# Learning an Unknown Network State in Routing Games

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

| Citation       | Wu, Manxi, and Saurabh Amin. "Learning an Unknown Network State in Routing Games." IFAC-PapersOnLine, 52, 20 (2019): 345-350. © 2019 IFAC-PapersOnLine. All rights reserved. |
|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| As Published   | http://dx.doi.org/10.1016/j.ifacol.2019.12.179                                                                                                                                                      |
| Publisher      | Elsevier BV                                                                                                                                                                                        |
| Version        | Final published version                                                                                                                                                                             |
| Citable link   | https://hdl.handle.net/1721.1/125223                                                                                                                                                               |
| Terms of Use   | Creative Commons Attribution-NonCommercial-NoDerivs License                                                                                                                                          |
| Detailed Terms | http://creativecommons.org/licenses/by-nc-nd/4.0/                                                                                                                                                 |
Learning an Unknown Network State in Routing Games

Manxi Wu *, Saurabh Amin **

* Institute for Data, Systems and Society, MIT, manxiwu@mit.edu
** Civil and Environmental Engineering, Institute for Data, Systems and Society, MIT, amins@mit.edu

Abstract: We study learning dynamics induced by myopic travelers who repeatedly play a routing game on a transportation network with an unknown state. The state impacts cost functions of one or more edges of the network. In each stage, travelers choose their routes according to Wardrop equilibrium based on public belief of the state. This belief is broadcasted by an information system that observes the edge loads and realized costs on the used edges, and performs a Bayesian update to the prior stage’s belief. We show that the sequence of public beliefs and edge load vectors generated by the repeated play converge almost surely. In any rest point, travelers have no incentive to deviate from the chosen routes and accurately learn the true costs on the used edges. However, the costs on edges that are not used may not be accurately learned. Thus, learning can be incomplete in that the edge load vector at rest point and complete information equilibrium can be different. We present some conditions for complete learning and illustrate situations when such an outcome is not guaranteed.

© 2019, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: Strategic learning, Routing games, Bayesian state estimation.

1. INTRODUCTION

Transportation networks are prone to disruptions resulting from random infrastructure breakdowns and adverse events such as natural disasters and security attacks. In many cases, the impact of these disruptions can be modeled as a change in the latent condition (or state) that influences travel costs on one or more network edges. Major disruptions, such as the 2007 collapse of I-35W bridge over Mississippi River in Minneapolis, can result in an abrupt interruption of nominal flow patterns on the network and trigger repeated learning and adjustment of travel decisions over a period of time (Zhu et al. (2010)).

In this paper, we focus on learning of an unknown network state when travelers have access to an information system that repeatedly updates and broadcasts state information (in the form of public belief). This problem is relevant to situations when a disruption or replacement of an infrastructure facility (e.g., bridge, bypass highway) results in travelers’ reliance on a public source of information to learn and adjust their route choices. In recent years, such information sources have become widely available and increasingly sophisticated in terms of their ability to collect traffic loads and costs on network edges via a variety of heterogeneous data sources (e.g., fixed sensors, GPS-enabled mobile sensors, and crowdsourcing services). Our generic learning model considers strategic travelers with access to imperfect information of the state from a public information system.

In particular, we consider learning dynamics resulting from a routing game of non-atomic travelers who repeatedly use the network with an unknown state (Sec. 2). The state represents the latent network condition and affects the travel costs on edges. In each state, the cost of any edge is an increasing function of the edge load (i.e., aggregate traffic load on that edge). For simplicity, we assume that travelers have identical preferences and that the demand is fixed. At the beginning of each stage, all travelers have access to the most recent public belief of the state (i.e., the probability of each possible state) through the public information system. Travelers then myopically choose routes with the smallest expected cost based on the public belief of the state. Thus, the routing strategy in each stage is a Wardrop equilibrium corresponding to the current public belief. Using classical results (Sandholm (2001)) one can conclude the essential uniqueness of the induced equilibrium load. Furthermore, the realized edge costs are random functions of the edge load and the true state. Again, for simplicity, we assume that cost distributions are normally distributed.

Importantly, in our model, the public information system acts as an “aggregator”: In each stage, it collects the loads and realized costs of all edges that were used in that stage, and updates the belief of state according to Bayes’ rule. However, the costs of edges that are not used are not available to the information system. The model captures several important and practically relevant aspects of how travelers learn the new network state: (i) travelers’ knowledge of the state, represented as the public belief, gets repeatedly updated by the information system; (ii) travelers are strategic (selfish) in that they choose routes with the lowest costs based on the current belief; (iii) the Bayesian update on the public belief is based on the available realizations of the loads and costs on edges that are used in each stage of repeated routing. Thus, the history of play affects the set of used edges and the
realized costs through belief update, and in turn affects the outcomes in future stages. Consequently, classical Bayesian learning, e.g. results in Blackwell and Dubins (1962), cannot be applied to our problem, due to the correlated and non-identical distribution of stage outcomes and the strategic behavior of travelers.

Our work contributes to the literature on learning and information effects in routing games. Prior work has studied the impact of heterogeneous state information on travel decisions in static settings; see for e.g. Arnott et al. (1991); Khan and Amin (2018); Wu et al. (2018). On one hand, the transportation community has studied travelers’ response to real-time traffic information using dynamic behavioral models and network simulation (Ben-Akiva et al. 1991; Mahmassani and Jayakrishnan 1991). In a related work Jha et al. (1998), the authors propose a Bayesian model to capture how travelers update their perception of travel times based on the received information. Their simulation-based approach identifies the evolution and stabilization of travel patterns. On the other hand, a variety of learning dynamics (Fudenberg et al. 1998) have been particularized to routing games. In the classical fictitious play setting, players best respond to the belief of opponents’ strategies in each stage and revise their beliefs based on the observed actions (Monderer and Shapley 1996; Marden et al. 2009)). Another well-known setting is pay-off based learning (Marden and Shamma 2012; Cominetti et al. (2010)), in which players’ strategies follow a controlled dynamics based on the history of their own realized costs.

In contrast, our learning setup focuses on modeling how strategic travelers make route choices, and how the belief of the uncertain state is updated based on the routing game outcomes. In related work Meigs et al. (2017), the authors consider a repeated routing game, in which travelers learn the slope of linear cost functions using a least square estimator based on the realized costs. In this paper, we take a Bayesian viewpoint and do not assume a specific functional form on the edge costs. Moreover, in their article, the costs of untaken edges are the intercepts of the linear cost functions, which are known by all travelers. However, in our model, travelers only know noisy realizations of costs on taken edges.

We now summarize our main result (Theorem 1) and analysis approach. We show that the public belief and the equilibrium edge load in each stage almost surely converge to a rest point that satisfies two properties: (i) The edge load is induced by a Wardrop equilibrium that corresponds to the public belief at convergence; (ii) Travelers accurately estimate the expected costs of all edges that are used. These two properties ensure that, at any rest point, the realized edge costs do not provide new information about the state, and travelers have no incentive to change route choices (Sec. 3).

Notably, our concept of rest point is similar to the self-confirming equilibrium studied in Fudenberg and Levine (1993), Fudenberg and Kreps (1995). These papers consider a strategic learning model, in which players myopically play an extensive-form game in each stage, and update their belief of opponents’ strategies based on the observed actions. Eventually, players accurately predict and best respond to opponents’ strategies on the reached information sets, but may continue to maintain incorrect beliefs on the unreached ones. Similarly, in our model, travelers learn the true cost of the used edges at rest points, but may persistently have an incorrect estimate of costs for the unused edges because no additional information is available to revise the belief. Consequently, some edges that are taken in the complete information equilibrium may not be taken at a rest point.

Moreover, we highlight some properties of rest points. We find that if the network is series-parallel, the average cost experienced by travelers at any rest point is no less than that in complete information equilibrium (Proposition 1). We also provide a set of conditions, under each of which travelers end up only using the edges that are part of the complete information equilibrium; i.e., they eventually make route choice as if they know the true state (Proposition 2). In this case, the learning is complete in that the edge load vector converges almost surely to the complete information equilibrium (Sec.4). Finally, we conclude our work in Sec. 5.

2. MODEL

In this section, we introduce our model of learning dynamics in which travelers repeatedly play a routing game in a transportation network. Travelers have access to a public information system, which updates and broadcasts the probability of each possible state (public belief). We first describe the routing game and travelers’ routing decisions in each stage. Then, we present how the information system updates the belief of the state.

2.1 Routing Games

Consider a transportation network modeled as a directed graph with a single origin-destination pair. Let $E$ denote the set of edges and $R$ denote the set of routes (sequences of edges connecting the origin and destination nodes). The uncertain network state, denoted $s$, represents the unknown infrastructure condition after disruption. The cost (travel time) function of each edge $e$, denoted $\ell_e^k(\cdot)$, depends on the state. We assume that state is grounded in a set $S$, and remains fixed throughout the learning dynamics.

In each stage game $G^k$, where $k = 1, 2, \ldots$, travelers receive the public belief of the state from the public information system. The belief is denoted as $\theta^k = (\theta^k(s))_{s \in S} \in \Delta(S)$, where $\theta^k(s)$ is the probability of state $s$.

Travelers are non-atomic players with a total demand of $D$, i.e. each player constitutes an infinitesimal part of the total demand. Travelers’ routing strategy is $q^k \triangleq (q^k_e)_{r \in R}$, where $q^k_e$ is the demand of travelers using route $r$. A strategy is feasible if it satisfies $\sum_{r \in R} q^k_r = D$ and $q^k_r \geq 0$ for any $r \in R$. The set of feasible strategy is denoted by $Q$. Given any strategy $q^k \in Q$, the aggregate edge load on edge $e \in E$ in stage $k$ is $w^k_e = \sum_{r \in R} q^k_r$, the aggregate edge load on edge $e \in E$ in stage $k$ is $w^k_e = \sum_{r \in R} q^k_r$, and the random variable $\ell^k_e$ representing the noise of the realized cost:

$$c^k_e = \ell^k_e(w^k_e) + \epsilon^k_e, \quad \forall e \in E,$$  

(1)
We make the following assumptions on the edge cost functions and the random noise:

(A1) For any \( e \in \mathcal{E} \), the edge cost function \( \ell^e_k \) is strictly increasing in \( w^e_k \), and the derivative on \( w^e_k \) is bounded from below by a positive number \( \alpha > 0 \), i.e.
\[
\frac{d\ell^e_k(w^e_k)}{dw^e_k} \geq \alpha, \quad \forall e \in \mathcal{E}, \quad \forall s \in \mathcal{S}, \quad \forall w^e_k > 0.
\]

(A2) The realized stage noise \( e_k = (\epsilon^e_k)_{e \in \mathcal{E}} \) has Gaussian distribution with mean \( 0 \) and a non-degenerate covariance matrix \( \Sigma \). The random variables \( \{e^k\}_{k=1}^{\infty} \) are independently and identically distributed.

With a slight abuse of notation, we denote the cost function of each route \( r \in \mathcal{R} \) in state \( s \) as \( \ell^r_s(w^r_k) = \sum_{e \in r} \ell^e_k(w^e_k) \). The realized cost on \( r \) is \( c^r_s = \ell^r_s(e_k) + \sum_{e \in r} \epsilon^e_k \). Using (A2), we can compute the expected cost of each \( r \in \mathcal{R} \) based on the stage belief \( \theta^k \) as follows:
\[
E_{\theta^k}[\ell^r_s(q^k)] = \sum_{s \in \mathcal{S}} \theta^k(s) \ell^r_s(q^k).
\]

We assume that travelers are myopic players in that they minimize the expected cost of the chosen route in each stage, without considering the impact of their route choice on the future outcomes. Then, travelers must take routes with the smallest expected cost based on the current public belief. This leads to the concept of Wardrop equilibrium defined as follows:

**Definition 1.** A feasible strategy profile \( q^k \in \mathcal{Q} \) is a Wardrop equilibrium corresponding to the stage belief \( \theta^k \in \Delta(\mathcal{S}) \) if the following condition is satisfied:
\[
q^k_r > 0 \Rightarrow E_{\theta^k}[\ell^r_s(q^k)] = \min_{r' \in \mathcal{R}} E_{\theta^k}[\ell^{r'}_s(q^k)], \quad \forall r \in \mathcal{R}.
\]

Since the edge cost functions are increasing in the aggregate edge load, the equilibrium is *essentially unique* in that given any \( \theta^k \), the corresponding equilibrium edge load \( w^k = (w^e_k)_{e \in \mathcal{E}} \) is unique (see Sandholm (2001)).

If \( \theta \) assigns probability 1 on state \( s \), then Definition 1 reduces to the classical Wardrop equilibrium under complete information of the state \( s \). We denote the complete information equilibrium edge load vector as \( w^* \).

### 2.2 Information and Belief

We introduce a public information system, which can observe the stage game outcomes – the load and the realized costs on edges that are used in each stage. The information system then updates the public belief based on the stage game outcomes according to Bayes’ rule, and broadcasts the updated public belief to all travelers.

The initial public belief, denoted \( \theta^0 \), represents the information of the network condition that is available to the public information system. We assume that the initial belief does not exclude any possible state:
\[
\theta^0(s) > 0, \quad \forall s \in \mathcal{S}.
\]

We now introduce our learning dynamics. The initial time can be set to the time of change in the existing network condition (for example, after a disruption as motivated in Sec. 1). In each stage \( k \in \{1, 2, \ldots\} \):

**Step 1:** The information system broadcasts the public belief of the state \( \theta^{k-1} \) to all travelers.

**Step 2:** Traveler populations play the routing game \( G^k \) based on the public belief \( \theta^{k-1} \) according to Wardrop equilibrium \( q^{k-1} \). The induced edge load vector is \( w^{k-1} \).

**Step 3:** The information system observes the aggregate load vector \( w^k \), and the realized costs on the taken edges \( c^k \). We denote the set of edges that are used in stage \( k \) as \( \mathcal{E}^k = \{e \in \mathcal{E} | w^e_k > 0\} \), and thus \( e^k = (e^k_e)_{e \in \mathcal{E}^k} \).

In stage \( s \), the observed costs on these edges has the following Gaussian distribution:
\[
c^k_s \sim N(\ell^s(e^k), \Sigma^k),
\]

where \( \ell^s_{e^k} = (\ell^s(e^k_e))_{e \in \mathcal{E}^k} \), and \( \Sigma^k \) is the sub-matrix of \( \Sigma \) with rows and columns corresponding to edges in \( \mathcal{E}^k \).

Then, the probability density function of \( c^k \) for a state \( s \) and edge load vector \( w^k \) can be written as follows:
\[
\phi^k_s[w^k][c^k] = \frac{\exp \left\{ -\frac{1}{2}(c^k - \ell^s(e^k))^\top (\Sigma^k)^{-1} (c^k - \ell^s(e^k)) \right\}}{(2\pi)^{|\mathcal{E}^k|/2} \sqrt{|\Sigma^k|}}.
\]

According to Bayes’ rule, the updated state belief \( \theta^{k+1} \) is the posterior distribution of the state based on the belief in stage \( k \) and the distribution of the realized costs:
\[
\theta^{k+1}(s) = \frac{\theta^k(s) \cdot \phi^k_s[w^k][c^k]}{\sum_{s' \in \mathcal{S}} \theta^k(s') \cdot \phi^k_s[w^k][c^k]}, \quad \forall s \in \mathcal{S}.
\]

We emphasize that travelers’ routing decisions impact the belief update. Secondly, in each stage, if an edge is not used by travelers, then the information system does not observe the cost of that edge. Thus, travelers cannot learn how the state impacts the cost on unused edges.

Note that the distribution of \( e^k \) only depends on the state \( s \) and the edge load \( w^k \). Then, given any two stages \( k \) and \( k' \), the realized costs \( e^k \) and \( e^{k'} \) are independent conditional on the edge loads \( w^k \) and \( w^{k'} \). We define \( W^{k+k'} = (w^j)_{j=1}^{k+k'} \) as the history of equilibrium edge load vectors, and \( C^k = (c^j)_{j=1}^k \) as the history of realized costs until the end of stage \( k \). Then in stage \( s \), the probability density function of \( C^k \) conditioned on \( W^{k+k'} \), can be written as follows:
\[
p^k(C^k | W^{k+k'}) = \prod_{j=1}^{k} \phi^j(s, w^j)[c^j], \quad \forall k.
\]

By iteratively applying (5), we can derive \( \theta^k \) directly from the initial belief \( \theta^0 \) and the history of realized costs \( C^k \) and edge loads \( W^{k+k'} \):
\[
\theta^k(s) = \frac{\theta^0(s) \cdot p^k(C^k | W^{k+k'})}{\sum_{s' \in \mathcal{S}} \theta^0(s') \cdot p^k(C^k | W^{k+k'})}, \quad \forall s \in \mathcal{S}, \quad \forall k.
\]

The main advantages of introducing the information system in our model is that it observes the stage game outcomes, and computes belief updates, so that travelers do not need to observe and update by themselves. Moreover, this setup allows for different sets of travelers to participate in different stages, i.e. an individual traveler may participate in one or multiple routing games as long as they are informed of the public belief of the state in the participated stages.
In our learning model, the public belief of the state \( \theta^k \) changes based on the stage game outcomes, and in turn influences the edge load \( w^{k*} \) induced by travelers’ route choices. Therefore, the tuple \((\theta^k, w^{k*})\) governs how travelers make route choices according to our learning dynamics, and whether or not they learn the true state of the network.

3. CONVERGENCE OF LEARNING DYNAMICS

In this section, we show that the sequence \( \{(\theta^k, w^{k*})\}_{k=1}^{\infty} \) generated by our learning dynamics converges to a rest point with probability 1.

Before presenting our theorem, we first introduce the concept of distinguishable states. To avoid confusion, we denote the true state as \( s \). For any load vector \( w \), we say that the state \( s \) is distinguishable from the true state \( s \) if:

\[
\exists e \in \{C|w_e > 0\}, \ s.t. \ \ell_s^e(w_e) \neq \ell_w^e(w_e).
\]

The set of distinguishable states given \( w \) is denoted as \( S^I(w) \). The distribution of the realized costs in any distinguishable state \( s \in S^I(w) \) is different (as in (3)) from that in the true state \( s \). Hence, state \( s \) can be distinguished from the true state \( s \) based on the realized costs. On the other hand, if the cost functions in state \( s \) are different from that in \( s \) only on a subset of edges and no edge in that set is used in \( w \), then \( s \) is indistinguishable, i.e. \( s \in S \setminus S^I(w) \).

**Theorem 1.** For any true state \( s \in S \), and any initial belief \( \theta^0 \) that satisfies (2), we have:

\[
\lim_{k \to \infty} \theta^k = \bar{\theta}, \ w.p.1, \quad \text{(8a)}
\]

\[
\lim_{k \to \infty} w^{k*} = \bar{w}, \ w.p.1, \quad \text{(8b)}
\]

where \( \bar{\theta} \) is a probability vector that satisfies:

\[
\bar{\theta}(s) = 0, \quad \forall s \in S^I(\bar{w}), \quad \text{(9)}
\]

and \( \bar{w} \) is the equilibrium edge load vector corresponding to the public belief \( \bar{\theta} \).

From Theorem 1, we know that the public belief \( \theta^k \) and the associated equilibrium edge load \( w^{k*} \) in the learning dynamics eventually converge to a rest point \((\bar{\theta}, \bar{w})\), which satisfies the following two properties:

1. **Equilibrium under imperfect state information:** Travelers have no incentive to deviate from the chosen routes based on public information of the state.

2. **Consistency:** Since the public belief \( \bar{\theta} \) excludes any distinguishable states, for any \( s \) such that \( \theta(s) > 0 \) and any \( e \in \tilde{E} \), we must have \( \ell_s^e(\bar{w}) = \ell_w^e(\bar{w}) \). Then, travelers accurately learn the cost of edges that are used, i.e. \( E[\ell_s^e(\bar{w})] = \ell_w^e(\bar{w}) \) for any \( e \in \tilde{E} \).

Additionally, we can show that the convergent public belief \( \bar{\theta} \) is a fixed point of the belief update function in (5). If the initial belief \( \theta^0 = \bar{\theta} \), then in any stage \( k \), the public belief is \( \theta^k = \bar{\theta} \), and the equilibrium edge load is \( w^{k*} = \bar{w} \).

Due to the space limit, we only present the main intuition behind the proof of Theorem 1.

Firstly, we show that the process of the public beliefs \( \{\theta^k\}_{k=0}^{\infty} \) is a bounded martingale. Hence, \( \theta^k \) converges to a random probability vector \( \bar{\theta} \) with probability 1. Secondly, by adapting the sensitivity analysis approach in Dafermos and Nagurney (1984) to our problem, we show that the equilibrium edge load \( w^{k*} \) in each stage is a continuous function of the public belief \( \theta^{k-1} \). Therefore, from continuous mapping theorem and the fact that \( \theta^k \) converges to \( \theta \), we can show that the unique equilibrium edge load \( w^{k*} \) also converges to \( \bar{w} \), which is the equilibrium edge load corresponding to \( \bar{\theta} \).

Finally, it remains to be shown that \( \bar{\theta} \) satisfies (9) with probability 1. To start with, we view the problem of distinguishing any state \( s \) from the true state \( s \) as a hypothesis testing problem. Next, we can analyze the log-likelihood ratio of the sequence of realized cost conditional on the edge loads until stage \( k \) in state \( s \) and the true state \( s \), denoted by \( \frac{p^*(C^k|W^{k*})}{p^*(C^k|W^s)} \). In fact, we can conclude that for any distinguishable state \( s \in S^I(\bar{w}) \),

\[
\lim_{k \to \infty} \log \left( \frac{p^*(C^k|W^{k*})}{p^*(C^k|W^s)} \right) = \infty \quad \text{with probability 1}. \quad \text{(10)}
\]

Hence, \( \bar{\theta} \) satisfies (9).

4. REST POINT ANALYSIS

In this section, we first compare travelers’ route choices and average cost at a rest point with that in complete information equilibrium. Then, we provide conditions under which travelers eventually make route choice as if they know the true state with probability 1. Finally, we illustrate some interesting aspects of learning dynamics and rest points through examples.

4.1 Comparison with Complete Information Equilibrium

We say the learning is complete if the edge load at the rest point is \( \bar{w} = w^{**} \) (i.e. is the complete information equilibrium in the true state \( s \)) with probability 1.

However, learning may converge to other rest points \((\bar{\theta}, \bar{w})\) such that \( \bar{w} \neq w^{**} \). This is because \( \bar{\theta} \) may assign positive probability on another state \( s \in S \setminus S^I(\bar{w}) \) that is not distinguishable from the true state \( s \) given the converged load vector \( \bar{w} \). Therefore, although the estimated costs on the used edges are identical to the true costs, the estimated cost on an unused edge based on \( \bar{\theta} \) may be higher than the actual cost in the true state. Consequently, if travelers know the true state, they may have the incentive to deviate from the rest point.

We denote the average cost of travelers at a rest point with edge load \( \bar{w} \) as \( C(\bar{w}) \triangleq \sum_{e \in \tilde{E}} \bar{w}_e \ell^e_w(\bar{w}) \), and the average cost in the complete information equilibrium as \( C(w^{**}) \triangleq \sum_{e \in \tilde{E}} w^{**}_e \ell^e_w(w^{**}) \).

The next proposition shows that if the network is series-parallel (i.e. the network does not have an embedded wheatstone network, see Milchtaich (2006)), the average cost at any rest point is no less than that in complete information equilibrium.

**Proposition 1.** If the network is series-parallel, then \( C(\bar{w}) \geq C(w^{**}) \) at any rest point \((\bar{\theta}, \bar{w})\).
The idea of the proof is that the edge load at any rest point is equivalent to the complete information equilibrium in a routing game on a subnetwork. In other words, travelers make route choice as if they only know a subset of the available routes in the original network. Then, based on Theorem 1 in Milchtaich (2006), we can show that if the network is series-parallel, the equilibrium average cost on the subnetwork is no less than that on the original network.

4.2 Conditions for Complete Learning

We now provide a set of conditions, under each of which the learning is complete with probability 1.

Proposition 2. For any true state \( s \in S \), the learning is complete, i.e. \( \lim_{k \to \infty} w^{ks} = w^{s*} \) with probability 1, if any of the following conditions is satisfied:

1. **Fully distinguishable states:** For any \( w \), \( S^i(w) = S \setminus \{ s \} \), i.e. for any \( s \in S \setminus \{ s \} \), \( s \) is distinguishable from \( s \).
2. **State-independent free flow travel time:** For any \( e \in E \) and any \( s \in S \), \( \ell^i(e) \) is identical across states.
3. **All edges are utilized:** For any \( s \in S \) and any \( e \in E \), \( w^{s*} > 0 \).

Each condition ensures that travelers repeatedly use the set of edges that should be taken in complete information equilibrium. Then, travelers will eventually learn the costs on these edges, and choose routes as if they know the true state.

In practice, condition (1) is relevant if every state impacts the costs on all edges. Condition (2) requires that the state only impacts the costs when there is congestion \( (w_e > 0) \). For example, lane closure does not change the cost when there is no traffic, but significantly aggregates congestion due to the loss of capacity. Condition (3) requires that all edges are utilized regardless of the state. This will hold when the traffic demand is high so that all routes are taken.

4.3 Example and Discussion

Consider a three-edge series-parallel network with set of states: \( \{ e_1, e_2, e_3, \emptyset \} \), where \( s = e \) means edge \( e \) has a disruption, and \( s = \emptyset \) means all edges are in nominal condition. The cost of edge \( e \) is \( \ell_e(w_e) \) if \( s \neq e \), and \( \ell_e^0(w_e) \) if \( s = e \). See Fig. 1 for the network and cost functions.

![Fig. 1. Three-edge network](image)

We assume that the noise term in each stage is \( \epsilon^k = \epsilon_e^k \epsilon_{\in E} \sim N(0, \Sigma) \), where \( \Sigma \) is a three-dimensional identity matrix. The true state is \( s = \emptyset \). Total demand is 1.

We can check that given any public belief, travelers will always take route \( e_3 - e_1 \). Then, two cases naturally arise at the rest point:

1. **Travelers take both** \( e_3 - e_1 \) and \( e_2 - e_1 \). In this case, travelers eventually learn the true state. The rest point is \( \theta^s = (0, 0, 0, 1) \), \( w^s = (1, 0.5, 0.5) \).
2. **Travelers only take** \( e_3 - e_1 \). In this case, travelers can distinguish states \( s = e_1, e_2 \) from the true state \( s = \emptyset \), but not state \( s = e_2 \). Moreover, if travelers believe that the probability of \( s = e_2 \) is no less than \( 0.2 \), then exclusively taking \( e_3 - e_1 \) is an equilibrium strategy. Therefore, the set of rest points in this case is \( \bar{\theta} = (0, x, 0, 1-x) \) for any \( x \geq 0.2 \), and \( \bar{w} = (1, 0, 1) \).

Since the network is series-parallel, we can check that the average costs \( C(\bar{w}) = 12 > 11.5 = C(w^{s*}) \) (Proposition 1).

We simulate the learning dynamics with the initial public belief \( \theta^0 = (1/4, 1/4, 1/4, 1/4) \), under which the equilibrium strategy is to take route \( e_3 - e_1 \). Due to the randomness of noise terms in (1), different realizations of stage costs lead to different belief updates and routing strategies, and eventually can lead to different rest points. We demonstrate one realization of the public belief \( \theta^k \) and the equilibrium edge load \( w^{s*} \) that converges to the complete information equilibrium \( (\theta^*, w^{s*}) \) in Fig. 2a - 2b. We can see that the learning dynamics indeed converges to one rest point (Theorem 1).

The main reason of different rest points is that travelers learn to take both routes after the first few stages in first realization, but always exclusively take route \( e_3 - e_1 \) in the second realization. Particularly, in the first realization, the costs on \( e_3 \) are high in the first few stages, which lead to a high belief on state \( s = e_3 \) (disruption on \( e_3 \)). Consequently, travelers start to take route \( e_2 - e_1 \), and eventually completely learn the true state. On the other hand, in the second realization, the realized costs of \( e_3 \) are low, so no traveler takes route \( e_2 - e_1 \). Therefore, travelers cannot distinguish state \( s = e_2 \) from the true state \( s = \emptyset \), and learning is incomplete.

Moreover, we consider another initial public belief \( \theta^0 = (0, 0.1, 0.0, 0.9) \), which is 90% accurate. Under this initial public belief, travelers take both \( e_2 - e_1 \) and \( e_3 - e_1 \). Fig. 2e - 2f demonstrates the beliefs and equilibrium edge loads of one realization such that the realized cost of \( e_2 \) is high in stage 1, and hence the updated belief \( \theta^1 \) in (5) assigns high probability on state \( s = e_2 \). Consequently, the equilibrium strategy is to take route \( e_3 - e_1 \) exclusively. Then, travelers cannot distinguish the state \( s = e_2 \) from the true state \( s = \emptyset \), i.e. the learning is incomplete.

Additionally, if the edge cost on \( e_2 \) when it is facing disruption were \( \ell^0_e = 2w_2 + 5 \), then, the cost functions satisfy the condition (2) in Proposition 2. We can check from Fig. 2g - 2h that the learning is complete.

5. CONCLUDING REMARKS

In this article, we study how strategic travelers learn the uncertain state after infrastructure disruptions and adjust
their route choices dynamically with the access of a public information system. Our results include the convergence of belief and edge loads, comparison of rest points with complete information equilibrium, and conditions that guarantee complete learning.

All our results hold for networks with multiple origin-destination pairs. One future extension is to permit the noise terms to be realized from a variety of distributions that depend on the state and edge load. Another direction of interest is to analyze the learning dynamics in asymmetric and incomplete information environment. In practice, travelers may obtain private information of the state after infrastructure disruption, or have private perception of the belief provided by the public information system. To solve this problem, we need to analyze how the asymmetric information environment affects travelers’ route choices in each stage and the updates of beliefs.

ACKNOWLEDGEMENTS

We are grateful to Prof. Asu Ozdaglar and Prof. Muhamet Yildiz for useful discussions. The first author was supported by the IDSS Hammer Fellowship, and the second author was party supported by NSF CAREER Award, and AFOSR grant: Building Attack Resilience into Complex Networks.

REFERENCES

Arnott, R., De Palma, A., and Lindsey, R. (1991). Does providing information to drivers reduce traffic congestion? Transportation Research Part A: General, 25(5), 309–318.

Ben-Akiva, M., De Palma, A., and Isam, K. (1991). Dynamic network models and driver information systems. Transportation Research Part A: General, 25(5), 251–266.

Blackwell, D. and Dubins, L. (1962). Merging of opinions with increasing information. The Annals of Mathematical Statistics, 33(3), 882–886.

Cominetti, R., Melo, E., and Sorin, S. (2010). A payoff-based learning procedure and its application to traffic games. Games and Economic Behavior, 70(1), 71–83.

Dafermos, S. and Nagurney, A. (1984). Sensitivity analysis for the asymmetric network equilibrium problem. Mathematical programming, 28(2), 174–184.

Fudenberg, D., Drew, F., Levine, D.K., and Levine, D.K. (1998). The theory of learning in games, volume 2. MIT Press.

Fudenberg, D. and Kreps, D.M. (1995). Learning in extensive-form games I. self-confirming equilibria. Games and Economic Behavior, 8(1), 20–55.

Fudenberg, D. and Levine, D.K. (1993). Self-confirming equilibrium. Econometrica: Journal of the Econometric Society, 523–545.

Jha, M., Madanat, S., and Peeta, S. (1998). Perception updating and day-to-day travel choice dynamics in traffic networks with information provision. Transportation Research Part C: Emerging Technologies, 6(3), 189–212.

Khan, Z. and Amin, S. (2018). Bottleneck model with heterogeneous information. Transportation Research Part B: Methodological, 112, 157–190.

Mahmassani, H.S. and Jayakrishnan, R. (1991). System performance and user response under real-time information in a congested traffic corridor. Transportation Research Part A: General, 25(5), 293–307.

Marden, J.R., Arslan, G., and Shamma, J.S. (2009). Joint strategy fictitious play with inertia for potential games. IEEE Transactions on Automatic Control, 54(2), 208–220.

Marden, J.R. and Shamma, J.S. (2012). Revisiting log-linear learning: Asynchronous, completeness and payoff-based implementation. Games and Economic Behavior, 75(2), 788–808.

Meigs, E., Parise, F., and Ozdaglar, A. (2017). Learning dynamics in stochastic routing games. In Communication, Control, and Computing (Allerton), 2017 55th Annual Allerton Conference on, 259–266. IEEE.

Milchtaich, I. (2006). Network topology and the efficiency of equilibrium. Games and Economic Behavior, 2(57), 321–346.

Monderer, D. and Shapley, L.S. (1996). Potential games. Games and Economic Behavior, 14(1), 124–143.

Sandholm, W.H. (2001). Potential games with continuous player sets. Journal of Economic theory, 97(1), 81–108.

Wu, M., Amin, S., and Ozdaglar, A.E. (2018). Value of information systems in routing games. arXiv preprint arXiv:1808.10590.

Zhu, S., Levinson, D., Liu, H.X., and Harder, K. (2010). The traffic and behavioral effects of the I-35W mississippi river bridge collapse. Transportation Research Part A: Policy and Practice, 44(10), 771–784.