2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

# Modeling Multiday Route Choices of Strategic Commuters: A Mean Field Game Approach

Minghui Wu, Yafeng Yin, and Jerome P. Lynch

*Abstract*— **In the era of connected and automated mobility, commuters will possess strong computation capabilities, enabling them to make foresighted and strategic route choices. This paper investigates the implications of such strategic planning on traffic patterns by modeling the commute problem as a mean field game, where every traveler plans for sequential route choices over a span of several days. We examine the concept of multiday user equilibrium, a special mean field equilibrium under commuter interactions, to derive network traffic flow patterns. Under mild conditions, our analysis establishes the existence and uniqueness of the equilibrium flow pattern and explores its relationship with conventional Wardrop equilibrium.**

## I. Introduction

With rapidly-advancing communications and vehicular technologies, commuters are increasingly connected. Connectivity allows drivers to access various decision-support technologies such as a navigation app that assist them in making route choices. Such decision support is expected to become even stronger in the coming decades with the development and deployment of driving automation when drivers feel comfortable with rendering more driving and travel agency to machine. In the connected and automated mobility era, commuters (connected drivers or automated vehicles) will possess strong learning and computation capability, enabling them to make more foresighted and strategic travel decisions. Instead of being myopic, commuters can optimize their decision sequences for a longer time range, which results in a lower total cost because there is inertia or cost associated with switching route choices. Such strategic planning can profoundly impact traffic flow and the overall equilibrium of the transportation network. It is thus intriguing to investigate how the behavior of, and interaction among strategic commuters would dictate traffic patterns.

To address these challenges, this paper models the problem as a mean field game (MFG). The strategic planning behavior of individual commuters is explicitly modeled as a Markov optimal control problem, while the aggregate population behavior dictates the traffic evolution in the planning horizon. The multiday user equilibrium (MUE) is a special mean field equilibrium of the proposed model, where no commuter can reduce their overall cost by altering their policy sequence. We then conduct a thorough analysis of the properties of the MUE such as its existence, uniqueness, and relationship with conventional Wardrop equilibrium. Due to the page limit, this paper focuses on presenting the model with homogeneous commuters.

This work was partly supported by research grants from National Science Foundation (CMMI-1904575 and CMMI-2233057) and the USDOT University Transportation Center for Connected and Automated Transportation (CCAT).

Minghui Wu is with the Department of Civil and Environmental Engineering, University of Michigan. minghuiw@umich.edu

Yafeng Yin is with the Department of Civil and Environmental Engineering, Departure of Industrial and Operations Engineering, University of Michigan. yafeng@umich.edu

Jerome P. Lynch is with the Department of Civil and Environmental Engineering, Duke University. jerome.lynch@duke.edu

### A. Prior work

Wardrop equilibrium (WE) has been a widely utilized notion for analyzing and modeling transportation systems. Initially introduced by Wardrop [1], WE, also known as user equilibrium (UE), characterizes a delicate state where no commuter can unilaterally change their route choices to reduce travel costs. Over time, the concept has been extended under different behavioral considerations and real-world needs. Traditionally, WE models have primarily focused on one-shot or stateless games. This approach fails to account for commuters' need for strategic planning of their travel over multiple days. In reality, route choices between days are interdependent due to the presence of inertia, which refers to commuters' reluctance in adjusting their choices [2]. In such cases, it is rather necessary to plan the trajectory of route choices, striking a balance between minimizing travel time and avoiding adjustment. To address this issue, this paper extends the framework to a Markov game setting, where commuters plan for sequential travel decisions. We will theoretically analyze the resulting traffic flow pattern when multiple individuals' sequential decision-making processes are coupled together.

As mentioned, the methodology in the paper mainly lies in the field of MFG, which involves a game played by an infinite number of infinitesimal players [3], [4]. Our model differs from traditional finite-horizon MFG [5]–[7] in that it does not necessitate an exogenous initial distribution. This is because when modeling route choices, it is impractical to cyclically enforce a specific traffic flow upon all commuters as their starting point. In fact, the starting distribution should emerge as a natural outcome of the commuters' interaction process, rather than being externally prescribed. Thus, the approach without an exogenous initial condition aligns more closely with the dynamic nature of daily commute choices, where individuals adapt and react to changing conditions over time.

### B. Contributions

To the best of our knowledge, this paper is the first to model and analyze commuters' strategic planning behavior and its consequent traffic flow pattern, advancing our understanding of traffic dynamics in the connected and automated mobility era. In response to this challenge, we introduce a novel concept of MUE for transportation network equilibrium analysis. This concept can be seen as a special mean field equilibrium, which is the steady state of the commuters' interaction process. At equilibrium, the starting and ending distributions of the planning horizon should be the same. Otherwise, commuters' interactions will yield a different sequence. Furthermore, our work establishes a crucial connection between MUE and the conventional Wardrop equilibrium in three distinct scenarios: no inertia, short planning horizons, and long planning horizons.

## II. Model

We consider a planning period of $N$ days, with a day $n \in \mathcal{N} = \{0, 1, ..., N-1\}$. An infinite number of commuters[1] are making their route choices every day on a graph $(\mathcal{V}, \mathcal{L})$, where $\mathcal{V}, \mathcal{L}$ are the set of all the nodes and links respectively. An origin-destination (OD) pair corresponds to two nodes in $\mathcal{V}$, which are connected by several paths. Each path is comprised of links in $\mathcal{L}$. In this paper, we only consider homogeneous commuters, which means that there is only one OD pair and all commuters share the same cost preference. The set of available paths for the OD pair is denoted as $\mathcal{S} = \{s_1, ..., s_M\}$. Since each commuter selects one route on each day, the route choice can be viewed as the state of the commuter on the day. In this sense, the set $\mathcal{S}$ becomes a (finite) state space of commuters.

The distribution of states over the population is called the mean field (MF) distribution. On day $n$, it is denoted as $\mu_n \in \mathcal{P}(\mathcal{S})$, where $\mathcal{P}(\mathcal{S})$ refers to the probability mass function defined on $\mathcal{S}$. We use a bold notation $\boldsymbol{\mu} = \{\mu_n\}_{n\in\mathcal{N}} \in \mathcal{M}$ to denote the MF distribution sequence over the horizon $\mathcal{N}$, where $\mathcal{M}$ refers to the domain of all possible MF distribution sequences. For a fixed total demand $\xi$, the flow on link $l \in \mathcal{L}$ is $x(l, \mu_n) : \mathcal{L} \times \mathcal{P}(\mathcal{S}) \to \mathbb{R}$, where $x(l, \mu_n) = \sum_{s\in\mathcal{S}} \xi\mu_n(s)\delta_{l,s}$. $\delta_{l,s}$ equals 1 if link $l$ is on route $s$, and 0 otherwise. Denote the link flow vector as $x(\mu_n) = \{x(l, \mu_n)\}_{l\in\mathcal{L}}$. Thus, by introducing the link-path incidence matrix $\Delta = [\delta_{l,s}]_{l\in\mathcal{L},s\in\mathcal{S}} \in \mathbb{R}^{|\mathcal{L}|\times|\mathcal{S}|}$, we can write $x(\mu_n) = \xi\Delta\mu_n$.

On each day, commuters can switch to another route or stay with the previous one. We consider the action on day $n$ is to pick the route for day $n+1$. Subsequently, the day-to-day route choice over the planning period can be modeled as a finite-horizon Markov decision process (MDP), and the action space $\mathcal{A}$ is identical to the state space $\mathcal{S}$. In this model, we seek a time-varying, feedback control policy $\pi_n(a|s) : \mathcal{S}\times\mathcal{S} \to [0, 1]$. We use a similar notation $\boldsymbol{\pi} = \{\pi_n\}_{n\in\mathcal{N}} \in \Pi$ to represent the policy sequence, where $\Pi$ is the domain for all possible policy sequences. For simplicity, we denote that $\pi_n(\cdot|s) \in \mathcal{P}(\mathcal{S}), \pi_n \in \mathcal{S} \times \mathcal{P}(\mathcal{S})$.

We metrize $\mathcal{P}(\mathcal{S})$ with the distance $d_f(\mu, \mu') = \max_{s\in\mathcal{S}} |\mu(s) - \mu'(s)|$, $\mu, \mu' \in \mathcal{P}(\mathcal{S})$. Then, we define the metrics for $\mathcal{M}$ and $\Pi$ with sup metrics.

[1]In this paper, we use players, travelers and commuters interchangeably

Meanwhile, the system dynamic is $P(s_{n+1} = s'|s_n = s, a_n = a) = \begin{cases} 1, & s' = a \\ 0, & \text{else} \end{cases}$. It means that although commuters can choose their actions from a stochastic policy, if they choose $s'$ for the next day, their next state will always be $s'$. This dynamic is flexible and can be easily generalized to consider uncertainty in the transition process.

On each day $n$, each commuter will experience a cost, which is modeled by $c(s, s', \mu_n, \pi_{n,s}) = f(s, \mu_n) + d(s, s') + \frac{1}{\theta}\ln\pi_n(s'|s)$, where $\pi_{n,s}$ refers to $\pi_n(\cdot|s)$. The first cost $f(s, \mu_n) : \mathcal{S} \times \mathcal{P}(\mathcal{S}) \to \mathbb{R}$ is the travel time of route $s$. Let $t_l(x(l, \mu_n))$ denote the link travel time on $l$, then the path travel cost is $f(s, \mu_n) = \sum_{l\in\mathcal{L}} t_l(x(l, \mu_n))\,\delta_{l,s}$. For example, the link travel time $t_l(x)$ can be chosen as the so-called BPR function [8] $t_l(x) = t_l^0\left[1 + \beta_l\left(\frac{x}{c_l}\right)^4\right]$, where $t_l^0$ is the free-flow travel time; $c_l$ is the capacity and $\beta_l$ is a parameter. Here we assume that $t_l(x)$ is continuous for all $l \in \mathcal{L}$, thus $f(s, \mu_n)$ is continuous with respect to $\mu_n$. The second cost $d(s, s')$ is a general distance function that captures user inertia, which refers to commuters' disutility in adjusting their routes [2]. We will mainly use the formulation $d(s, s') = \epsilon \cdot \mathbf{1}_{s\neq s'}$ [9], where one receives a small penalty $\epsilon$ whenever they switch routes. The third cost is used to reflect the random residue in the value function, which follows i.i.d. Gumbel distribution. Some have also used it as entropy regularization or penalization [5], [10].

### A. Individual behavior

Given the population behavior $\boldsymbol{\mu}$, commuters find the optimal policy sequence

$$\min_{\boldsymbol{\pi}\in\Pi} J_{\boldsymbol{\mu}}(\boldsymbol{\pi}) = E\left[\sum_{n=0}^{N-1} c(s_n, a_n, \mu_n, \pi_{n,s_n})\right] \tag{1}$$

subject to

$$s_0 \sim \mu_0, a_n \sim \pi_{n,s_n}, s_{n+1} = a_n \tag{2}$$

where $\mu_0$ is the starting distribution in $\boldsymbol{\mu}$.

To characterize the optimality, for the given $\boldsymbol{\mu}$, the value function of a policy sequence $\boldsymbol{\pi}$ and the optimal value function on day $n$ are defined respectively as follows

$$V_n^{\boldsymbol{\mu},\boldsymbol{\pi}}(s) = E\left[\sum_{k=n}^{N-1} c(s_k, a_k, \mu_k, \pi_{k,s_k})\right] \tag{3}$$

$$V_n^{\boldsymbol{\mu}}(s) = \inf_{\boldsymbol{\pi}} E\left[\sum_{k=n}^{N-1} c(s_k, a_k, \mu_k, \pi_{k,s_k})\right] \tag{4}$$

subject to similar constraints as (2), where $s \in \mathcal{S}$. For any given value function $V$, we can define two Bellman equations, for the policy and optimal value functions respectively, by the following two Bellman operators:

$$\mathcal{G}_\mu^\pi V(s) = \sum_{s'\in\mathcal{S}} \pi(s'|s)\left(c(s, s', \mu, \pi_s) + V(s')\right) \tag{5}$$

$$\mathcal{G}_\mu V(s) = \inf_\pi \sum_{s' \in \mathcal{S}} \pi(s'|s)\left(c(s, s', \mu, \pi_s) + V(s')\right) \quad (6)$$

where $s \in \mathcal{S}, \mu \in \mathcal{P}(\mathcal{S})$, $\pi \in \mathcal{S} \times \mathcal{P}(\mathcal{S})$. Hereinafter, we will call the former policy Bellman operator and the latter optimal Bellman operator or simply Bellman operator. If a policy sequence $\boldsymbol{\pi} \in \Pi$ is optimal with respect to the population behavior $\boldsymbol{\mu} \in \mathcal{M}$, then for all $n \in \mathcal{N}$ and $s \in \mathcal{S}$, there must be

$$\mathcal{G}_{\mu_n} V^{\boldsymbol{\mu}}_{n+1}(s) = \mathcal{G}^{\pi_n}_{\mu_n} V^{\boldsymbol{\mu}}_{n+1}(s) = V^{\boldsymbol{\mu}}_n(s) \quad (7)$$

Under the cost formulation of the proposed model, given $V$ as the value function for the next day, we obtain the unique optimal policy by solving a strictly convex problem

$$\pi(s'|s) = \frac{e^{-\theta(d(s,s')+V(s'))}}{\sum_{a \in \mathcal{S}} e^{-\theta(d(s,a)+V(a))}} \quad s' \in \mathcal{S} \quad (8)$$

Correspondingly, given the current MF distribution $\mu \in \mathcal{P}(\mathcal{S})$, substituting the optimal policy yields:

$$\mathcal{G}_\mu V(s) = f(s, \mu) - \frac{1}{\theta} \ln \left[ \sum_{a \in \mathcal{S}} e^{-\theta(d(s,a)+V(a))} \right] \quad (9)$$

It is commonly assumed that the travel time $f$ and inertia cost $d$ are non-negative and upper bounded by some constant $C$. Assume $\pi_s$ is the optimal policy for state $s$ given the value function $V$ and MF distribution $\mu$, then $\mathcal{G}_\mu V(s) = f(s, \mu) + \sum_{a \in \mathcal{S}}[d(s, a) + \frac{1}{\theta} \ln \pi_s(a) + V(a)]\pi_s(a)$. For other state $s'$, $\pi_s$ may not be optimal, thus $\mathcal{G}_\mu V(s') \leq f(s', \mu) + \sum_{a \in \mathcal{S}}[d(s', a) + \frac{1}{\theta} \ln \pi_s(a) + V(a)]\pi_s(a)$. Hence,

$$\begin{aligned} &\mathcal{G}_\mu V(s') - \mathcal{G}_\mu V(s) \leq f(s', \mu) - f(s, \mu) \\ &+ \sum_{a \in \mathcal{S}} [d(s', a) - d(s, a)]\pi_s(a) \leq 2C \end{aligned} \quad (10)$$

Switching $s$ and $s'$ yields $|\mathcal{G}_\mu V(s) - \mathcal{G}_\mu V(s')| \leq 2C$. For any MF distribution sequence $\boldsymbol{\mu}$, since the final value $V_N(s) = 0$ for all states, we can prove by induction that the optimal value $\{V_n\}_{n \in \mathcal{N}}$ satisfies $|V_n(s) - V_n(s)| \leq 2C$ for all $n$ and $s$. Combining it with (8) and the bound for $d$ yields the fact that the optimal policy sequence $\boldsymbol{\pi}$ satisfies

$$\begin{aligned} \pi_n(s'|s) &= \frac{e^{-\theta\left[d(s,s')+V_{n+1}(s')\right]}}{\sum_{x \in \mathcal{S}} e^{-\theta[d(s,x)+V_{n+1}(x)]}} \\ &\geq \frac{e^{-\theta\left[C+V_{n+1}(s')\right]}}{\sum_{x \in \mathcal{S}} e^{-\theta[V_{n+1}(s')-2C]}} = \frac{1}{Me^{3\theta C}} \end{aligned} \quad (11)$$

for all $n, s, s'$. We denoted the lower bound by $\omega$.

Note that it is the relative value between states that matters rather than the absolute value. Suppose we add a constant on the value of all states, it will have no influence on the system. Therefore, we say that the value function $V$ is defined on $\mathbb{R}^M/\mathbb{R}$, and the norm for the value function is $\|V\|_\# = \inf_{\lambda \in \mathbb{R}} \|V + \lambda\|$, where $\|\cdot\|$ is the $L_2$-norm [5].

### *B. Population behavior*

For a given policy $\pi \in \mathcal{S} \times \mathcal{P}(\mathcal{S})$ and an MF distribution $\mu \in \mathcal{P}(\mathcal{S})$, we can define the operator

$$\mathcal{K}_\pi \mu(s) = \sum_{s' \in \mathcal{S}} \mu(s')\pi(s|s') \quad s \in \mathcal{S} \quad (12)$$

which outputs the induced next MF distribution. Thus, if a policy sequence $\boldsymbol{\pi} \in \Pi$ can induce an MF distribution sequence $\boldsymbol{\mu} \in \mathcal{M}$, for any $n$ and $s \in \mathcal{S}$, there must be

$$\mathcal{K}_{\pi_n} \mu_n(s) = \mu_{n+1}(s) \quad (13)$$

## III. Multiday User Equilibrium

Section II discusses the criteria for determining the optimality of a policy sequence and whether it can induce an MF distribution sequence. These two concepts will be used in this section to define the multiday user equilibrium or MUE.

### *A. Interaction process*

Before formally defining the equilibrium, we first present a motivating example to illustrate how strategic commuters interact with each other. Suppose the horizon length is $N = 3$ and the distribution sequence in the first episode $\mu_0 \sim \mu_2$ is randomly generated. As in Figure 1, at the end of day 2, after observing the MF distribution sequence in the first episode, the strategic commuters can calculate the best response $\pi^*_0 \sim \pi^*_2$ by sequentially using (8) and (9) backward. Note that since $V_3(s)$ always equals 0 for all states, $\pi^*_2$ is a trivial uniform policy to maximize entropy. On contrary, determining $\pi^*_0$ and $\pi^*_1$ requires the knowledge of $\mu_1$ and $\mu_2$. Also note that $\mu_0$ does not appear in the calculation process, hence it will not influence the optimal policy sequence. Then, $\pi^*_0$ and $\pi^*_1$ will be used as the policy on day 2 and 3 respectively, which induces the MF distribution $\mu_3$ and $\mu_4$. To facilitate analysis of the system, we treat $\mu_2$ along with $\mu_3$ and $\mu_4$ as the next MF distribution sequence. In this sense, the starting distribution of the next sequence is always the ending distribution of the last sequence. Similarly, at the end of day 4, commuters can solve and implement the new policy sequence, and the episode-by-episode process will go on in a similar way.

Compared to previous works to analyze traffic equilibrium such as [1], a key difference in the proposed model is that commuters consider route choice over the planning horizon as a whole. Because the commuters are implementing the best response over a planning horizon, this strategic planning behavior achieves a balance between minimizing travel costs and avoiding adjustments under the presence of user inertia, which dictates a new traffic flow pattern. The steady state of the interaction process will be defined as the MUE to analyze the system in the next section.

### *B. Definition*

Formally, we first denote the mapping from an MF distribution sequence to its unique optimal policy sequence as $\Phi : \mathcal{M} \to \Pi$. Besides, denote the mapping from a policy sequence to its induced MF distribution sequence starting

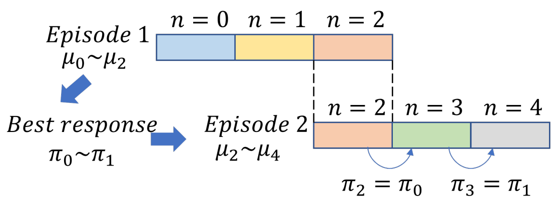


Fig. 1. Illustration of the interaction process

from a specific $\mu \in \mathcal{P}(\mathcal{S})$ as $\Psi_\mu : \Pi \times \mathcal{P}(\mathcal{S}) \to \mathcal{M}$. Then, the definition of the MUE is given as follows

*Definition 1:* A pair $(\boldsymbol{\pi}, \boldsymbol{\mu})$ is called an MUE if

$$\boldsymbol{\pi} = \Phi(\boldsymbol{\mu}), \quad \boldsymbol{\mu} = \Psi_{\mu_{N-1}}(\boldsymbol{\pi})$$

Given an MF sequence $\boldsymbol{\mu}$, $\Psi_{\mu_{N-1}}(\boldsymbol{\pi})$ means that the policy sequence $\boldsymbol{\pi}$ induces the next sequence from the final distribution of $\boldsymbol{\mu}$, which matches with the interaction process in the Section III-A. Note that to ensure the MUE is in a steady state, it must have the same starting and ending distribution. We can write the two operators together as $\Gamma(\boldsymbol{\mu}) = \Psi_{\mu_{N-1}}(\Phi(\boldsymbol{\mu}))$.

*Remark 1:* It is worth stressing the difference between the MUE and conventional mean field equilibrium or MFE. For any given initial distribution $\nu$, an MFE $(\boldsymbol{\pi}, \boldsymbol{\mu})$ should satisfies $\boldsymbol{\pi} = \Phi(\boldsymbol{\mu}), \boldsymbol{\mu} = \Psi_\nu(\boldsymbol{\pi})$ [7]. Unlike the conventional MFE, the definition of MUE does not rely on any exogenous variable, such as $\nu$.

The following sections will analyze the properties of the MUE and its relationship with traditional WE.

### C. Existence and uniqueness

We start with proving the existence of the MUE due to the continuity of $f(s, \mu)$,

*Proposition 1:* If the cost $f(s, \mu)$ is continuous, there always exists at least one MUE $(\boldsymbol{\pi}, \boldsymbol{\mu})$.

*Proof sketch:* We identify $\mathcal{M}$ with simplex $\mathcal{S}_{|\mathcal{N}|(|\mathcal{S}|-1)} \subseteq \mathbb{R}^{|\mathcal{N}|(|\mathcal{S}|-1)}$ and $\Pi$ with simplex $\mathcal{S}_{|\mathcal{N}||\mathcal{S}|(|\mathcal{S}|-1)} \subseteq \mathbb{R}^{|\mathcal{N}||\mathcal{S}|(|\mathcal{S}|-1)}$. Define the mapping $\Gamma : \mathcal{S}_{|\mathcal{N}|(|\mathcal{S}|-1)} \to \mathcal{S}_{|\mathcal{N}|(|\mathcal{S}|-1)}$, which satisfies $\Gamma(\boldsymbol{\mu}) = \Psi_{\mu_{N-1}}(\Phi(\boldsymbol{\mu}))$. It maps from an MF distribution sequence $\boldsymbol{\mu}$ to another sequence induced by the optimal policy sequence and starts from the final distribution of $\boldsymbol{\mu}$. We can then prove the existence by Brouwer's fixed-point theorem. □

Meanwhile, we assume that the link performance function $t_l(v)$ is strictly monotone, which is mild and has been widely adopted in the literature. While the MUE itself may not be unique, we can demonstrate that every MUE should have distinct starting and ending distributions.

*Proposition 2:* Under the monotonicity assumption, if two MUEs $(\boldsymbol{\pi}, \boldsymbol{\mu})$ and $(\tilde{\boldsymbol{\pi}}, \tilde{\boldsymbol{\mu}})$ satisfy that $\mu_0 = \tilde{\mu}_0$, there must have $\mu_n = \tilde{\mu}_n$ and $\pi_n = \tilde{\pi}_n$ for all $n \in \mathcal{N}$.

*Proof sketch:* For simplicity, denote $x_n(l) = x(l, \mu_n)$, and $\tilde{x}_n(l) = x(l, \tilde{\mu}_n)$. Motivated by Proposition 1 in [11] and using the path-link relationship, we can prove that $\sum_{l\in\mathcal{L}} (x_n(l) - \tilde{x}_n(l)) [t_l(x_n(l)) - t_l(\tilde{x}_n(l))] \leq 0$. Hence, all equality must hold due to the strict monotonicity, which leads to $x_n(l) = \tilde{x}_n(l)$ for all $n, l$. Therefore, $f(s, \mu_n) = f(s, \tilde{\mu}_n)$ holds for all $s, n$ by definition, thus $\mathcal{G}_{\mu_n} V = \mathcal{G}_{\tilde{\mu}_n} V$ for all $V$. By induction, the two MUE have the same value functions $\{V_n\}_{n\in\mathcal{N}}$. As a result, $\pi_n = \tilde{\pi}_n$ for $n \in \mathcal{N}$ based on (8), and the proposition holds. □

### D. Connection with Wardrop Equilibrium

This section discusses the relationship between MUE and conventional Wardrop Equilibrium or WE, which in this paper, refers to a general concept, including UE [1] and other variants.

*1) No inertia:* As we discussed, user inertia influences travel choices across adjacent days. When there is no inertia, commuters do not need to be foresighted. In this case, the following proposition demonstrates that the MUE simplifies to logit-based stochastic user equilibrium (logit-SUE) [12], an extension of UE to capture bounded rationality.

*Proposition 3:* When $d(s, s') = 0$ for all $s, s'$, there exists a unique MUE $(\boldsymbol{\pi}, \boldsymbol{\mu})$. Furthermore, let the MF distribution $\mu_{SUE}$ denote the logit-based SUE distribution, then $\mu_n = \mu_{SUE}$ holds for all $n$.

*Proof sketch:* In this special case, the optimal policy $\pi_n(s'|s)$ in (8) has nothing to do with the previous state $s$. Hence, $\pi_n(s'|s) = \mu_{n+1}(s')$ fo all $n, s, s'$. Substituting in (8) and taking log on both sides yields the following result for all $n$

$$V_n(s) + \frac{1}{\theta} \ln \mu_n(s) = -\frac{1}{\theta} \ln \sum_{x\in\mathcal{S}} e^{-\theta V_n(x)} \tag{14}$$

By using (9), (14) as well as the final value $V_N(s) = 0$ for all $s$, we can prove that $f(s, \mu_n) + \frac{1}{\theta} \ln \mu_n(s)$ is the same for all $s$ and $n$, which matches the condition for logit-SUE. □

*2) Short planning horizon:* In the presence of user inertia, when the planning horizon is very short (i.e. $N = 2$), commuters essentially plan only for the next day. In this scenario, the framework reduces to state-dependent SUE (SDSUE) [13], [14], which is similar to SUE but considers the inertia between path choices. The following proposition provides further insight into this concept.

*Proposition 4:* When $N = 2$ and $d(s, s') = \epsilon \cdot \mathbf{1}_{s\neq s'}$ with $\epsilon \neq 0$, denote $\mu_{SDSUE}$ as the SDSUE distribution, then $\mu_0 = \mu_1 = \mu_{SDSUE}$.

*Proof sketch:* Since the MUE shares the same starting and ending distribution, we denote it as $\mu = \mu_0 = \mu_1$. It can be proved that the optimal policy on day 0 is

$$\pi_0(s'|s) = \frac{e^{-\theta(d(s,s')+f(s',\mu))}}{\sum_{x\in\mathcal{S}} e^{-\theta(d(s,x)+f(x,\mu))}} \tag{15}$$

By definition, the optimal policy $\pi_0$ can maintain $\mu$ to be invariant, which matches the definition of SDSUE [14]. Hence, $\mu = \mu_{SDSUE}$. □

*3) Long planning horizon:* Additionally, the MUE reveals certain patterns when the planning horizon tends toward infinity. To discuss the pattern, we first introduce a special multiday equilibrium, named stationary equilibrium (SE). This concept is built upon the idea of stationary solutions presented in [5] and maintains a time-invariant MF distribution and policy across consecutive days.

*Definition 2:* A pair $(\bar{V}, \bar{\mu})$, where $\bar{\mu} \in \mathcal{P}(\mathcal{S})$ and $\bar{V} \in \mathbb{R}^M/\mathbb{R}$, is called an SE if it satisfies the following conditions for all $s \in \mathcal{S}$:

- There exists a constant $\bar{\lambda}$ such that $\mathcal{G}_{\bar{\mu}}\bar{V}(s) = \bar{V}(s) + \bar{\lambda}$
- $\mathcal{K}_{\bar{\pi}}\bar{\mu}(s) = \bar{\mu}(s)$, where $\bar{\pi}$ is the unique optimal policy determined by $\bar{V}$ and $\bar{\mu}$.

As in the following lemma, the existence of SE is generally assured.

*Lemma 1 (Theorem 3 in [5]):* Under the continuous and bounded cost function, there always exists an SE.

Note that Section II-A shows that $\pi_n(s'|s) \geq \omega$ holds for all $n \in \mathcal{N}$. Assume that every link is covered by some paths, then $x(l, \mu_n) \geq \omega$, which means that the resultant link flow will share a common lower bound. Consequently, the link travel time function is actually strongly monotone. This assumption is not restrictive. For example, the BPR function satisfies this assumption with a lower bound on link flow

$$(x_1 - x_2)[t_l(x_1) - t_l(x_2)] \geq \frac{4\beta_l t_l^0 \omega^3}{c_l^4}(x_1 - x_2)^2 \quad (16)$$

Denote the coefficient as $\eta_l$. By picking $\eta = \min_{l \in \mathcal{L}} \eta_l$, for any link flow $\{x(l)\}_{l \in \mathcal{L}}$ and $\{x'(l)\}_{l \in \mathcal{L}}$

$$\sum_{l \in \mathcal{L}} [t_l(x(l)) - t_l(x'(l))] [x(l) - x'(l)] \geq \eta \|x - x'\|^2 \quad (17)$$

where $\|\cdot\|$ is the $L_2$-norm.

Under the strong monotonicity assumption, the following proposition proves that all SEs have the same link flow.

*Proposition 5:* Under the continous and monotone cost assumption, if $(\bar{V}_1, \bar{\mu}_1)$ and $(\bar{V}_2, \bar{\mu}_2)$ are both SE, there must have $x(\bar{\mu}_1) = x(\bar{\mu}_2)$.

*Proof sketch:* For simplicity, denote $\bar{x}_1(l) = x(l, \bar{\mu}_1), \bar{x}_2(l) = x(l, \bar{\mu}_2)$.

Motivated by Proposition 7 in [5] and using the path-link relationship, we can derive

$$\begin{aligned}&\sum_{l \in \mathcal{L}} (\bar{x}_1(l) - \bar{x}_2(l)) [t_l(\bar{x}_1(l)) - t_l(\bar{x}_2(l))] \\ &+ 2\phi\|\bar{V}_1 - \bar{V}_2\|_{\#} \leq 0\end{aligned} \quad (18)$$

Due to the strict monotonicity, all equality must hold. Therefore, $\bar{x}_1(l) = \bar{x}_2(l)$ for all $l \in \mathcal{L}$. □

Now, we are prepared to establish the relationship between the MUE and SE. The following proposition demonstrates that, as the planning horizon tends to infinity, the SE emerges within the MUE, either centrally or at the two extremes.

*Proposition 6:* Without losing generality, assume the episode length is odd, and denote the episode as $\mathcal{N} = \{0, 1, ..., 2k\}$. When the episode length is $2k+1$, denote the corresponding MUE as $(\boldsymbol{\pi}^k, \boldsymbol{\mu}^k)$. Under the continuous and monotone cost assumption, for every $\epsilon > 0$, there exists $K$ such that either $\Delta\mu_0^k = \Delta\bar{\mu}$, or $\Delta\mu_k^k$ is $\epsilon$-close to $\Delta\bar{\mu}$ for all $k \geq K$.

*Proof sketch:* Motivated by Theorem 7 in [5] and using the path-link relationship, we first prove that for any $\nu$, if $\|\Delta\nu - \Delta\bar{\mu}\| > 0$, the mean field equilibrium or MFE with horizon length $2k+1$ and initial distribution $\nu$, $(\boldsymbol{\pi}^{MFE}, \boldsymbol{\mu}^{MFE})$, satisfies that

$$\|\Delta\mu_k^{MFE} - \Delta\bar{\mu}\|^2 + \|V_k^{MFE} - \bar{V}\|_{\#}^2 \leq B\left(\frac{E}{E+1}\right)^{k-1} k^2 \quad (19)$$

where $V^{MFE}$ is the value function defined for the MFE, and $B, E$ are two constants independent of $k$ and $\nu$. Therefore, for any $\epsilon > 0$, there exist $K$ (independent of $\nu$) such that $\|\Delta\mu_k^{MFE} - \Delta\bar{\mu}\|$ is below $\epsilon$ for all $k \geq K$.

Now, suppose there exists $p \geq K$ such that with horizon length $2p+1$, one of the resulting MUE $(\boldsymbol{\pi}^p, \boldsymbol{\mu}^p)$ satisfies $\|\Delta\mu_0^p - \Delta\bar{\mu}\| > 0$ and $\|\Delta\mu_p^p - \Delta\bar{\mu}\| > \epsilon$. Now if we use $\mu_0^p$ as the initial distribution, the resulting MFE with horizon length $2p+1$ will be the same as the MUE according to Proposition 2, which contradicts the bound we got. Thus we prove the proposition by contradiction. □

## IV. Numerical Examples

The proposed model is applied to a three-by-three grid network taken from [15], as shown in Figure 2. The coefficients in the BPR function are randomly generated and provided in Table I. All commuters travel from node 1 to 9, and the total inflow is 2,000. Commuters have 6 paths to choose from. The path-link relationship is given in Table II. The planning horizon spans seven days, with $N = 7$. Lastly, we take $d(s, s') = \epsilon \cdot \mathbf{1}_{s \neq s'}$ as the adjustment cost, and set the coefficient $\theta = 1$. Fictitious play [6] is used as the algorithm to numerically solve the MUE.

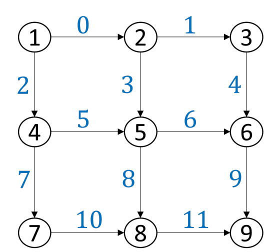


Fig. 2. Test network with nine nodes and twelve links.

TABLE I
Link travel cost information

| Link | $c$ | $\beta$ | $t^0$ | Link | $c$ | $\beta$ | $t^0$ |
|---|---|---|---|---|---|---|---|
| 0 | 600 | 0.23 | 15 | 6 | 500 | 0.16 | 17 |
| 1 | 600 | 0.29 | 12 | 7 | 500 | 0.24 | 19 |
| 2 | 600 | 0.22 | 14 | 8 | 500 | 0.18 | 11 |
| 3 | 500 | 0.18 | 12 | 9 | 800 | 0.19 | 17 |
| 4 | 900 | 0.21 | 14 | 10 | 700 | 0.23 | 10 |
| 5 | 600 | 0.2 | 17 | 11 | 600 | 0.16 | 16 |

The resulting MUE, represented in path flow evolution, is shown in Figure 3. Each sub-figure plots the flow dynamic of the path during the seven-day period. The blue curve corresponds to the case with inertia, where $\epsilon$ is set to 1. As we have analyzed, the flow starts and ends at the same position, but there will always be within-horizon fluctuation.

TABLE II
PATH-LINK RELATIONSHIP

| Path | Link |
|---|---|
| 0 | 0, 1, 4, 9 |
| 1 | 0, 3, 6, 9 |
| 2 | 0, 3, 8, 11 |
| 3 | 2, 7, 10, 11 |
| 4 | 2, 5, 8, 11 |
| 5 | 2, 5, 6, 9 |

The green curve plots the MUE without user inertia. As can be seen from the figure, the curve is stable in the sense that it roughly maintains the same value, which matches our analysis. For better demonstration, we also set the horizon length to 2 and 20 to solve for the SDSUE and SE respectively, which are also presented in the figure using the red and yellow lines respectively.

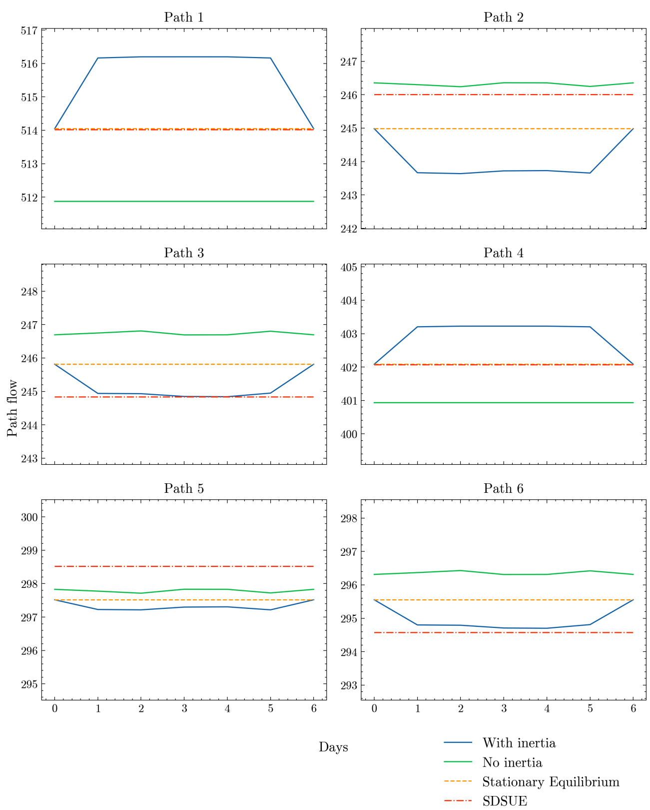


Fig. 3. Path flow evolution

## V. Conclusion and Future Work

In this research, we have developed a mean field game model to capture the multiday route choices of strategic and foresighted commuters. Multiday user equilibrium or MUE is defined for the game, which represents a special fixed-point solution that eliminates the need for exogenous initial distributions. Under mild conditions, we have demonstrated the existence and uniqueness of the equilibrium. Furthermore, we have shown how the MUE reduces to the conventional Wardrop equilibrium in two specific cases, while also exploring its asymptotic behavior in more general scenarios when the time horizon approaches infinity.

Broadly speaking, our study enhances the understanding of the future mobility system when commuters are increasingly connected and automated. Moreover, by characterizing traffic flow as the outcome of the MUE, the MUE-based analysis could potentially serve as an alternative to the traditional static equilibrium paradigm for analyzing transportation systems.

There are several interesting open problems to be solved. In the current work, players are assumed to have perfect information. To enhance the model's applicability to real-world scenarios, it is important to consider cases with imperfect and incomplete information. It is also interesting to investigate how to incorporate the learning process in the framework, where agents update their policies while observing the information based on their daily experiences.

## References

[1] J. G. Wardrop, "Road paper. some theoretical aspects of road traffic research," *Proc. Inst. Civ. Eng.*, vol. 1, no. 3, pp. 325–362, 1952.

[2] H. S. Mahmassani and G.-L. Chang, "On boundedly rational user equilibrium in transportation systems," *Transportation science*, vol. 21, no. 2, pp. 89–99, 1987.

[3] M. Huang, R. P. Malhamé, and P. E. Caines, "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle," *Commun. Inf. Syst.*, vol. 6, no. 3, pp. 221–252, 2006.

[4] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese journal of mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[5] D. A. Gomes, J. Mohr, and R. R. Souza, "Discrete time, finite state space mean field games," *Journal de Mathématiques Pures et Appliquées*, vol. 93, no. 3, pp. 308–328, 2010.

[6] S. Perrin, J. Perolat, M. Laurière, M. Geist, R. Elie, and O. Pietquin, "Fictitious play for mean field games: Continuous time analysis and applications," *Advances in Neural Information Processing Systems*, vol. 33, pp. 13 199–13 213, 2020.

[7] K. Cui and H. Koeppl, "Approximately Solving Mean Field Games via Entropy-Regularized Deep Reinforcement Learning," in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 130:1909-1917, 2022.

[8] B. of Public Roads, *Traffic Assignment Manual*. U.S. Dept. of Commerce, Urban Planning Division, 1964.

[9] P. Delle Site, "A mixed-behaviour equilibrium model under predictive and static Advanced Traveller Information Systems (ATIS) and state-dependent route choice," *Transp. Res. Part C Emerg. Technol.*, vol. 86, pp. 549–562, 2018.

[10] Q. Xie, Z. Yang, Z. Wang, and A. Minca, "Learning while playing in mean-field games: Convergence and optimality," *in International Conference on Machine Learning*, PMLR, pp. 11 436–11 447, 2021.

[11] J. Perolat et al., "Scaling Mean Field Games by Online Mirror Descent," *Proc. Int. Jt. Conf. Auton. Agents Multiagent Syst. AAMAS*, vol. 2, no. Aamas, pp. 1028–1037, 2022.

[12] Y. Sheffi, "Urban transportation networks: Equilibrium analysis with mathematical programming methods," 1985.

[13] C. Castaldi, P. D. Site, and F. Filippi, "Stochastic user equilibrium in the presence of inertia," in *Transportation Research Procedia*, Elsevier B.V., pp. 13–24, 2017.

[14] C. Castaldi, P. Delle Site, and F. Filippi, "Stochastic user equilibrium in the presence of state dependence," *EURO J. Transp. Logist.*, vol. 8, no. 5, pp. 535–559, 2019.

[15] X. He and S. Peeta, "A marginal utility day-to-day traffic evolution model based on one-step strategic thinking," *Transp. Res. Part B Methodol.*, vol. 84, pp. 237–255, 2016