Multi-Agent Reinforcement Learning (MARL) Schema for Dynamic Traffic Assignment

Introduction

The 4-step transportation modeling process is a systematic approach used in urban and transportation planning to analyze and optimize transportation systems. The process involves four key stages: (1) Trip Generation, where the total number of trips originating in each zone of the network is estimated; (2) Trip Distribution, which determines the origins and destinations of these trips, and creates the OD matrix; (3) Mode Choice, where the transportation mode for each trip is selected; and (4) Traffic Assignment, the critical step that assigns trips to specific routes in the transportation network, considering factors like road capacity, congestion, and travel time. The last step is a dynamic and competitive process. This traffic assignment problem (TAP) becomes vital due to the finite road resources available (supply), leading travelers to compete for the most efficient routes to minimize their travel times (Shou et al.). The intricacies of road congestion underscore the interconnectivity of travelers' decisions, where the choices of one driver can significantly impact the travel times of others (Grunitzki et al.).

Conventionally, TAP employs one of two key concepts: User Equilibrium (UE) or System Optimum (SO). UE posits that the assignment process is akin to a Markov game with a Nash equilibrium solution, where no individual can enhance their travel time by independently altering their route choice (Sheffi). Nevertheless, it's recognized that Nash equilibria frequently lead to suboptimal outcomes when compared to socially optimal assignments (Zhou et al.). SO, on the other hand, is achievable only when all motorists collectively opt for actions that minimize the total system travel time rather than prioritizing their individual travel times [Sheffi].

Traditional TAP, whether considering UE or System SO, is typically addressed through mathematical modeling and solving (Zhou et al.). However, this approach often necessitates the use of simplifying assumptions, resulting in models that may diverge from real-world complexities. Moreover, these models are often difficult to solve. Therefore, in this project, our goal is to employ a Reinforcement Learning (RL) framework to model route choice behavior. This approach treats route selection as an adaptive decision-making process situated within a complex environment, enabling the model to learn optimal actions through iterative interactions with its surroundings.

Numerical Examples

The model's effectiveness presented in this study will be demonstrated by applying it to a real-world transportation network. Transportation networks are graphs where nodes represent intersections, and links correspond to roads through which traffic flows. Each link is associated with a latency function, typically a non-linear function (e.g., the commonly used BPR function), which models the travel time of the link based on its traffic flow (Shou et al.). We employed the Nguyen–Dupuis network to evaluate the model, a well-established benchmark in similar research efforts (Zhou et al.). This network comprises 13 nodes, 19 links, 25 routes, and 4 origin-destination (O-D) pairs.

It's important to highlight that, prior to deploying the model in a complex network, we should conduct preliminary testing in simpler network configurations to facilitate debugging. Our initial step involves developing a single-agent model in a network with an O-D pair and a single link. Once we have successfully defined the necessary variables and conducted the initial testing of this basic model, we have established a research trajectory. This trajectory will guide us in building a multi-agent model within a simplified network that features an O-D pair and five links, providing three distinct paths. Subsequently, we can expand the operational environment to assess the model's effectiveness in a more intricate network setting.

We also intend to conduct a comparative analysis of our Multi-Agent Reinforcement Learning (MARL) solution method with traditional approaches to evaluate its accuracy and effectiveness. To facilitate this comparison, we will leverage AequilibraE, a Python package designed for transportation modeling (AequilibraE), to perform the traffic assignment within the network and subsequently compare the obtained results. This comparative assessment will help us gauge the performance and advantages of our novel MARL approach in the context of transportation modeling.

Methodology

Reinforcement learning (RL) revolves around the challenge of enabling an agent to acquire behavior through dynamic interactions with its environment. Typically, the RL problem is characterized by a set of states (S), actions (A), a state transition function (T), and a reward function (R) that yields the reward $R(s, a)$ following a specific action. The primary objective is to determine a policy (π) that maximizes the agent's cumulative future rewards. At each state, the agent faces the crucial choice of either exploiting familiar actions or exploring new alternatives, potentially leading to the discovery of more advantageous actions (Grunitzki et al.).

In this problem, a state ($s \in S$) comprises two main components: the node and the time. For an agent currently in the state $s = (n, t)$, the available actions ($a \in A$) can either be selecting one of the outbound links from the node or, alternatively, utilizing one of the predefined routes from the origin to the destination. The reward function is typically defined as a negative value representing the travel cost associated with the state transition. This reward function is designed so that as travel time decreases, the reward increases, motivating drivers to seek to minimize their travel times (Shou et al.).

Timeline of the project

Week	W1	W2	W3	W4	W5	W6	W7	After	Responsibility
Tasks	Oct. 18	Oct. 25	Nov. 1	Nov. 8	Nov. 15	Nov. 22	Nov. 29	~	
1. Literature review on Traffic Assignment problems									SB
2. Literature review on RL methods									AK
3. Create the initial model and implement it on the simple network									IS and AK
4. Debug the model and implement it on the 2nd. network									IS and SB
5. Finalize the model and implement it on Nguyen network									IS, AK, and SB
6. Solve the problem using traditional approaches and compare the results									SB and IS

References

- Camargo, P., Zill, J., & O’Brien, A. (2018). AequilibraE. 2018
- Sheffi, Y. (1985). Urban transportation networks (Vol. 6). Prentice-Hall, Englewood Cliffs, NJ
- Shou, Z., Chen, X., Fu, Y., & Di, X. (2022). Multi-agent reinforcement learning for Markov routing games: A new modeling paradigm for dynamic traffic assignment. Transportation Research Part C: Emerging Technologies, 137, 103560
- Grunitzki, R., de Oliveira Ramos, G., & Bazzan, A. L. C. (2014, October). Individual versus difference rewards on reinforcement learning for route choice. In 2014 Brazilian Conference on Intelligent Systems (pp. 253-258). IEEE
- Zhou, B., Song, Q., Zhao, Z., & Liu, T. (2020). A reinforcement learning scheme for the equilibrium of the in-vehicle route choice problem based on congestion game. Applied Mathematics and Computation, 371, 124895