

# Сравнение алгоритмов управления сигналами светофоров в крупномасштабном сценарии моделирования движения транспортных средств

А.А. Агафонов  
Самарский национальный  
исследовательский университет им. акад.  
С.П. Королева  
Самара, Россия  
ant.agafonov@gmail.com

Е.Ю. Ефименко  
Самарский национальный  
исследовательский университет им. акад.  
С.П. Королева  
Самара, Россия  
evgeniaefimenko27@gmail.com

**Аннотация**—Задача управления сигналами светофоров остается важной задачей, решаемой в интеллектуальных транспортных системах. Развитие методов машинного обучения привело к активному развитию адаптивных методов управления сигналами светофоров, основанных на обучении с подкреплением. В то же время, сравнение классических подходов к управлению с методами на основе обучения с подкреплением в существующих работах проводится либо на синтетических сценариях, либо на сценариях, включающих малое количество перекрестков. В данной работе мы проводим сравнение алгоритмов управления сигналами светофоров в крупномасштабном сценарии моделирования движения транспортных средств в системе моделирования SUMO. Для проведения исследований был разработан сценарий моделирования на основе сценария SUMO «TAPAS Cologne».

**Ключевые слова**— обучение с подкреплением, детерминированная модель, управление сигналами светофоров, SUMO.

## 1. ВВЕДЕНИЕ

Задача оптимального управления транспортными потоками является одной из ключевых задач в современных транспортных системах. Управление транспортными потоками позволит сократить транспортные расходы на совершение поездок в сети, снизить потребление топлива и уменьшить уровень загрязнения окружающей среды.

Развитие информационно-коммуникационных технологий, интернета вещей (IoT), подключенных и автономных транспортных средств привело к увеличению объема данных, которые могут использоваться для решения задачи адаптивного управления сигналами светофоров. В частности, активно развиваются методы управления, основанные на алгоритмах обучения с подкреплением.

Адаптивное управление транспортными потоками путем управления сигналами светофоров основывается на информации, полученной от детекторов транспортных средств, видеокамер, подключенных транспортных средств и др. для выбора следующей фазы светофора. Обзор и сравнение классических методов управления (SOTL, MaxPressure) на основе детерминированных моделей было проведено в [1, 2]. В работе [3] авторы предложили адаптивный метод управления на основе детерминированной прогнозной модели. Предлагаемый метод основывается на выборе фазы светофорного цикла, прогнозируемый поток для которой оказывается максимален. В последние два десятилетия наибольший интерес представляет разработка методов адаптивного

управления на основе методов машинного обучения, и, в частности, на основе обучения с подкреплением [4-6]. Однако эти методы могут работать нестабильно в сложных сценариях с зашумленными данными и слишком чувствительны к гиперпараметрам [7]. Более того, исследования разработанных алгоритмов часто проводятся на синтетических сценариях моделирования с перекрестками типовой структуры, либо рассматривают малое число перекрестков в реальных сценариях.

В данной работе было проведено сравнение алгоритмов адаптивного управления сигналами светофоров на разработанном крупномасштабном сценарии моделирования движения транспортных средств. В следующем разделе приводится краткое описание сравниваемых алгоритмов. В разделе 3 представлено описание сценария, постановка и результаты экспериментов. Заключение работы представлено в разделе 4.

## 2. АЛГОРИТМЫ АДАПТИВНОГО УПРАВЛЕНИЯ СИГНАЛАМИ СВЕТОФОРОВ

### А. Детерминированные методы

Детерминированные методы используют заранее определенные правила для выбора следующей фазы светофорного цикла.

- SOTL [1] оценивает зеленые «запросы» от текущей фазы и других конкурирующих фаз, а затем решает, сохранить или изменить текущую фазу. В частности, сохранение текущей фазы светофора может определяться наличием непрерывного потока транспортных средств.
- MaxPressure [2] стремится сбалансировать длину очереди между соседними перекрестками за счет минимизации «давления» фаз на перекресток.
- MaxFlow [3] основывается на выборе фазы светофорного цикла, прогнозируемый поток для которой оказывается максимален.

### Б. Методы на основе обучения с подкреплением

Обозначим через  $S$  множество состояний некоторого объекта. Пусть  $A$  обозначает множество возможных действий. Упрощенно задача состоит в построении отображения  $S \rightarrow A$ , которое для каждого конкретного состояния отображает действие, которое необходимо совершить в указанном состоянии в определенном смысле оптимальным образом. Для решения этой задачи обычно используют подход на основе марковского

процесса принятия решений  $\langle S, A, P, R, \gamma \rangle$ , где  $P$  определяет вероятности перехода между состояниями, величина  $R$  характеризует «награду»,  $\gamma$  – коэффициент дисконтирования. Задача обучения с подкреплением заключается в нахождении оптимальной политики  $\pi: S \rightarrow A$ , которая максимизирует награду  $R$ .

В качестве алгоритмов на основе обучения с подкреплением в работе рассматриваются:

- IDQN [8] – алгоритм на основе Q-обучения; каждый агент контролирует отдельный светофор и обучается независимо.
- IPPO [8] – алгоритм на основе оптимизации политики.

### 3. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Экспериментальные исследования разработанных алгоритмов были проведены в системе моделирования движения транспортных средств SUMO [9]. Для проведения исследований был разработан сценарий моделирования, основанный на сценарии движения SUMO «TAPAS Cologne» [10].

Сценарий содержит 2928 перекрестков различной конфигурации, в т.ч. 316 сигнализированных перекрестков, 5808 сегментов. Движение транспортных средств рассматривалось в утренний час пик в период с 6 до 8 утра. Дорожная сеть рассматриваемого сценария представлена на Рис. 1.



Рис. 1. Дорожная сеть сценария моделирования движения

Сравнение алгоритмов проводилось по двум метрикам: среднее время ожидания и среднее время движения. Среднее времени ожидания (в секундах) показывает среднее время, которое транспортные средства провели без движения на перекрестке, среднее время движения (в секундах) — это среднее время, затрачиваемое транспортными средствами на совершение поездки в сети. Результаты сравнения алгоритмов представлены в Таблице I.

### 4. ЗАКЛЮЧЕНИЕ

В работе проведено сравнение адаптивных детерминированных алгоритмов и алгоритмов на основе

обучения с подкреплением для решения задачи управления сигналами светофоров. Для проведения экспериментальных исследований был разработан крупномасштабный сценарий моделирования движения транспортных средств. Результаты показывают, что алгоритм IDQN на основе обучения с подкреплением, обученный независимо для каждого агента, превосходит остальные алгоритмы. Однако данный подход не является масштабируемым, что требует разработки новых подходов к управлению.

Таблица II. СРАВНЕНИЕ ЭФФЕКТИВНОСТИ УПРАВЛЕНИЯ ТРАНСПОРТНЫМИ ПОТОКАМИ

	Среднее время движения	Среднее время ожидания
IDQN	319,21	19,54
IPPO	371,32	67,28
SOTL	361,14	44,04
MaxPressure	334,8	28,59
MaxFlow	327,69	23,68

### БЛАГОДАРНОСТИ

Работа выполнена при поддержке Российского научного фонда (проект № 21-11-00321, <https://rscf.ru/en/project/21-11-00321/>).

### ЛИТЕРАТУРА

- [1] Papageorgiou, M. Review of road traffic control strategies / M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, Y. Wang // Proceedings of the IEEE. – 2003. – Vol. 91(12). – P. 2043-2065.
- [2] Varaiya, P. The Max-Pressure Controller for Arbitrary Networks of Signalized Intersections / P. Varaiya // Advances in Dynamic Network Modeling in Complex Transportation Systems: Complex Networks and Dynamic Systems. – 2013. – P. 27-66.
- [3] Мясников, В.В. Детерминированная прогнозная модель управления сигналами светофоров в интеллектуальных транспортных и геoinформационных системах / В.В. Мясников, А.А. Агафонов, А.С. Юмаганов // Компьютерная оптика. – 2021. – Т. 45, № 6. – P. 917-925. DOI: 10.18287/2412-6179-CO-1031.
- [4] Yau, K.-L.A. A survey on Reinforcement learning models and algorithms for traffic signal control / K.-L.A. Yau, J. Qadir, H.L. Khoo, M.H. Ling, P. Komisarczuk // ACM Computing Surveys. – 2017. – Vol. 50(3).
- [5] Greguric, M. Application of Deep Reinforcement Learning in Traffic Signal Control: An Overview and Impact of Open Traffic Data / M. Greguric, M. Vujic, C. Alexopoulos, M. Miletic // Applied Sciences. – 2020. – Vol. 10(11). – P. 4011.
- [6] Qadri, S.S.S.M. State-of-art review of traffic signal control methods: challenges and opportunities / S.S.S.M. Qadri, M.A. Gökçe, E. Öner // European Transport Research Review. – 2020. – Vol. 12(1). – P. 55.
- [7] Genders, W. An Open-Source Framework for Adaptive Traffic Signal Control / W. Genders, S. Razavi // ArXiv preprint: 1909.00395, 2019.
- [8] Ault, J. Learning an Interpretable Traffic Signal Control Policy / J. Ault, J.P. Hanna, G. Sharon // ArXiv preprint: 1912.11023, 2020.
- [9] Lopez, P.A. Microscopic Traffic Simulation using SUMO / P.A. Lopez // 21st International Conference on Intelligent Transportation Systems (ITSC). – 2018. – P. 2575-2582.
- [10] TAPASCologne – SUMO Documentation [Electronic resource]. – Access mode: <https://sumo.dlr.de/docs/Data/Scenarios/TAPASCologne.html> (21.11.2021).