

РАСПРЕДЕЛЕНИЕ ПОСТАВОК СВЕТЛЫХ НЕФТЕПРОДУКТОВ НА ОСНОВЕ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Козляк Р.В. (ИТМО)

Научный руководитель – кандидат технических наук Митягин С.А.
(ИТМО)

Введение.

В данном исследовании рассматривается применение методов обучения с подкреплением (Reinforcement Learning, RL) к задаче перераспределения поставок на нефтяные станции (Petroleum Station Reshipment Problem, PSRP).

Задача пополнения запасов нефтяных станций PSRP заключается в доставке различных нефтепродуктов из центрального хранилища в сеть автозаправочных станций в рамках многодневного планирования. При этом необходимо учитывать ограничения по вместимости резервуаров станций и отсекам транспортных средств, а также минимизировать транспортные издержки и повышать операционную эффективность. Существующие подходы, основанные преимущественно на точных или эвристических методах оптимизации, как правило, решают эту задачу в однодневном формате. Однако исследования [1] показывают, что многодневное планирование позволяет находить более эффективные решения, предотвращая ненужные ежедневные поставки.

С этой целью был использован алгоритм Recurrent Proximal Policy Optimization (RecurrentPPO) для решения многопериодной версии PSRP с целью оценки его эффективности в оптимизации стратегий перераспределения поставок во времени. В качестве базового подхода применялся точный солвер CPSAT, а эксперименты проводились на небольшом синтетическом датасете.

Основная часть.

Точные методы [1], применяемые для решения задачи PSRP, имеют ряд ограничений, таких как необходимость перерасчета решения при изменении условий. Кроме того, эти методы плохо масштабируются с увеличением числа переменных, что особенно усложняет задачу в многодневном горизонте планирования. Подходы, основанные на эвристиках [2], [3], [4], также требуют пересчета задачи, хотя выполняют его быстрее, чем точные методы.

Предполагается, что использование методов обучения с подкреплением позволит избежать необходимости перерасчета всей задачи, так как обученный агент сможет динамически предсказывать действия на основе изменяющегося состояния системы станций. Данное исследование направлено на формулирование задачи PSRP в рамках RL-подхода и предоставление первичного решения для дальнейших экспериментов. Среда основана на математическом описании задачи, представленном в работе [1]. Таким образом, одним из ключевых результатов нашей работы является создание среды для обучения RL-агента. Она предоставляет основу для моделирования процессов, происходящих при транспортировке топлива на заправочные станции, что способствует дальнейшему развитию RL-методов для решения задачи PSRP. Среда реализована в Gymnasium и совместима с библиотекой Stable Baselines 3, что обеспечивает поддержку универсального интерфейса.

Выводы. Полученные результаты продемонстрировали способность модели RL адаптироваться к ряду требований, предъявляемых к данной задаче в условиях ограниченного масштаба. Кроме того, в рамках работы предложена среда для обучения алгоритмов RL, предназначенная для решения задачи PSRP, что открывает возможности для дальнейших экспериментов в области оптимизации транспортных процессов.

Список использованных источников:

1. Boers L. et al. The multi-period petrol station replenishment problem: Formulation and solution methods //Computational Logistics: 11th International Conference, ICCL 2020, Enschede, The Netherlands, September 28–30, 2020, Proceedings 11. – Springer International Publishing, 2020. – C. 600-615.
2. Cavaliere F., Bendotti E., Fischetti M. An integrated local-search/set-partitioning refinement heuristic for the capacitated vehicle routing problem //Mathematical Programming Computation. – 2022. – T. 14. – №. 4. – C. 749-779.
3. Cornillier F., Boctor F., Renaud J. Heuristics for the multi-depot petrol station replenishment problem with time windows //European Journal of Operational Research. – 2012. – T. 220. – №. 2. – C. 361-369.
4. Shaw P. Using constraint programming and local search methods to solve vehicle routing problems //International conference on principles and practice of constraint programming. – Berlin, Heidelberg : Springer Berlin Heidelberg, 1998. – C. 417-431.