

# ФОРМИРОВАНИЕ АЛГОРИТМА УПРАВЛЕНИЯ АГЕНТА-ПЕРЕХВАТЧИКА В ГЕТЕРОГЕННОМ ТРАНСПОРТНОМ ПОТОКЕ

**Н.В. Быков**

*Российский университет транспорта (РУТ-МИИТ)*  
Россия, 127994, Москва, Образцова ул., 9с9  
E-mail: nik.bkv@gmail.com

**В.А. Кульметьева**

*Российский университет транспорта (РУТ-МИИТ)*  
Россия, 127994, Москва, Образцова ул., 9с9  
E-mail: varko2007@gmail.com

**Ключевые слова:** алгоритмы управления, транспортный поток, клеточные автоматы, обучение с подкреплением, DQN.

**Аннотация:** Внедрение беспилотных автомобилей несет в себе дополнительные угрозы, которые связаны с возможным взломом и захватом управления автомобилем. В этом случае важно обеспечить возможность перехвата взломанного беспилотного автомобиля. В работе для этого предлагается использовать агентов-перехватчиков, которые движутся в транспортном потоке вместе с целью. Транспортный поток, являющийся средой, описывается в рамках модели Нагеля-Шрекенберга, основанной на клеточных автоматах. Используются как прямые подходы к формированию алгоритмов управления агента-перехватчика, так и подход глубокого машинного обучения, основанный на алгоритме DQN.

## 1. Введение

Прогресс в развитии индивидуального транспорта характеризуется увеличением количества разработок в области беспилотных автомобильных транспортных средств (БАТС), предназначенных для автоматического движения по автомобильным дорогам. БАТС могут значительно снизить частоту дорожно-транспортных происшествий и уменьшают заторы на дорогах, поскольку обеспечивают достаточно большую точность скоординированного управления по сравнению с управляемыми человеком автомобильными транспортными средствами (АТС).

В последнее десятилетие разработка беспилотных транспортных средств переживает технологический бум в автомобильной отрасли всех ведущих стран мира. Наиболее активно работы по созданию беспилотных транспортных средств ведутся в США, Германии, Японии, Китае, Великобритании, Швеции, Франции, Корее. Многие из этих стран уже допустили БАТС разной степени автономности на дороги общего пользования.

В России беспилотные автомобили находятся на стадии тестирования и разработки. В 2018 году правительство России утвердило законопроект о беспилотных автомобилях, который предусматривает возможность экспериментальной эксплуатации таких автомобилей на дорогах страны. В середине марта 2022 года Правительство РФ

приняло постановление «Об установлении экспериментального правового режима в сфере цифровых инноваций и утверждении Программы экспериментального правового режима в сфере цифровых инноваций по эксплуатации высокоавтоматизированных транспортных средств». Это постановление говорит о том, что беспилотные автомобили станут частью городской инфраструктуры к 2025 году при условии успешного прохождения испытаний. Также осенью 2022 года Росстандарт утвердил восемь ГОСТов для беспилотных автомобилей с искусственным интеллектом, эти ГОСТы вступили в силу с 1 января 2023. В июле 2023 года на трассе М-11 «Нева» запустили первые беспилотные коммерческие перевозки. Также в настоящее время в России проводятся испытания беспилотных автомобилей на закрытых территориях, например, на территории инновационного центра «Сколково», в Набережных Челнах (полигон КамАЗ) и в Иннополисе (город-спутник Казани). Что касается правовой составляющей, то сейчас правительство РФ разрабатывает закон «О высокоавтоматизированных транспортных средствах». В будущем он будет регламентировать все правовые отношения, касающиеся беспилотных автомобилей.

Несмотря на многочисленные испытания, необходимо в первую очередь решить вопросы, связанные с технической безопасностью и обеспечением качественной связи между автомобилем и системами управления дорожным движением. Это создает необходимость создания новых математических и компьютерных моделей, описывающих совместное движение БАТС и управляемых человеком АТС.

Компании могут успешно тестировать автономные автомобили в закрытых помещениях, но, чтобы оценить реальный уровень безопасности, производителям придется развернуть парк беспилотных автомобилей на дорогах общего пользования. Это означает, что большое число людей фактически станет частью большого эксперимента.

Например, даже если беспилотный транспорт в перспективе можно сделать очень надежным, он не будет на 100% защищен от угрозы взлома программного обеспечения (ПО) извне. Такие эксперименты уже проводились в США (хакеры успешно отключили тормозную систему у БАТС). Следовательно, БАТС может потерять управление в связи с выходом из строя одной из внутренних систем, или же ввиду захвата злоумышленником. В этом случае необходимо учитывать влияние потенциально опасного БАТС на транспортный поток, и обезвредить это неисправное АТС и удалить его из транспортного потока.

## 2. Модели и методы

### 2.1. Модель среды

В моделях на клеточных автоматах (КА) дорога разбивается на малые сегменты (клетки), дискретным считается и время. Преимущество данной модели состоит в ее простоте, и она не требует значительных вычислительных ресурсов для описания транспортного потока, поскольку поведение автомобиля описано с использованием терминологии дискретных систем.

Основополагающей в ряду моделей на клеточных автоматах (КА) является модель Нагеля и Шрекенберга [1, 2], в которой скорость АТС представлена несколькими дискретными значениями.

Основное отличие БАТС от управляемого человеком АТС состоит во времени реакции на изменение поведения впереди едущего АТС. Эта реакция определяется типами текущего и впереди движущегося АТС.

В гетерогенном транспортном потоке из двух типов АТС может быть четыре типа следования [3]:

- (a) управляемое человеком АТС – управляемое человеком АТС;
- (b) управляемое человеком АТС – БАТС;
- (c) БАТС – управляемое человеком АТС;
- (d) БАТС – БАТС.

Если АТС управляется человеком, то он, как правило, не знает, как управляется АТС впереди, поэтому можно считать, что в случаях (a) и (b) модель движения АТС будет одной и той же. Процесс следования БАТС за управляемым человеком АТС (случай (c)) определяется системой автоматического круиз-контроля. Существенное отличие случая (d) от остальных состоит в том, что группа из последовательно движущихся БАТС может образовывать кластер из нескольких БАТС, который перемещается как единое целое за счет обмена информацией между БАТС через интерфейс V2V, характерное время которого много меньше шага по времени в модели (обычно этот шаг составляет 1 с, характерное же время обмена информацией для БАТС порядка 0,1 с).

Традиционный подход КА для построения качественной модели гетерогенного трафика с учетом возможности образования кластеров на основе развития правил 184 и 240 для элементарных КА применен в [4].

В работе [5] было предложено поведение БАТС в рамках S-NFS модели. Подобное поведение включает в себя кластеризацию при отсутствии случайного торможения. Принципы поведения БАТС предложенные в данной статье легко применимы и для модели Нагеля-Шрекенберга.

## 2.2. Подход глубокого обучения с подкреплением

Игровое поле (Среда), представляет собой КА, основанный на модели Нагеля-Шрекенберга. Так как для успешного осуществления перехвата необходимо наличие более чем одной полосы движения, вводятся правила перестроения АТС из одной полосы в другую.

Агент-перехватчик (Игрок) – это алгоритм обучения с подкреплением, который взаимодействует с окружающей средой посредством выбираемых действий. Среда реагирует на действия агента изменением своего состояния и подачей сигнала подкрепления.

Ввиду того, что Игрок является участником транспортного потока, его действия аналогичны действиям частиц в потоке: агент может двигаться вперед и менять полосу движения. Для успешного перехвата агент может двигаться быстрее остальных участников потока. Смена полосы и изменение скорости, как и в случае с АТС, происходит последовательно за один временной шаг.

Игрок получает награду, если он успешно перехватывает проблемного участника движения (Цель). После осуществления перехвата Игроку назначается новая Цель, и игра возобновляется в той же среде. Если действие, выбранное Игроком, приводит к столкновению – Игрок получает штраф и игра заканчивается. Игрок вознаграждается, если выбранное им действие уменьшает расстояние до Цели. Это расстояние рассчитывается так, как если бы топология дороги представляла собой бесконечную прямую линию, а не окружность.

Для реализации механизма взаимодействия Игрока и Среды используется алгоритм Deep Q Network (DQN) [6]. Этот алгоритм относится к value-based RL-методам, и в своей основе использует алгоритм Q-обучения. В классической реализации Q-обучения формируется таблица из всех возможных состояний среды и всех возможных действий. Однако зачастую среда имеет слишком много состояний и действий, поэтому явно

создать Q-таблицу невозможно. Чтобы решить эту проблему, DQN использует нейронные сети для прогнозирования, а не для хранения Q-значений функции.

DQN обучает две модели. Первая модель (policy model) соответствует каждому действию агента. Вторая модель (target model) обновляется в соответствии с весами политики каждые  $k$  действий. Это связано с тем, что изначально все, от весов политики до действий агента, является случайным. Из-за этой случайности модели сложнее выполнять правильные действия, но, когда есть другая модель, которая использует знания, полученные первой моделью, каждые  $k$  действий, можно добиться некоторой степени согласованности.

Начальное состояние. При обучении перехватчика использовались следующие значения параметров Среды:  $p_1 = 0.3$ ;  $p_2 = 0.7$ ;  $p_3 = 0.99$ ;  $G = 15$ . Число полос – 3. Число клеток в каждой полосе  $n = 1000$ , число шагов по времени на установление  $T_s = 1000$ . Плотность потока равна 0,5, транспортные средства распределены в системе равномерно. Доля человекоуправляемых АТС в транспортном потоке - 25%. Размер кластера составлял  $S = 3$ .

Скорость всех транспортных средств в начальный момент равна нулю. После достижения транспортным потоком установившегося состояния агентом-перехватчиком становится любое транспортное средство. Цель выбирается случайным образом на расстоянии  $0.25n$ .

Архитектура DQN. В ходе исследования применялись как сложные модели, так и простые. И после всех поисков выбор был сделан в пользу архитектуры с одним сверточным слоем с 32 фильтрами, размером пакета 128, тремя плотными (полносвязанными) слоями по 32 узла каждый.

Входной слой: входная форма такая же, как форма массива, представляющего игровое поле (количество полос \* количество клеток).

Сверточные слои: один слой Conv2D с 32 фильтрами размером  $2*2$ .

Отсев: 20%

Сглаживание: преобразование выходных данных сверточного слоя из 2D в 1D-массив.

Плотные (полносвязанные) слои: три плотных слоя имеют по 32 узла каждый.

Выходной слой: выходной слой содержит 4 выходных узла, каждый узел представляет действие [ускориться, замедлиться, повернуть налево, повернуть направо].

### 3. Заключение

Произведено моделирование агента-перехватчика, как с использованием математического моделирования, так и с помощью нейросетевого подхода. По результатам исследования было изучено влияние агента-перехватчика на характеристики транспортного потока.

Предложенный метод глубокого обучения показывает удовлетворительные результаты. Полученные результаты позволяют решить задачу управления агентом в случае перехвата беспилотных транспортных средств, и при соответствующем обеспечении уровня плотности транспортных средств, обеспечить возможность нейтрализации потенциально враждебного или вышедшего из строя беспилотного транспортного средства.

В качестве дальнейших приоритетных направлений исследований и развития созданной системы можно выделить следующие:

- реализация многомерной нейросетевой модели перехватчика;

- добавление в систему противодействующего агента с необходимостью его обнаружения перехватчиком.  
Исследование поддержано грантом РФФ, проект № 24-21-00306.

## Список литературы

1. Nagel K., Schreckenberg M. A cellular automaton model for freeway traffic // Journal de Physique I. 1992. Vol. 2 (12). P. 2221-2229.
2. Schreckenberg M., Schadschneider A., Nagel K., Ito N. Discrete stochastic models for traffic flow // Physical Review E. 1995. Vol. 51 (4). P. 2939.
3. Jiang Y., Wang S., Yao Z., Zhao B., Wang Y. A cellular automata model for mixed traffic flow considering the driving behavior of connected automated vehicle platoons // Physica A: Statistical Mechanics and its Applications. 2021. Vol. 582. P. 126262.
4. Bykov N.V. Cellular Automata Simulation of Heterogeneous Freeway Traffic Flow in a Smart City // Proceedings of the 2022 4th International Youth Conference on Radio Electronics, Electrical and Power Engineering. REEPE 2022. P. 1-5.
5. Masaka J., Sueyoshi F., Hossain M.A., Utsumi S., Tanimoto J. Can the introduction of CAVs mitigate social dilemmas causing traffic jams on highways? // Physics Open. 2023. Vol. 17. P. 100176.
6. Mnih V., Kavukcuoglu K., Silver D., Graves A., Antonoglou I., Wierstra D. and Riedmiller M. Playing atari with deep reinforcement learning. 2013. arXiv preprint arXiv:1312.5602.