Post-Wardropian traffic assignment framework for fleets of autonomously routing vehicles

Rafał Kucharski[1]

[1]Jagiellonian University, Poland

Short summary

Connected autonomous vehicles will be able to apply any arbitrary routing strategies, not only behavioural decisions of random utility maximisers but also collective ones given any objective. Such a route choice process will not fit into a well-established Wardropian paradigm of traffic assignment. To this end, we introduce a new *Post-Wardropian* framework to better model, understand, predict, and control fleet routing behaviour in future CAV cities.
Keywords: Autonomous Vehicles, Traffic Assignment, User Equilibrium.

1 Introduction

Let's consider the following scenario for the future urban mobility system:
Human drivers daily make individual routing decisions, and despite some inherent variability both in the supply and demand, their routing behaviour is somehow stable. Then, some CAV fleet providers offer autonomous routing services which, at first, are not different from human routing. Some drivers join the fleet for the comfort, and modal shifts are observed towards the fleet operator who increases his own market share. Beyond reaching the critical mass, a fleet operator realises the potential of controlling a fleet and assigning controlled vehicles to arbitrary routes for their own benefit. The operator develops a routing strategy optimising some latent goal (e.g., profit or market share). Starting from this point, the routing becomes potentially very different from the human routing. At any point, some new fleet operator may enter the system and compete with both the human drivers and against other fleets, with tailored pricing and routing strategies. Eventually, the city, observing undesired fleet actions and their negative impact, may start acting and regulating them. All the above inducing two kinds of network dynamics: i) joining and leaving the CAV fleets and ii) drivers adapting behaviour to the new situation
At the moment the unified framework to simulate such scenarios and understand their impact is missing. In this paper we reformulate the well-known day-to-day traffic assignment problem for the case of fleets of connected autonomous vehicles (CAVs).

2 Framework

Building on the classic day-to-day route choice problem, we adjust it for the future case of fleets of autonomous vehicles. We reformulate the standard definitions of choice set, utility and its expectation and the choice. Forming a complete agent-based framework, where humans follow behavioural principles, while the fleet(s) apply arbitrary strategies to reach their own objectives.

Wardropian user equilibria for perfectly rational non-collaborative individuals.

We begin with standard abstract formulation aligned with the classic Wardropian notion for individual human drivers. Let's consider some urban mobility network, represented with a directed weighted graph $G(N, A)$, where the demand of $\mathbf{Q}$ drivers want to travel between their origins and destinations at respective departure times $\mathbf{Q} \ni Q_i = \{o_i, d_i, \tau_i, \boldsymbol{\theta_i}\}$. Each individual driver i, based on personal traits, knowledge and experience ($\boldsymbol{\theta_i}$), selects a subjectively optimal path k_i from her own choice set of considered paths K_i and selects the path that maximises expected utility:

$$k_i = argmax_{k \in K_i} u_{k,i}. \tag{1}$$

The choice is made based on individual expectations of some path utility $u_{k,i}$, typically expressed as a function of expected travel times $t_{k,i}$ or costs $c_{k,i}$, weighted with individual perceptions, biases or random terms (included within $\boldsymbol{\theta_i}$).
Those choices shape travel times in the network $\boldsymbol{t}_a : a \in A$ and on the paths $t_{k,i}$ - the pivot variable of the assignment problem, being on the one hand the fundamental component of utility upon which the choices are made and on the other resulting from those choices. Travel times depend most importantly on the congestion pattern yielded by the joint route-choice action of all the drivers.
The travel times $\boldsymbol{t}$ are derived from some model $\mathcal{M}$, ranging from macroscopic static to microscopic agent-based models representing the actual traffic: $\boldsymbol{t} = \mathcal{M}(\mathcal{K}_i : i \in \boldsymbol{Q}, \dots)$. Travel times are experienced by drivers as a consequence of their choices and build expectations for the new day. Resulting in the standard day-to-day process, hypothetically equilibrating, i.e., reaching a state where no driver can unilaterally change her path choice to improve own utility.
Such (loosely formulated) Wardrop Equilibria are the state-of-practice in modelling urban demand patterns. Will it remain capable of representing the future system as well? We argue not and propose the postulated new framework.

Differences

When the fleets of autonomous vehicles will start replacing the human-driven vehicles, the paths chosen by CAVs k_i will be under the control of some AV operator, who will be in full charge of:

1. defining the choice set of paths potentially considered by each AV:
$$K_i \to K_i^* \tag{2}$$
, potentially including paths not considered by human drivers before.
2. defining the utility of the path:
$$u_{k,i} \to u_{k,i}^* \tag{3}$$
, with any arbitrary function.
3. predicting path travel times:
$$t_k \to t_k^* \tag{4}$$
, potentially better than humans.
4. executing routing, i.e., selecting a path that a CAV will follow.
$$k_i = \text{argmax}_{k' \in \boldsymbol{K_i}} u_{k,i} \to k_i^* \tag{5}$$
, not necessarily maximising individual utility as we will consider below.

Those four aspects exploit the differences between human and algorithmic routing behaviour that we postulate to include in the Post-Wardropian framework introduced here.

1. Choice set Even now, the route guidance can suggest paths which would not be considered by the driver. Yet it may be potentially amplified, as the CAV user is likely to accept more exotic paths [5]. This may extend the search space and identify more complex network loading patterns.

2. Arbitrary utility Despite being latent to the external observer and, to some extent, also to the decision-maker, individual path utility $u_{k,i}$ remains roughly predictable, and the choice made upon it remains rational [2]. This rationality can be relaxed with CAVs, as any components may be arbitrarily included in utility, as it is in the full control of the fleet operator. Even when simplifying the utility to a function of travel time only, a CAV may now include the following:

$t(k_i)$ own travel time expected on the selected path.

$T(k_i)$ system-wide total travel time (given individual path choice k_i)

$T_F(k_i)$ total travel time of the fleet to which one belongs.

$T_{-F}(k_i)$ total travel time beyond one's own fleet.

Which, weighted with γ's, yields the general postulated form of path utility for CAVs:

$$u_{k,i} = -(\gamma_i t(k_i) + \gamma_T T(k_i) + \gamma_F T_F(k_i) + \gamma_{-F} T_{-F}(k_i)) \tag{6}$$

Each utility maximising CAV will *behave* in routing according to the values of γ parameters. The natural case for a human driver would be with $\gamma_i = 1, \gamma_T, \gamma_F, \gamma_{-F} = 0$, which is in line with Wardrop's first principle [6]. Under the utopian, idealistic view, the potential of CAVs can then be used with $\gamma_i = 0, \gamma_T = 1, \gamma_F, \gamma_{-F} = 0$ - leading to System Optimal solution, minimising total travel costs. If we consider the fleet as a collaborative unit, cooperating to minimise group reward, one would set $\gamma_i, \gamma_T = 0, \gamma_F = 1, \gamma_{-F} = 0$. Those are the three most natural cases: *Egoistic, Social and Selfish* respectively. Followed by less immediate but also possible ones, specifically, with $\gamma_T, \gamma_{-F} < 0$ and the antisocial or competitive behaviour may emerge [3].

3. Predictive capabilities Accurate predictions of joint actions of all vehicles $\mathcal{K} = \{k_i : i \in Q\}$ will likely be a holy grail of efficient CAV fleet routing operations.
Classically, the route-choice estimation aims at predicting the actual choices made by human drivers with some probability model p calibrated with empirical data:

$$\theta^* = \text{argmax}_\theta \sum_{i \in Q} \sum_{k \in \boldsymbol{K_i}} ln(p(k_i, \theta_i)) * \mathcal{I}(k_i) \tag{7}$$

where θ represents calibrated model parameters (like β's in utility formulation for typical Discrete Choice Models), p is the trained route choice model (parameterized with θ) which yields route choice probabilities $p(k_i, \theta_i)$ and the classical log-likelihood formula aims to maximise the probability of the actually selected route $\mathcal{I}(k_i)$.
If we consider a single vehicle equipped with a good predictive module p, it predicts system-wide travel times $t_a : a \in A$, resulting from predicted route-choice actions of all other users, $\mathcal{K}_{-i} = \{k_{i'} : i' \in Q \setminus i\}$ and uses it to act optimally. This still does not differ much from the current situation with real-time route guidance.
The things get more interesting when we consider the fleet F of CAVs. If the fleet solves the above task well, it can predict the choices of human drivers, i.e., a joint action $\mathcal{K}_{-F} = \{k_i : i \notin F\}$ for all vehicles not in the direct control of the fleet. If so, then the route choice can become a Stackelberg game [7], where actions are taken simultaneously (at the departure) however the fleet operator knows upfront what the actions (strategies) of other human players will be. The CAV fleet operator is then a leader, taking advantage of predicting the actions of the followers. Notably, with large fleets, this is likely to make the system non-stationary and the route guidance inaccurate.
Even more interestingly, for the case of competing fleets $F, F' \in \mathcal{F}$ which both predict human actions: $\mathcal{K}_{-F} = \{k_i : i \notin \mathcal{F}\}$, but cannot predict strategic actions of their competitors. Leading to intriguing competition patterns, where human drivers are likely to be just a background for a game-theoretical interplay.

4. Route choice or fleet assignment? Eventually, given a greater choice set K_i, a customised objective function (rather than utility) u, well-predicted choices of others $\mathcal{K}$ and expected path utilities (based on expected travel times t_k) a CAV can now make a route choice – i.e., select a path k_i. While humans behaviourally execute it by maximising subjective utility, $\text{argmax}_{k' \in K_i} u_{k,i}$ the CAV algorithms are not constrained behaviourally and are free to act. Which we shall not denote anymore as a choice, but rather *assignment*, as the behavioural component is now algorithmised.
Notably, from the perspective of the fleet, which controls multiple vehicles, the action space grows exponentially with the fleet size, as the central controller can now assign a path to each of the controlled vehicles, which yields the action space size.

$$\mathcal{K}_F \in \prod_{i \in F} K_i, \tag{8}$$

This is now significant, as even with 50 controlled vehicles and 10 possible paths, the search space is now 10^{50} (in general, $|K|^{|Q|}$ for K-choice made by Q agents). And within this huge space $\mathcal{K}_F$, the fleet may use any possible methods to identify the optimal joint action:

$$\mathcal{K}^*{}_F = \text{argmin}_{\mathcal{K}_F \in \prod_{i \in F} K_i} u_F(\mathcal{K}_F, \boldsymbol{\gamma}) \tag{9}$$

,where $u_F(\mathcal{K}_F, \boldsymbol{\gamma})$ is utility (parameterized with $\boldsymbol{\gamma}$ in eq. 6) resulting from the joint route-choice action of a fleet operator. Any possible joint action (including those very different from human choices) can be the maximiser of utility, specifically with non-zero γ's.
In abstract toy networks, the optimal fleet strategy $\mathcal{S}_F$ may be found analytically. Yet arguably only for small instances and under quite strict assumptions (e.g., perfect prediction, deterministic travel times and/or choices, convex delay functions) hardly met in practice. In real-world scenarios we may safely assume that fleet operators will resort to machine learning to optimise assignments. Then, the problem (eq. 10) becomes a machine learning problem, where some algorithm finds the optimal joint action. A natural setting for this is Reinforcement Learning (as in [1]), where the policy π is iteratively learnt from interaction with environment $\mathcal{M}$. Then, a single episode is one day of commute, the actions are the route choices, $\mathcal{K}_F$ and a reward is the utility. The challenge is then interpretation of training, which can be done either along consecutive simulation days or externally (then a single simulated day may incorporate multiple internal runs to train the model).

Fleet market maximising strategy To illustrate the arising complexity and advocate for the need, let's assume that the fleet, just like any other player on the market, aims to increase its own market share. Then any $\mathcal{K}_F$ can be chosen as a strategy S to reach this goal. For instance, the one maximising the expected probability of joining the fleet, maximising profit or expected long-term market share. This may go beyond the utility formulation of eq. 6 and we say that a joint action ${\mathcal{K}^*}_F$ realises fleet strategy $\mathcal{S}_F$ if:

$$ {\mathcal{K}^*}_F = \text{argmax}_{\mathcal{K}_F \in \prod_{i \in F} K_i} \mathcal{S}_F(\mathcal{K}_F) \tag{10} $$

We can safely say that the probability $\pi_{F,i}$ of joining the fleet F by an individual i will be driven by:

- individual perception of CAV routing, expressed either as a travel time multiplier $\beta_{t,F,i}$ or an alternative-specific constant $\beta_{F,i}$. Based on stated preference [4] we may expect the CAV fleet to be perceived as more attractive than regular vehicles.
- costs $c_{F,i}$, either a one-time purchase or a subscription.
- advantage that it provides, measured for instance with $\Delta t_{i,F} = \bar{t}_{i,-F} - \bar{t}_{i,F}$ a difference between expected travel time in the fleet and beyond.

Leading to a general fleet joining probability formula:

$$ \pi_{F,i} \propto (\beta_{t,F,i}, \beta_{F,i}, c_{F,i}, \Delta t_{i,F}, \dots) \tag{11} $$

We may wish to simulate the strategy S of the fleet interested in maximising the probability of joining it (and of staying for those when already joined). Then, the eventual strategy $\mathcal{S}_F$ will result from a complex interplay between routing efficiency $\{\Delta t_{i,F} : i \in F\}$, individual characteristics $(\beta_{t,F,i},\ \beta_{F,i})$ and pricing strategy $c_{F,i}$, where routing efficiency will likely play a pivotal role. Understanding it and optimising calls for a cautious and detailed model representing underlying dynamics.

Post-Wardropian traffic assignment framework

Algorithm 1 presents a complete framework to model the state and dynamics of future urban mobility systems where the road network is shared between behavioural human drivers and competing fleets of autonomous vehicles.
Here, starting from the typical input of the traffic assignment (network, demand and traffic flow model). We perform the repeated daily simulations. It may start with no CAVs and simulate only humans would learn subjectively optimal choices to eventually stabilise close to User Equilibrium. This is the likely starting point when CAVs will start operations. As the fleet enters the system, we assume that every day it: i) predicts human choices, ii) predicts choices of other fleets, iii) picks its own strategy and sets prices for its users and, finally, iv) assigns vehicles to routes. Competing fleets will likely do it independently and simultaneously and later jointly execute their strategies on the networks. Such joint action of agents affects the system performance with the traffic flow model within the simulation. From these resulting network-wide travel times, each fleet, human driver and the city may learn and update or change strategy.
This framework, despite being closely similar to the classic Wardropian iterative process, employs several highly challenging components. It shall be considered a generic framework where specific elements need to be filled with external models and dedicated solutions, such as: a calibrated

Algorithm 1 Post-wardropian traffic assignment framework

Require: network $G(N, A)$, demand $\boldsymbol{Q}$, traffic flow model $\mathcal{M}$

```
for each day d do                                          ▷ day-to-day simulation
    for each fleet F ∈ 𝓕 do                                ▷ fleets act independently
        k_i : i ∈ −F ← (eq. 11))                           ▷ predict human choices
        k_i : i ∈ 𝓕 \ Fs                                   ▷ predict other fleets
        𝒮*_F ← argmax_{S∈𝒮} (eq. 10))                      ▷ pick own strategy
        c_{F,i}                                            ▷ set prices for individual users
        𝒦_F ← (eq. 9))                                     ▷ assign fleet vehicles to routes
    end for
    for each agent i do
        𝒦 ← k_i (eq. 1))                                   ▷ decision (humans) or assignment (CAVs)
    end for
    t(k_i : i ∈ Q) ← ℳ(𝒦)                                  ▷ simulate the system
    C(t)                                                   ▷ measure the system performance
    for each fleet F ∈ 𝓕 do                                ▷ fleet learning
        𝒮_F                                                ▷ Learn outcomes of applying the fleet strategy
    end for
    for each agent i ∉ F do                                ▷ human learningn/daptation
        t̄_{k,i} ← f(t_{k,i})                               ▷ update cost expectation (learn/adapt)
    end for
end for
return
```

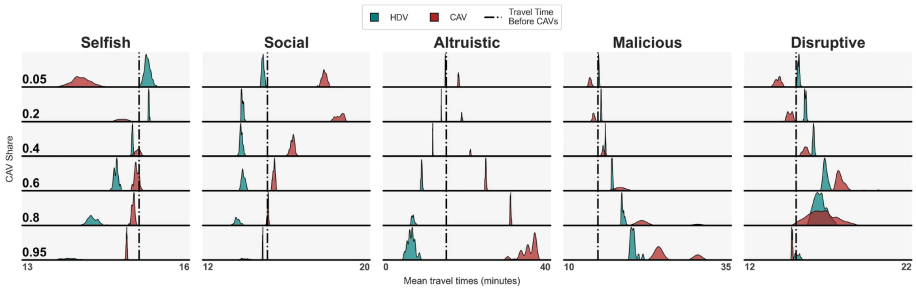


Figure 1: Resulting travel time distributions in a simple two-route scenario with varying strategies (eq. 6) and market penetrations – adapted from [3]

and accurate route-choice model (eq. 11), a game-theoretic model anticipating competitors' moves, selecting an optimal strategy to reach operators' objectives, an algorithm to identify a fleet assignment maximising strategy (eq. 9) - possibly trained externally, detailed microscopic traffic simulation, and a human adaptation model. Most remaining open research problems with no established methods and algorithms.

3 Results

In the full paper we will report a series of illustrative examples resulting from applying the above framework. Findings demonstrate that the fleet routing capabilities may significantly affect the system performance for all the users, induce instabilities and lead to both better and worse performance, depending on design choices made by fleet operators. Here we report illustrative cases demonstrating the impact on the system (Figure 1) and the process of learning oprimal RL strategies (Figure 2).

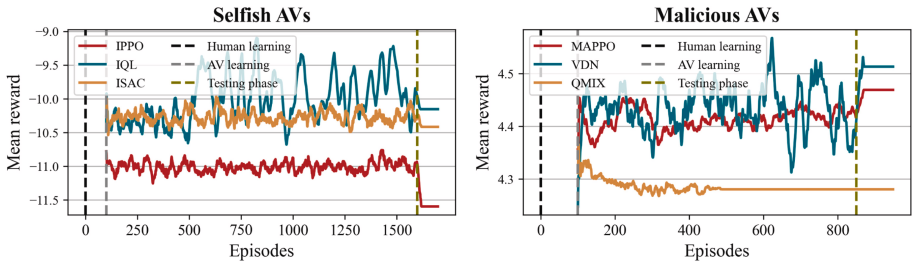


Figure 2: Challenges of training RL algorithm with various RL algorithms and objectives on a St. Arnault SUMO network – adapted from [1])

4 Conclusions

The initially stable traffic system, where human agents learned subjectively optimal routing decisions and the experienced costs met the expected ones, is likely to destabilise as CAVs will be introduced at scale. Since:

1. Individual, algorithmically driven CAV agents may apply different models to make routing decisions (eq. 6)).
2. The fleet operator may apply their own strategy (eq. 10).
3. Humans may adapt to the new strategies.
4. Fleet strategies may predict and anticipate those adaptations (eq. 7)
5. Fleets may compete for customers by applying various strategies (eq. 11).

The above framework will be instrumental in understanding the dynamics and equilibria of such future traffic systems. Human drivers, in such a system, can be in equilibrium when they, for instance, do not change day-to-day their routing probabilities or choices (stationarity) and when they do not decide to join the fleet (their probabilities of joining the fleet). Fleet members may be assigned by the fleet operator to any routes, and as long as they do not reconsider leaving the fleet or joining another fleet, they may be considered in equilibrium. The fleet remains in equilibrium when it does not adapt strategy. And the system is stable when distributions of traffic flows (and travel times) do not vary day-to-day. This brings several hypothetically stable consecutive states that may emerge:

1. when there are no CAVs and humans reach classic Wardrop Equilibrium.
2. when the fleet adopts routing that is indifferent to human routing;
3. when the fleet adopts a fixed arbitrary strategy and humans drivers adopt to it, reaching a new equilibrium (Nash);
4. when no user joins or leaves the fleet (market shares are fixed);
5. when no human drivers are left in the system and only CAVs are routing (monopoly);
6. when several fleets compete with individual strategies and reach some kind of equilibrium (either in a mixed system – with human drivers, or in a CAV-only system).

Most of the above not introduced formally, not considered and not experimented to understand the properties. They call for formalisation, understanding their properties, uniqueness and dynamics, which becomes possible with the proposed framework.

References

[1] Ahmet Onur Akman, Anastasia Psarou, Łukasz Gorczyca, Zoltán György Varga, Grzegorz Jamróz, and Rafał Kucharski. Routerl: Multi-agent reinforcement learning framework for urban route choice with autonomous vehicles. *SoftwareX*, 31:102279, 2025. ISSN 2352-7110. doi: https://doi.org/10.1016/j.softx.2025.102279.

[2] Carlos F Daganzo and Yosef Sheffi. On stochastic models of traffic assignment. *Transportation science*, 11(3):253–274, 1977.

[3] Grzegorz Jamróz, Ahmet Onur Akman, Anastasia Psarou, Zoltán György Varga, and Rafał Kucharski. Social implications of coexistence of cavs and human drivers in the context of route choice. *Scientific Reports*, 15(1):6768, 2025.

[4] Taha Hossein Rashidi, Travis Waller, and Kay Axhausen. Reduced value of time for autonomous vehicle users: Myth or reality? *Transport Policy*, 95:30–36, 2020.

[5] Tom Van Vuren and David Watling. A multiple user class assignment model for route guidance. *Transportation research record*, pages 22–22, 1991.

[6] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1:325–362, 1952.

[7] Hai Yang, Xiaoning Zhang, and Qiang Meng. Stackelberg games and multiple equilibrium behaviors on networks. *Transportation Research Part B: Methodological*, 41(8):841–861, 2007.