

УДК 004.896

АДАПТИВНОЕ УПРАВЛЕНИЕ СВЕТОФОРНЫМ ОБЪЕКТОМ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

© Остапенко П.В., Сапрыкин О.Н.

e-mail: pave-ostapenk@yandex.ru

*Самарский национальный исследовательский университет
имени академика С.П. Королёва, г. Самара, Российская Федерация*

Высокие темпы урбанизации и рост уровня автомобилизации бросают новые вызовы в области организации транспортной инфраструктуры городов. Среди прочих, можно указать следующие актуальные проблемы:

- увеличение транспортных расходов, связанных с дорожно-транспортными задачами, что негативно влияет на национальную производительность и конкурентоспособность;
- увеличение выбросов CO₂ от транспортных средств, из-за увеличения времени простоя;
- общественное недовольство по поводу отсутствия эффективного управления движением и увеличения время достижения пункта назначения поездки.

Один из способов решения данных проблем – интеллектуальное регулирование техническими средствами управления дорожным движением. В частности, в данной работе рассматривается адаптивное управление светофорными объектами. С появлением первых светофоров возникла потребность в оптимизации времени ожидания транспортных средств на перекрестке. Первые адаптивные методы регулирования светофора появились в начале 70-х годов [1]. Эти методы основаны на использовании датчиков движения, собирающих информацию о насыщенности потока на перекрестках. Системы SCOOTs, SCATS, RHODES и др. зарекомендовали себя во многих городах мира и по сей день используются для оптимизации дорожного движения.

Представленные выше методы требуют заранее определенной модели среды. То есть они работают эффективно, когда дорожное движение на перекрестке стабильно с течением времени. Но в реальных условиях, трафик имеет стохастическую природу. В связи с этим, для управления дорожным движением в последнее время стали использоваться методы машинного обучения, в частности – обучение с подкреплением. Методы, основанные на искусственном интеллекте, все чаще применяются для управления светофорами, благодаря своей возможности адаптироваться к изменениям трафика.

Для того, чтобы смоделировать перекресток, использован микросимулятор дорожного городского движения SUMO (Simulation of Urban MObility) [2]. Пакет SUMO позволяет строить различные типы дорожных сетей, добавлять автомобили различных типов, строить их маршруты, а также устанавливать светофорные объекты и датчики.

Для исследования выбран перекресток на пересечении улиц Партизанская и Аврора в г. Самара. На данном перекрестке наблюдается большое скопление транспортных средств, особенно в часы пик.

В исследовании использовался марковский алгоритм принятия решений Q-learning [3]. Для этого мы вводим функцию Q, отражающую ценность каждого возможного действия агента a (в нашем случае – светофора) для текущего состояния моделирования s , в котором он находится

$$Q(s, a) \quad (1)$$

Процесс обучения – итерационное уточнение функции Q на каждом шаге. Величина максимальной возможной награды на следующем шаге определяется как:

$$\max_a Q(s_{j+1}, a) \quad (2)$$

Состояние зависит от длины светофорной фазы, размера очереди и времени ожидания на перекрестке и имеет следующую зависимость:

$$(\text{lightphases}) * [(\text{queuesizes}) * (\text{waitingtimes})]^{(\text{edges})} \quad (3)$$

Величина награды, которую получит агент обозначим переменной r_t . Вознаграждение записывается по следующей формуле:

$$r = \sum_{\text{edge}i=1}^4 \beta_q (\text{queuesize}q)_i^{\theta q} + \beta_w (\text{waitingtime})_i^{\theta w} \quad (4)$$

С учётом дисконтирующего коэффициента γ , снижающего ценность будущих наград для светофора по сравнению с немедленными, была получена формула для функции Q.

Для того чтобы моделирование было максимально точным и приближённым к реальности, подсчитаны основные показатели исследуемого перекрёстка, такие как: длина цикла светофорного объекта, интенсивность движения транспортных средств на перекрёстке за 60 минут. Данные приведены к формату программы SUMO и записаны в файл модели.

В ходе работы рассмотрены современные методы адаптивного управления светофором. В программе SUMO создана модель перекрёстка улиц Авроры и Партизанской, для которого предложен метод оптимизации работы светофора на основе алгоритма Q-learning.

В дальнейшем планируется разработка нейросети, сравнение метода Q-learning с другими методами оптимизации и расширение моделируемой дорожной сети.

Библиографический список

1. Sims, A.G. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits [Текст]/ A.G. Sims, K.W. Dobinson// IEEE Transactions on vehicular technology, 1980 - 130–137p.
2. Krajzewicz, Recent development and applications of SUMO – Simulation of Urban MObility [Текст]/ Krajzewicz, Daniel, Erdmann, Jakob, Behrisch, Michael, and Bieker, Laura// International Journal On Advances in Systems and Measurements, 2012 -128–138p.
3. Arel, I. Reinforcement learning-based multi-agent system for network traffic signal control [Текст]/ I. Arel, C. Liu, T. Urbanik, A. Kohls// Intelligent Transport Systems, IET, 2010 -128–135p.