

# Online Operations of Automated Electric Taxi Fleets: An Advisor-student Reinforcement Learning Framework

Xindi TANG<sup>1</sup>, Meng LI<sup>2</sup>, Xi LIN<sup>3</sup> and Fang HE<sup>4</sup>

<sup>1</sup>*Department of Industrial Engineering, Tsinghua University, Beijing 100084, PR China.  
E-mail:txd16@mails.tsinghua.edu.cn*

<sup>2</sup>*Department of Civil Engineering, Tsinghua University, Beijing 100084, PR China.  
E-mail: mengli@tsinghua.edu.cn*

<sup>3</sup>*Department of Civil Engineering, Tsinghua University, Beijing 100084, PR China.  
E-mail: ghnttaifeng@126.com*

<sup>4</sup>*Department of Industrial Engineering, Tsinghua University, Beijing 100084, PR China.  
E-mail: fanghe@mail.tsinghua.edu.cn*

## Keywords

automated vehicle, electric taxi, reinforcement learning, online operation

## Introduction

With the advancement of technology and the appeal to future transportation system construction, automation and electrification have become inevitable trends in the development of intelligent vehicles. Autonomous vehicles liberate the drivers' workforce and possess the potential to be controlled in a centralized manner for the system-level benefits, while electric vehicles are environmentally friendly in both pollution discharge and noise making. It is envisioned that automated electric taxis will play a vital role in future transportation systems for providing customized travel service. Stakeholders of travel service industry are adapting to this coming new trend around the world. Waymo, a company developing autonomous vehicle, has taken a critical step in the commercialization of autonomous vehicles in the ride-hail service by allowing riders in greater Phoenix to book and pay for a ride in a driverless taxi using Waymo's technology. Uber, a ride-hail company, restarted autonomous vehicle testing on public roads in Pittsburgh. It is reported that all taxis in Beijing will be replaced by electric vehicles with fast-battery-swapping or fast-charging in the next 2-3 years (Sohu, 2019). City Shenzhen in China mandatorily require newly registered ride-hailing vehicles to be electric vehicles (Bituauto, 2019).

This paper aims at solving the online operation problem of electric automated taxi fleets. This is a novel topic about fleet operation, which is not well-solved, while some researchers have conducted preliminary exploration in the existing literature. Four categories of methods are employed. The first one builds models and uses exact optimization to achieve fleet control (Ramezani and Nourinejad, 2017; Billhardt and Bajo, 2019). The second type uses knowledge of markov property. It models the dispatching problem as dynamic programming and solves by approximate dynamic programming (Godfrey and Powell, 2002a; Godfrey and Powell, 2002b). The third category takes advantage of reinforcement learning and intends to tackle large scale problem (Xu et al., 2018; Jin et al., 2019; Ke et al., 2019). The last one utilizes rule-based principles or simulation to manage mobility services (Ota et al., 2016).

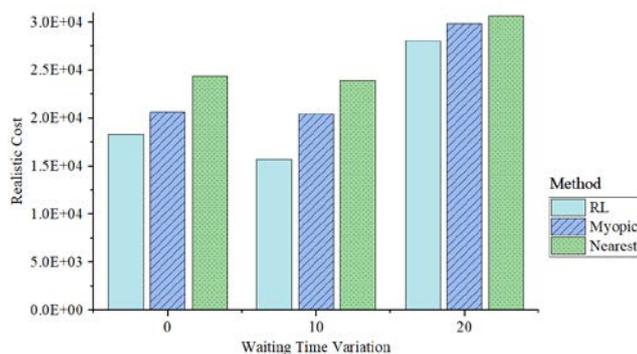
To verify the effectiveness of proposed method, we first design a hypothetical network with two nodes and two origin-destination (OD) pairs. The demands in one OD pair are relatively stable and in another OD pair are with peak-hour characteristic, and in such setting the spatial and temporal heterogeneity can be captured. We compare our framework with a myopic assignment model treated as a benchmark. The results show that the framework

outperforms the myopic model in realistic cost where the AETs can be dispatched to the zone with peak-hour demands in advance, which greatly enhance the performance of online operation system. Further experiments are conducted with a realistic road network in Tongzhou, Beijing, where the parameters of all 408 electric taxis are borrowed from the true dataset. Extensive managerial insights are drawn from the tests accordingly.

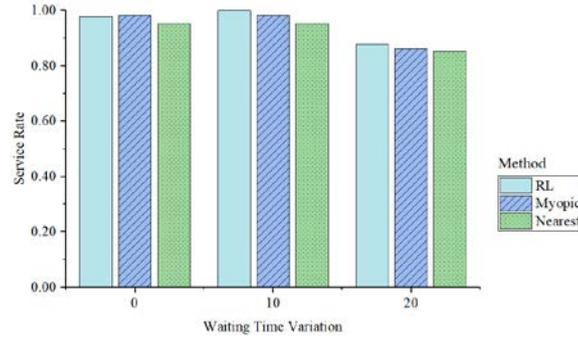
## Methodology

This study focuses on proposing a methodological framework to solve the online operation problem of electric automated taxi fleets. Specifically, the operation area is divided into a set of hexagons where the demands are generated stochastically, and certain number of electric automated taxis are distributed initially. A centralized control system is established to manage the whole taxi fleet. We discretize the time horizon into a set of operation intervals, and in each interval the taxis can be assigned to serve the demands generated in the corresponding regions, dispatched to other areas to eliminate supply-demand imbalance and forced to travel to charging stations to get refueled. In each operation interval, our framework includes two stage. At the first stage, we decide the number of taxis to serve demands, to be dispatched and to be recharged in each hexagon; this task is fulfilled with a reinforcement learning framework such that the information of current states, e.g., the distribution of demands, the distribution of vehicles, state of charge (SOC) of taxis, occupancy of charging stations, are incorporated to approximate the future system costs for aiding the current decision-making. At the second stage, a combinatorial optimization model is proposed to precisely decide the action of the each taxi. Each taxi will be assigned to a specific customer, dispatched to a specific area, or forced to a specific charging station and charge an amount of electricity. The results from the first stage acts as a reference such that the action obtained by the combinatorial optimization model generates similar number of taxis in serving customers, dispatching and recharging in each hexagon. In this way, the optimization model can avoid myopic decisions in that it considers the impact of current actions on future system states. Two system targets, i.e., the minimization of total system cost (including customer waiting, customer drop and taxi operating) and the maximization of agency profit, are considered in this study. Moreover, once the training is done, the optimization in both stages can be accomplished instantaneously, which is suitable for the implementation of online operation where timely decisions are required.

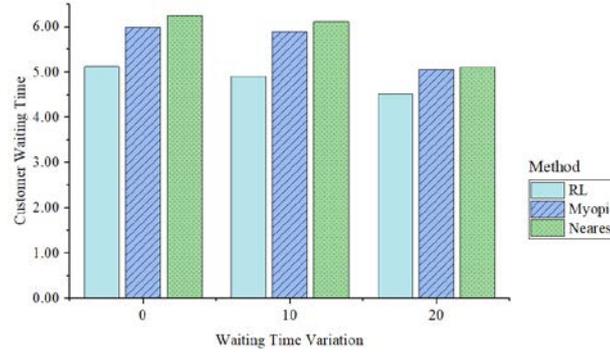
## Results



(a)



(b)



(c)

Fig. Indexes under different customer waiting variation. (a) Realistic Cost. (b) Service Rate. (c) Average Waiting Time.

The realized costs of RL are relatively lower than the values of MY and ND under various customer waiting time variance, attaining 5%-15% improvement compared to myopic model and 10%-35% improvement compared to nearest distance model. The results demonstrate that the RL can reduce system cost by dispatching under diverse scenarios. At the meantime, RL maintains a relatively satisfactory customer serving rate. By means of beforehand dispatching, the system with RL achieves lowest customer waiting time compared to MY and ND, validating effectiveness of such method.

## Conclusion

We investigate online dispatching and charging problem of large-scale electric automated taxi fleets, which is challenging under various stochasticity. A two-stage framework is proposed, combining reinforcement learning and combinatorial optimization. The results show that our framework outperforms benchmarks in multiple indexes, i.e., realistic cost, serving rate and customer waiting time. It has capability to dispatch in advance to eliminate hysteresis of handling imbalance, smooth charging demand to eliminate waiting queue at stations and achieve instant decision procedure. Further research can incorporate congestion into travel time among areas. Another direction is to synergistically optimize charging station planning and AETs' operation.

## References

- Billhardt, H., F. A. O. S. P. J. and Bajo, J. (2019). Taxi dispatching strategies with compensations. *Expert Systems with Applications*, 122:173–182.
- Bitauto (2019). Shenzhen electrification. <http://news.bitauto.com/hao/wenzhang/30917513>. Accessed October 7, 2019.
- Godfrey, G. A. and Powell, W. B. (2002a). An adaptive dynamic programming algorithm for dynamic fleet management, i: Single period travel times. *Transportation Science*, 36(1):21–39.
- Godfrey, G. A. and Powell, W. B. (2002b). An adaptive dynamic programming algorithm for dynamic fleet management, ii: Multiperiod travel times. *Transportation Science*, 36(1):40–54.
- Jin, J., Zhou, M., Zhang, W., Li, M., Guo, Z., Qin, Z., Jiao, Y., Tang, X., Wang, C., Wang, J., et al. (2019). Coride: Joint order dispatching and fleet management for multi-scale ride-hailing platforms. arXiv preprint arXiv:1905.11353.
- Ke, J., Xiao, F., Yang, H., and Ye, J. (2019). Optimizing online matching for ride-sourcing services with multi-agent deep reinforcement learning. arXiv preprint arXiv:1902.06228.
- Ota, M., Vo, H., Silva, C., and Freire, J. (2016). Stars: Simulating taxi ride sharing at scale. *IEEE Transactions on Big Data*, pages 1–1.
- Ramezani, M. and Nourinejad, M. (2017). Dynamic modeling and control of taxi services in large- scale urban networks: A macroscopic approach. *Transportation Research Part C: Emerging Technologies*.
- Sohu (2019). Beijing taxi electrification. [http://www.sohu.com/a/324305714\\_434156](http://www.sohu.com/a/324305714_434156). Accessed July 2, 2019.
- Xu, Z., Li, Z., Guan, Q., Zhang, D., Li, Q., Nan, J., Liu, C., Bian, W., and Ye, J. (2018). Large- scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 905–913. ACM.