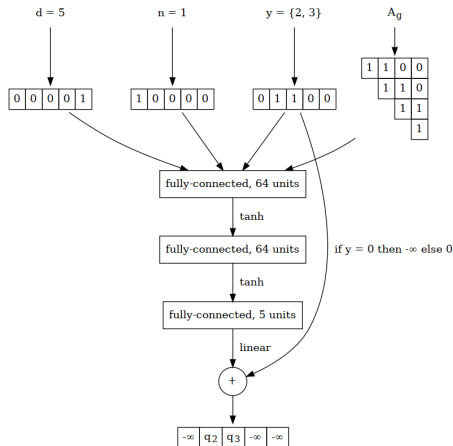
$$Q_x(d, y) \leftarrow (t_{finish} - t_{start}) + \max_{z \in \{V | (y, z) \in E\}} Q_y(d, z)$$

Что было сделано

Алгоритм DQN-routing⁷

- Объединяет link-state протокол с алгоритмом Q-routing
- Вход нейросети:
 - Номер текущего узла
 - Номер узла назначения
 - Узлы-соседи
 - Матрица смежности графа

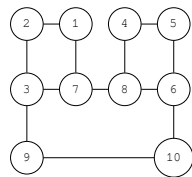
Был протестирован в сеттинге компьютерных и конвейерных сетей



⁷“Multi-agent deep learning for simultaneous optimization for time and energy in distributed routing system”.

Эксперименты в модели компьютерной сети

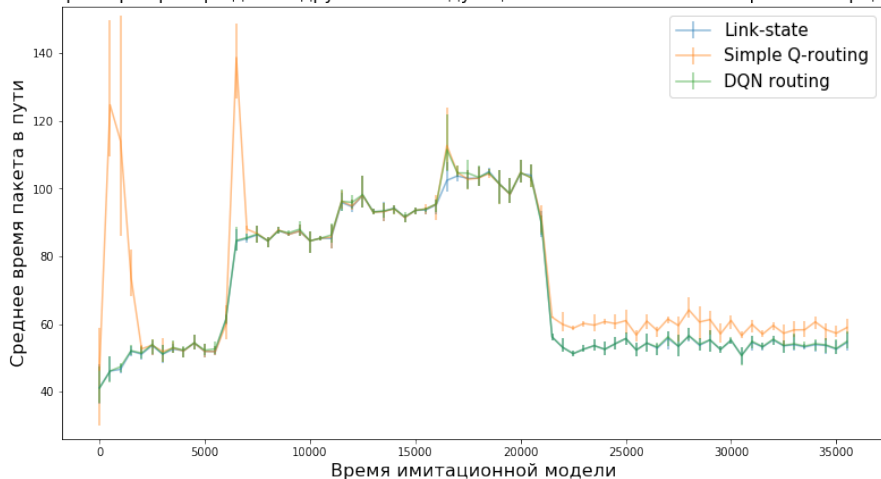
- Запуск экспериментов в симуляторе
- Сравнение архитектур между собой и сравнение с алгоритмами link-state и Q-routing
- Служебные сообщения доставляются бесплатно



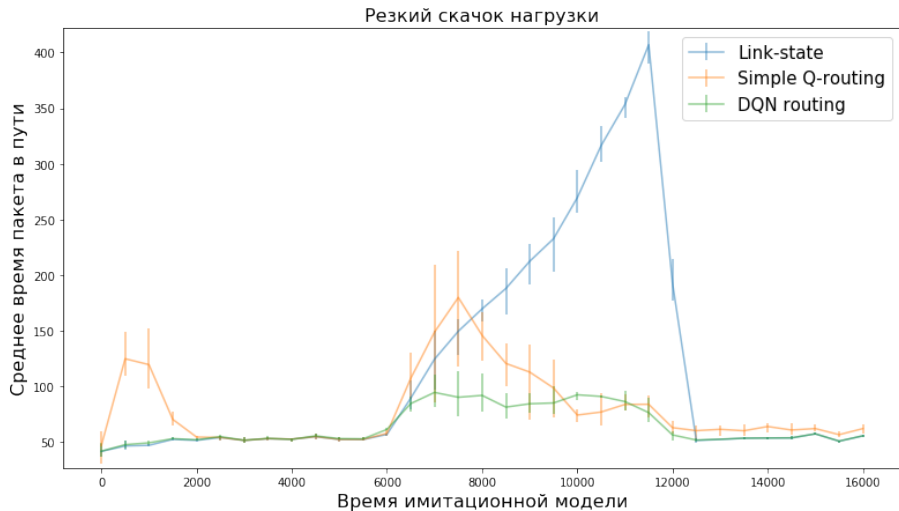
Топология сети для тестов

Изменение топологии сети: сравнение с link-state и Q-routing

Обрыв трех ребер одно за другим с последующим восстановлением в обратном порядке

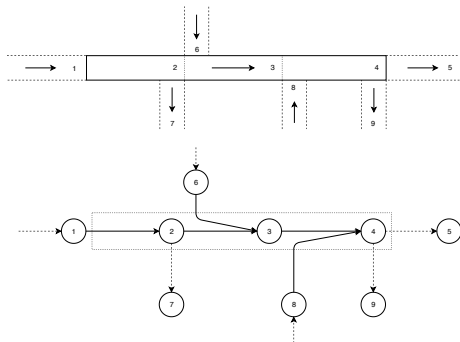


Изменение нагрузки: сравнение с link-state и Q-routing



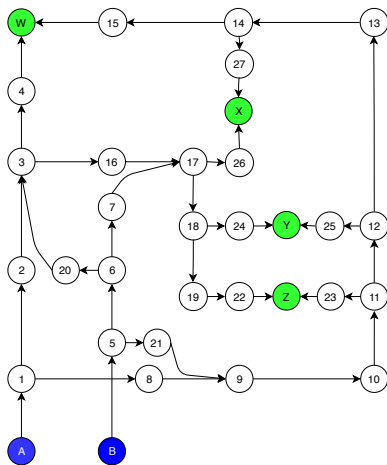
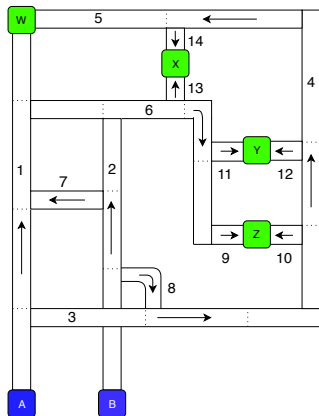
Эксперименты в модели системы багажных конвейеров

- Конвейерная сеть моделируется как ориентированный граф
- Каждая секция конвейера – отдельная вершина
- В оптимизируемую функцию включено *энергопотребление*
- На вход нейросети дополнительно подается информация о состоянии конвейеров
- Важность экономии энергии регулируется параметром α



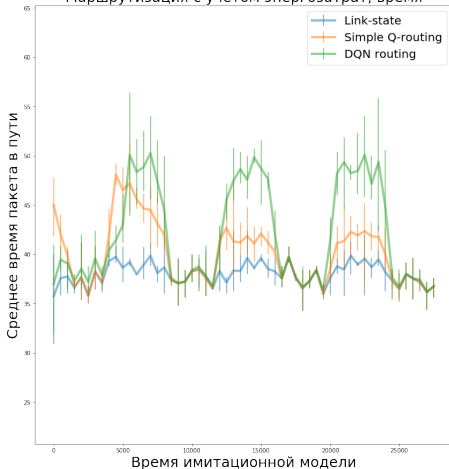
Участок конвейерной сети и соответствующий участок графа

Модель конвейерной системы для проведения экспериментов

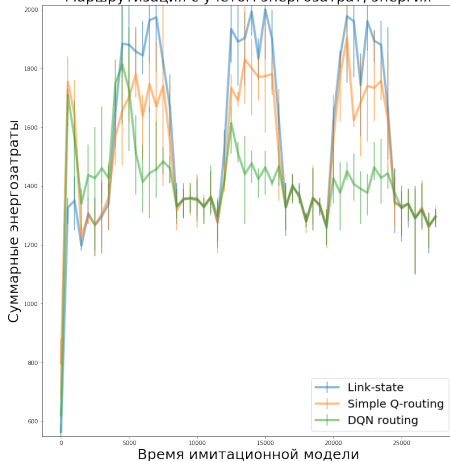


Неравномерный поток до выходных вершин: сравнение с link-state и Q-routing

Маршрутизация с учетом энергозатрат, время

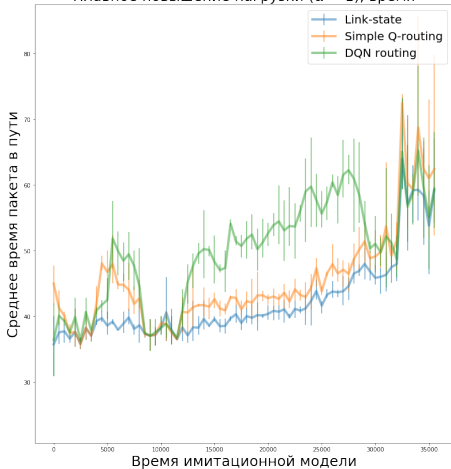


Маршрутизация с учетом энергозатрат, энергия

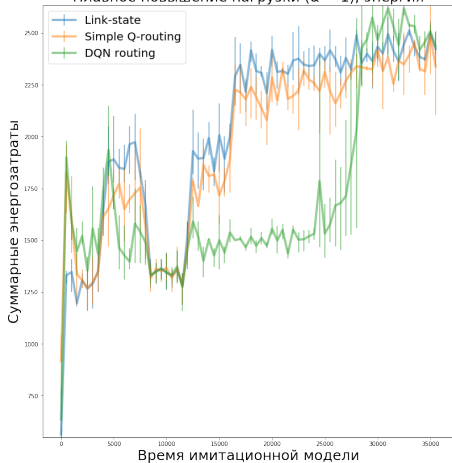


Плавное повышение нагрузки: $\alpha = 1$

Плавное повышение нагрузки ($\alpha = 1$), время

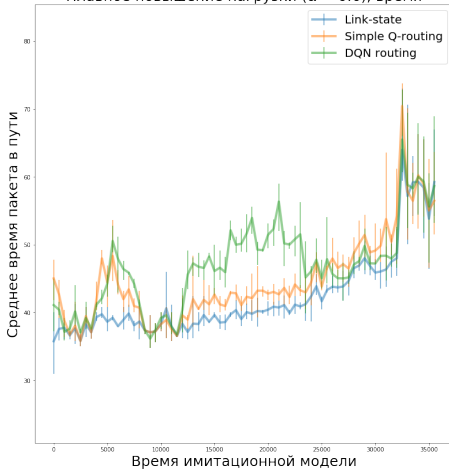


Плавное повышение нагрузки ($\alpha = 1$), энергия

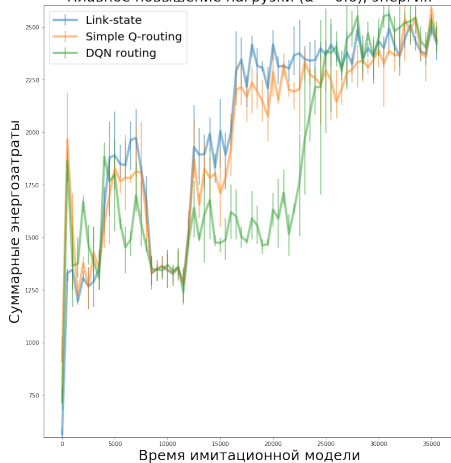


Плавное повышение нагрузки: $\alpha = 0.6$

Плавное повышение нагрузки ($\alpha = 0.6$), время



Плавное повышение нагрузки ($\alpha = 0.6$), энергия



- Плюсы

- Адаптация под изменения трафика
- Адаптация после поломок
- Оптимизация времени доставки и энергопотребления с заданным приоритетом

- Минусы

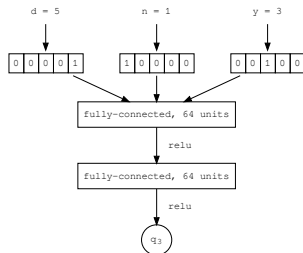
- Размер выходного слоя линейно зависит от размера графа
- Размер входного слоя квадратично зависит от размера графа
- Требуется предварительного обучения с учителем

Идеи усовершенствования алгоритма

- Предсказание Q-функции отдельно для каждого исходящего ребра
- Использование графовых эмбедингов
- Зональная маршрутизация
- Построение глобальной графовой нейронной сети (GG-NN)

Предсказание Q-функции для соседей по отдельности

- Только небольшое количество узлов в сети являются нашими соседями по исходящим ребрам
- Можем предсказывать значение Q-функции для каждого из них в отдельности, подавая номер каждого из них на вход
- Избавляемся от линейной зависимости размера выходного слоя от размера графа
- Качество работы остается прежним



Использование графовых эмбедингов: идея

- Вместо кодирования меток узлов унитарным кодом использовать их отображения в векторное пространство фиксированной размерности
- Отказ от подачи на вход матрицы смежности
 - Вместо этого пересчитывать эмбединги при изменении топологии
 - Эмбединги косвенно передадут информацию о топологии

Зональная маршрутизация: идея

- Ограничиваем рассматриваемый граф только узлами не далее k ребер от данного
- Если узел назначения находится вне текущей зоны — делаем запрос к узлам на границе зоны.

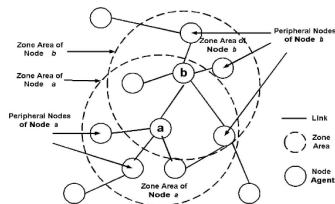
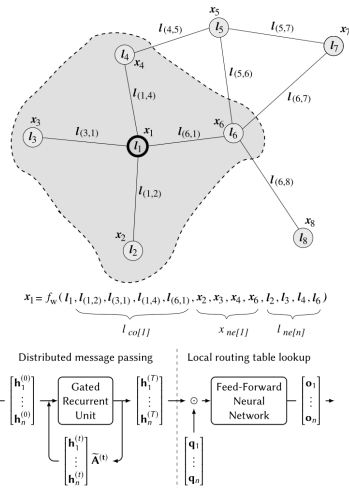


Figure 4. Zone boundary of node agent a and node agent b .

Глобальная графовая нейронная сеть: идея

- Рассматриваем весь граф с состояниями ребер и узлов как вход для графовой нейронной сети (GG-NN)
- Считается распределенно на физических узлах системы
- Промежуточные состояния и их градиенты передаются между узлами по сети
 - Применяется для компьютерных сетей с обучением с учителем⁸
 - Добавим обучение с подкреплением во время работы



⁸Geyer, Carle.

- Проведение большего числа экспериментов на различных топологиях
- Динамическое управление скоростями конвейеров

Спасибо за внимание!