Reinforcement Learning and Mechanism Design for Routing of Connected and Autonomous Vehicles

Doctoral Consortium

Behrad Koohy
Citizen-Centric Artificial Intelligence Systems, University of Southampton
Southampton, United Kingdom
behrad.koohy@soton.ac.uk

ABSTRACT
The data provided by Connected and Autonomous Vehicles (CAVs) is a powerful tool, providing insight into user incentives and preferences, and combined with existing road data sources, provides a number of new research avenues for intelligent traffic systems. In this paper, we propose the use of Reinforcement Learning (RL) for adaptive pricing of travel systems such as trains, buses and toll-road, in simulations which consider multiple transport providers and traffic management systems, known as the multi-market pricing problem. We also propose two research directions for this problem, the use of incentives when user preferences are included and development of detection and prevention of unintentional collusion between RL pricing agents.

KEYWORDS
Connected and Autonomous Vehicles; Vehicle Routing; Reinforcement Learning; Adaptive Pricing

ACM Reference Format:

1 INTRODUCTION
The emergence of Connected and Autonomous Vehicles (CAVs) offers significant opportunities for enhancing infrastructure efficiency and planning for new infrastructure. The data provided by CAVs, when combined with other data sources like induction loop sensors, offers significant potential for developing efficient, fair, and intelligent routing systems [3]. This increase in data from CAVs (combined with existing sources of data such as road-side sensors) have made it feasible to apply developments in Reinforcement Learning (RL) to intelligent traffic systems in contexts such as Traffic Signal Control (TSC) [13], electric vehicle charging [11] and autonomous vehicles [1]. One domain in intelligent traffic systems which could benefit from the ability to learn from complex, real-world data that is provided by RL is road and congestion pricing.

Congestion pricing models require a quantitative value for the marginal slowdown caused by the individual, which is often determined through stylized traffic models that assume deterministic conditions and may not reflect real-world scenarios [12]. A commonly proposed method of congestion pricing, known to increase traffic flow is to charge road users a cost proportional to the negative externalities (i.e. increased travel time) caused to other road users, known as a Pigouvian tax [10]. Sharon et al. [12] introduces this pricing strategy as a micro-tolling paradigm. RL has also found use in the context of road pricing, also known as congestion pricing; Mirzaei et al. [9] introduced enhanced delta-tolling, which uses RL to find the optimal parameters within the micro-tolling paradigm introduced by Sharon et al. Micro-tolling calculates the price for each link in a transport network by using two parameters: the known free-flow travel time and the current travel time [12]. This method of pricing improves on macro-tolling which is limited by assumptions of constant demand and capacity for links [12].

Existing approaches in the area of road pricing often view this problem from the perspective of a single travel provider or traffic management tool (i.e. TSC [2, 4, 8, 14] and micro-tolling [9, 12]. Furthermore, proposed adaptive solutions to this problem do not consider user preferences and intention, and are limited to demand-based adaptations. Our research abstracts the micro-tolling road routing problem, referred to as the multi-market pricing problem, and expands it to include multiple travel providers such as buses, trams, and toll roads. This approach provides a more comprehensive view of the road state and can inform optimal policy decisions for stakeholders such as users, local authorities, and government bodies. Additionally, it enables research into cooperation between systems and travel providers that traditionally may not interact.

In section 2, we introduce the Multi-Market Routing Problem (MMRP). In section 3, we discuss our previous work utilising RL for TSC, and how it relates to the MMRP. Section 4 introduces the in progress solution, and sections 5 & 6 are proposed areas of future work.

2 MULTI-MARKET ROUTING PROBLEM
The MMRP is defined by the tuple \((O, D, T, F, \pi(x), V, U)\), where \(o \in O\) is the set of origins, \(d \in D\) is the set of destinations, \(t_x \in T\) where \(T\) represents the set of travel providers and \(t_x\) is the provider which connects \(o_x\) to \(d_x\), \(F\) represents the set of free travel options which require a travel cost \(\epsilon\) to get to from all origins but do not cost the user anything to use, \(\pi(x)\) represents the set of pricing policies for the travel providers, \(v(x, c) \in V\) is the set of volume delay functions which determine the travel time on travel provider \(x\) where \(c\) is the current capacity and \(u \in U\) is the set of road users. Users \(u\) arrive at random to one of the origins \(o\) with a randomly chosen destination. We also provide each user \(u \in U\) with a maximum budget, where routes above this cost are infeasible,
We will also compare it against various agents such as deterministic policies (random pricing models, noisy pricing models, including adaptive noisy models). Additionally, we will test the effectiveness of the adaptive pricing policies under uncertainty by incorporating user behavior data and likely future road states into the pricing model.

5 INCENTIVES AND TRAFFIC MANAGEMENT

One area of future work is to introduce a mechanism which allows for the RL agents to offer incentives to users to take certain routes. In [16], an intention-aware routing algorithm is proposed with incentives, and the authors find that operational costs for a fleet of delivery vehicles is reduced by up to 30%. It is important to highlight that one of the methods of accessing road user intention is to incorporate data from CAVs. User preferences will be included when calculating incentives, including a dynamic response in decision-making to pricing policy changes.

An effective, equitable and reliable incentive system would allow for better usage of travel infrastructure, including public transport, where schemes such as “Park and Ride” [6] have been implemented in an attempt to reduce the number of vehicles on the road in city centres.

One further use case of incentives can be to manage environmental considerations in areas by balancing demand in specific areas at specific times (e.g. limiting air pollution around schools). This could be a valuable tool for local authorities, from known scenarios such as the school example, to situations such as accidents and disaster response where sections of the road network are required by emergency services.

6 UNINTENTIONAL COLLUSION AVOIDANCE

Kastius et al. [7] found that RL agents in oligopolies can be forced into collusion without direct communication. To prevent manipulation and protect the interests of road users, local authorities, and governments, it is essential to evaluate strategies that introduce collusion. This can be done by designing adaptive agents which have the objective of forcing collusion, and testing our proposed solution’s resilience to these strategies. Additionally, research should explore methods to identify unintentional collusion and adapt pricing strategies accordingly. The effects of these strategies should be evaluated in both early-stage and optimized agents, taking into account the non-stationarity of the problem, which can favor the antagonist agent. To ensure transparency and equity, this research will also draw on explainable RL techniques.

7 CONCLUSIONS

This extended abstract highlights multiple open research questions in the vehicle routing and congestion pricing problem, the feasibility of which is reliant on data from CAVs. The use of RL for this problem would introduce adaptive pricing strategies which can respond to traffic scenarios in a way that analytical solutions are not able to. The research also aims to investigate the use of mechanism design for incentives to manage congestion and the potential for collusion avoidance in RL pricing strategies, which could have wide-reaching implications beyond the transport sector.
ACKNOWLEDGMENTS
This research is supported by an ICASE studentship from EPSRC and Yunex Traffic, the EPSRC AutoTrust platform grant (EP/R029563/1), and an EPSRC Turing AI Acceleration Fellowship on Citizen-Centric AI Systems (EP/V022067/1).

REFERENCES


