When Congestion Games Meet Mobile Crowdsourcing: Selective Information Disclosure

Presenter: Lingjie Duan, Associate Professor, SUTD

Joint work with Hongbo Li, PhD student
Singapore University of Technology and Design (SUTD)

@ AAAI 2023 Conference, Feb 2023, Washington, DC, USA

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Motivation

1. Dynamic traffic information to learn:
    - Emerging traffic navigation platforms (e.g., Waze and Google Maps) crowdsource mobile users to learn and share their observed traffic conditions.
    - These platforms make all information public, and current users still choose the shortest path (Vasserman, Feldman, and Hassidim 2015; Zhang et al. 2018).
    - Such selfish decisions make the system arbitrarily bad in term of total travel cost.

Motivation

2. Congestion games literature about social planner with complete information of traffic conditions:
    - They implement payment (Ferguson, Brown, and Marden 2022; Li and Duan 2022) or non-monetary mechanism (Tavafoghi and Teneketzis 2017; Li, Courcoubetis, and Duan 2019) on users to regulate selfish routing.
    - Yet they limit attentions on one-shot static scenario with fixed traffic condition to regulate.

Literature Review: Information Sharing

There are some recent works studying information sharing among users in a dynamic scenario:

1. Information learning to accelerate convergence rates to Wardrop equilibrium for stochastic congestion games (Meigs, Parise, and Ozdaglar 2017; Wu and Amin 2019).

Literature Review: Information Sharing

There are some recent works studying information sharing among users in a dynamic scenario:

1. Information learning to accelerate convergence rates to Wardrop equilibrium for stochastic congestion games (Meigs, Parise, and Ozdaglar 2017; Wu and Amin 2019).
   However, these works do not consider mechanism design to motivate users to reach social optimum.

Literature Review: Information Sharing

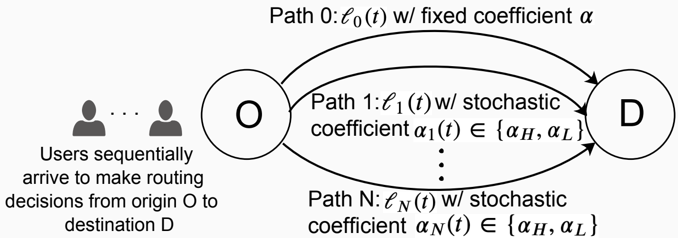
2. Travel cost minimization for multi-armed bandit (MAB) problems (Krishnasamy et al. 2021; Bozorgchenani et al. 2022).

Literature Review: Information Sharing

2. Travel cost minimization for multi-armed bandit (MAB) problems (Krishnasamy et al. 2021; Bozorgchenani et al. 2022).
   However, all of these MAB works overlook users' deviation to selfish routing.

System Model

Dynamic Congestion Model



- Parallel transportation network: one safe path and N risky/stochastic paths.
- Infinite discrete time horizon: $t \in \{1, 2, \cdots\}$.
- Travel latency of path $i \in \{0, 1, \cdots, N\}$ at time t: $\ell_i(t)$.
- Atomic users sequentially arrive to make routing choice: $\pi(t) \in \{0, 1, \cdots, N\}$.

Dynamic Congestion Model

Current travel latency $\ell_i(t)$ of each path $i \in \{0, 1, \cdots, N\}$ has linear correlation with last latency $\ell_i(t-1)$.

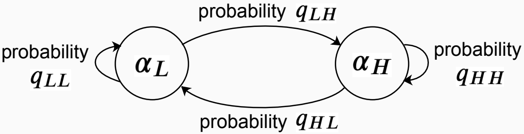
- For safe path 0 with fixed traffic coefficient α,

$$\ell_0(t+1) = \begin{cases} \alpha\ell_0(t) + \Delta\ell, \text{ if } \pi(t) = 0, \\ \alpha\ell_0(t), \text{ if } \pi(t) \neq 0, \end{cases}$$

where constant correlation coefficient $\alpha \in (0, 1)$ measures the leftover flow to be serviced over time, and $\Delta\ell$ is the addition introduced by current user to the next.

Dynamic Congestion Model

- On any risky path $i \in \{1, \cdots, N\}$, its coefficient $\alpha_i(t)$ is stochastic and alternates between $\alpha_L \in (0,1)$ and $\alpha_H \in [1, +\infty)$:



The Markov chain for $\alpha_i(t)$.

Then the travel latency $\ell_i(t+1)$ is estimated as:

$$\ell_i(t+1) = \begin{cases} \alpha_i(t)\ell_i(t) + \Delta\ell, & \text{if } \pi(t) = 0, \\ \alpha_i(t)\ell_i(t), & \text{if } \pi(t) \neq 0. \end{cases}$$

Partially Observable Markov Chain

Define a random observation set $\mathbf{y}(t) = \{y_1(t), \cdots, y_N(t)\}$ for N risky paths, where $y_i(t) \in \{0, 1, \emptyset\}$:

- $y_i(t) = 0$ tells that the current user observes a hazard after choosing path i.
- $y_i(t) = 1$ tells that the user does not observe any hazard on path i.
- $y_i(t) = \emptyset$ tells that this user travels on another path with $\pi(t) \neq i$.

Under the correlation state $\alpha_i(t) = \alpha_H$ or α_L, we respectively denote the probabilities for the user to observe a hazard as:

$$p_H = \mathbf{Pr}(y_i(t) = 1 | \alpha_i(t) = \alpha_H),$$
$$p_L = \mathbf{Pr}(y_i(t) = 1 | \alpha_i(t) = \alpha_L),$$

where $p_L < p_H$.

Partially Observable Markov Chain

The historical data of users' observations $(\mathbf{y}(1), \cdots, \mathbf{y}(t-1))$ and routing decisions $(\pi(1), \cdots, \pi(t-1))$ keep growing in the time horizon.

At the beginning of time t, we translate these data into a prior belief $x_i(t)$ for seeing bad traffic condition $\alpha_i(t) = \alpha_H$ using Bayesian inference:

$$x_i(t) = \mathbf{Pr}\big(\alpha_i(t) = \alpha_H | x_i(t-1), \pi(t-1), \mathbf{y}(t-1)\big).$$

Partially Observable Markov Chain

During time slot t, given prior probability $x_i(t)$, the platform will further update it to a posterior probability $x_i'(t)$ after a new users with $\pi(t)$ shares his observation $y_i(t)$ during the time slot:

$$x_i'(t) = \mathbf{Pr}\big(\alpha_i(t) = \alpha_H | x_i(t), \pi(t), \mathbf{y}(t)\big).$$

Besides the traveled path i, for any other path $y_j(t) = \emptyset$, we keep $x_j'(t) = x_j(t)$.

Partially Observable Markov Chain

At the end of time slot t, the platform estimates the posterior correlation coefficient:

$$\mathbb{E}[\alpha_i(t)|x_i'(t)] = x_i'(t)\alpha_H + (1 - x_i'(t))\alpha_L.$$

Then we obtain the expected travel latency on stochastic path i for time $t+1$ as:

$$\mathbb{E}[\ell_i(t+1)|x_i(t), y_i(t)] = \begin{cases} \mathbb{E}[\alpha_i(t)|x_i'(t)]\mathbb{E}[\ell_i(t)|x_i(t-1), y_i(t-1)] + \Delta\ell, & \text{if } \pi(t) = i, \\ \mathbb{E}[\alpha_i(t)|x_i'(t)]\mathbb{E}[\ell_i(t)|x_i(t-1), y_i(t-1)], & \text{if } \pi(t) \neq i. \end{cases}$$

The platform updates $x_i'(t)$ to $x_i(t+1)$ below:

$$x_i(t+1) = x_i'(t)q_{HH} + (1 - x_i'(t))q_{LH}.$$

POMDP Problem Formulations

Myopic Policy

We summarize the dynamics of expected travel latencies among all $N+1$ paths and the hazard beliefs of N stochastic paths into vectors:

$$\begin{aligned}\mathbf{L}(t) =& \Big\{\ell_0(t), \mathbb{E}\big[\ell_1(t)|x_i(t-1), y_i(t-1)\big], \cdots, \mathbb{E}\big[\ell_N(t)|x_N(t-1), y_N(t-1)\big]\Big\},\\ \mathbf{x}(t) =& \big\{x_1(t), \cdots, x_N(t)\big\}.\end{aligned}$$

We define the best stochastic $\hat{\imath}(t)$ to be the one out of N risky paths to provide the shortest expected travel latency at time t below:

$$\hat{\imath}(t) = \arg \min_{i \in \{1,\cdots,N\}} \mathbb{E}\big[\ell_i(t)|x_i(t-1), y_i(t-1)\big].$$

The selfish user will only choose between safe path 0 and this path $\hat{\imath}(t)$ to minimize his own travel latency.

Myopic Policy Cost Function

We formulate the ρ-discounted long-term cost function since time t under myopic policy as:

$$C^{(m)}\big(\mathbf{L}(t), \mathbf{x}(t)\big) = \begin{cases} \ell_0(t) + \rho Q_0^{(m)}(t+1), \\ \qquad\qquad\qquad \text{if } \mathbb{E}[\ell_{\hat{\imath}(t)}(t) | x_{\hat{\imath}(t)}(t-1), y_{\hat{\imath}(t)}(t-1)] \geq \ell_0(t), \\ \mathbb{E}[\ell_{\hat{\imath}(t)}(t) | x_{\hat{\imath}(t)}(t-1), y_{\hat{\imath}(t)}(t-1)] + \rho Q_{\hat{\imath}(t)}^{(m)}(t+1), \\ \qquad\qquad\qquad \text{otherwise.} \end{cases}$$

Socially Optimal Cost Function

Similarly, we formulate the social cost function under socially optimal policy below:

$$C^*\big(\mathbf{L}(t), \mathbf{x}(t)\big) = \min_{i \in \{1, \cdots, N\}} \big\{ \ell_0(t) + \rho Q_0^*(t+1), \ell_i(t) + \rho Q_i^*(t+1) \big\}.$$

Policies Comparison:
Myopic versus Socially Optimum

Threshold-type Solutions

Lemma (1)
The cost functions $C^{(m)}\big(\mathbf{L}(t), \mathbf{x}(t)\big)$ *and* $C^{*}\big(\mathbf{L}(t), \mathbf{x}(t)\big)$ *under both policies increase with* $\mathbf{L}(t)$ *and* $\mathbf{x}(t)$.

With this monotonicity result, we next prove that both policies are of threshold-type.

Threshold-type Solutions

Proposition (1)
Provided with $\mathbf{L}(t)$ *and* $\mathbf{x}(t)$*, the user under the myopic policy keeps staying with path 0, until the expected latency of the best stochastic path* $\hat{\imath}(t)$ *reduces to be smaller than the following threshold:* $\ell^{(m)}(t) = \ell_0(t)$.

Similarly, the socially optimal policy will choose stochastic path i *if* $\mathbb{E}[\ell_i(t)|x_i(t-1), y_i(t-1)]$ *is less than the following threshold:* $\ell_i^*(t) = \arg\max_z \left\{z|z \le \rho Q_i^*(t+1) - \rho Q_0^*(t+1) - \ell_0(t)\right\}$.

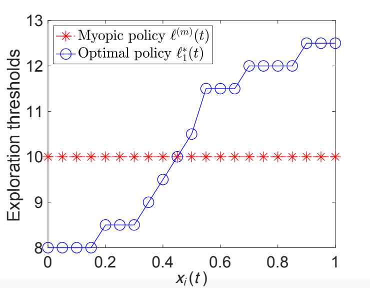
Policies Comparison

Proposition (2)

There exists a belief threshold x^{th} satisfying

$$\min\left\{\frac{\alpha-\alpha_L}{\alpha_H-\alpha_L}, \frac{1-q_{LL}}{2-q_{LL}-q_{HH}}\right\} \le x^{th} \le \max\left\{\frac{\alpha-\alpha_L}{\alpha_H-\alpha_L}, \frac{1-q_{LL}}{2-q_{LL}-q_{HH}}\right\}.$$

As compared to socially optimal policy, if risky path $i \in \{1, \cdots, N\}$ has weak hazard belief $x_i(t) < x^{th}$, myopic users will only over-explore this path with $\ell^{(m)}(t) \ge \ell_i^*(t)$. If strong hazard belief with $x_i(t) > x^{th}$, myopic users will only under-explore this path with $\ell^{(m)}(t) \le \ell_i^*(t)$.



Price of Anarchy (PoA) Analysis

We define the price of anarchy (PoA) to be the maximum ratio between the social cost under myopic policy and the minimal social cost, by searching all possible system parameters:

$$\mathrm{PoA}^{(m)} = \max_{\substack{\alpha,\alpha_H,\alpha_L,q_{LL},q_{HH},\\ \mathbf{x}(t),\mathbf{L}(t),\Delta\ell,p_H,p_L}} \frac{C^{(m)}\big(\mathbf{L}(t),\mathbf{x}(t)\big)}{C^{*}\big(\mathbf{L}(t),\mathbf{x}(t)\big)},$$

which is obviously larger than 1.

Price of Anarchy (PoA) Analysis

Proposition (3)
As compared to the social optimum, the myopic policy achieves $PoA^{(m)} \geq \frac{1}{1-\rho}$, which can be arbitrarily large for discount factor $\rho \to 1$.

In the worst-case PoA analysis, where the myopic policy always chooses safe path 0, but the socially optimal policy frequently explores stochastic path 1 to learn α_L.
(Myopic has zero-exploration of stochastic paths).

Selective Information Disclosure Mechanism Design

Benchmark: Information Hiding Mechanism

In the information hiding policy (Tavafoghi and Teneketzis 2017), the user without any information believes that $x_i(t)$ of any risky path $i \in \{1, \cdots, N\}$ has converged to its stationary hazard belief $\bar{x}$. Then he can only decide his routing policy $\pi^{\emptyset}(t)$ by comparing α to $\mathbb{E}[\alpha_i(t)|\bar{x}]$.

Proposition (4)
This hiding policy leads to $PoA^{\emptyset} = \infty$, regardless of discount factor ρ.

The worst case $\text{PoA}^{\emptyset}$ happens when maximum-exploration, which is opposite to the zero-exploration $\text{PoA}^{(m)}$.

Definition: Selective Information Disclosure (SID) Mechanism

Definition (SID)

1. Unlike the information hiding mechanism, if a user arrival is expected to choose a different route $\pi^{\emptyset}(t) \neq 0$ from optimal $\pi^*(t) = 0$ of path 0, then our SID mechanism will disclose the latest expected travel latency set $\mathbf{L}(t)$ to him.

Definition: Selective Information Disclosure (SID) Mechanism

Definition (1)

1. Unlike the information hiding mechanism, if a user arrival is expected to choose a different route $\pi^{\emptyset}(t) \neq 0$ from optimal $\pi^{*}(t) = 0$, then our SID mechanism will disclose the latest expected travel latency set $\mathbf{L}(t)$ to him.
2. Otherwise, unlike the myopic policy, our mechanism hides $\mathbf{L}(t)$ from this user.

Definition: Selective Information Disclosure (SID) Mechanism

Definition (1)

1. Unlike the full information hiding mechanism, if a user arrival is expected to choose a different route $\pi^{\emptyset}(t) \neq 0$ from optimal $\pi^{*}(t) = 0$, then our SID mechanism will disclose the latest expected travel latency set $\mathbf{L}(t)$ to him.
2. Otherwise, unlike the myopic policy, our mechanism hides $\mathbf{L}(t)$ from this user.
3. Besides, our mechanism always provides optimal path recommendation $\pi^{*}(t)$, without sharing hazard belief set $\mathbf{x}(t)$, routing history $\big(\pi(1), \cdots, \pi(t-1)\big)$, or past observation set $\big(\mathbf{y}(1), \cdots, \mathbf{y}(t-1)\big)$.

Finite PoA

Theorem (1)

Our SID mechanism results in $PoA^{(SID)} \leq \frac{1}{1-\frac{\rho}{2}}$ *, which is always no more than 2.*

Average Inefficiency Ratio

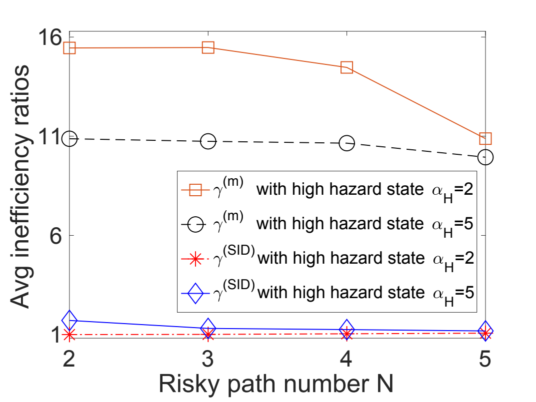
Define the average inefficiency ratios achieved by myopic policy and our SID mechanism:

$$\gamma^{(\mathrm{m})} = \frac{\mathbb{E}\left[C^{(\mathrm{m})}\left(\mathbf{L}(t), \mathbf{x}(t)\right)\right]}{\mathbb{E}\left[C^{*}\left(\mathbf{L}(t), \mathbf{x}(t)\right)\right]},$$

$$\gamma^{(\mathrm{SID})} = \frac{\mathbb{E}\left[C^{(\mathrm{SID})}\left(\mathbf{L}(t), \mathbf{x}(t)\right)\right]}{\mathbb{E}\left[C^{*}\left(\mathbf{L}(t), \mathbf{x}(t)\right)\right]}.$$

Average Inefficiency Ratio

After running 50 long-term experiments for averaging each ratio, we plot the following figure to compare $\gamma^{(m)}$ to $\gamma^{(\mathrm{SID})}$ versus risky path number N.



- As N increases, the travel latencies in risky paths decrease more, making the system better.
- As α_H increases, risky paths differ more from the safe path, such that $\pi^{(m)}$ approaches π^*.

Conclusion

1. Our study extends the traditional congestion games fundamentally to create positive information learning benefit generated by users dynamically.

Conclusion

1. Our study extends the traditional congestion games fundamentally to create positive information learning benefit generated by users dynamically.
2. Myopic routing policy is arbitrarily bad, as its PoA is larger than $\frac{1}{1-\rho}$.

Conclusion

1. Our study extends the traditional congestion games fundamentally to create positive information learning benefit generated by users dynamically.
2. Myopic routing policy is arbitrarily bad, as its PoA is larger than $\frac{1}{1-\rho}$.
3. Our selective information disclosure (SID) mechanism reduces PoA to be less than $\frac{1}{1-\frac{\rho}{2}}$.

Reference

- Bozorgchenani, A.; Maghsudi, S.; Tarchi, D.; and Hossain, E. 2022. Computation Offloading in Heterogeneous Vehicular Edge Networks: On-Line and Off-Policy Bandit Solutions. IEEE Transactions on Mobile Computing, 21(12): 4233–4248.
- Ferguson, B. L.; Brown, P. N.; and Marden, J. R. 2022. The Effectiveness of Subsidies and Tolls in Congestion Games. IEEE Transactions on Automatic Control, 67(6): 2729–2742.
- Krishnasamy, S.; Sen, R.; Johari, R.; and Shakkottai, S. 2021. Learning unknown service rates in queues: A multiarmed bandit approach. Operations Research, 69(1): 315-330.
- Li, H.; and Duan, L. 2022. Online Pricing Incentive to Sample Fresh Information. IEEE Transactions on Network Science and Engineering, forthcoming.
- Li, Y.; Courcoubetis, C.; and Duan, L. 2019. Recommending paths: Follow or not follow? In IEEE INFOCOM 2019-IEEE Conference on Computer Communications, IEEE.

Reference

- Meigs, E.; Parise, F.; and Ozdaglar, A. 2017. Learning dynamics in stochastic routing games. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE.
- Tavafoghi, H.; and Teneketzis, D. 2017. Informational incentives for congestion games. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE.
- Vasserman, S.; Feldman, M.; and Hassidim, A. 2015. Implementing the wisdom of waze. In Twenty-Fourth International Joint Conference on Artificial Intelligence.
- Wu, M.; and Amin, S. 2019. Learning an unknown network state in routing games. IFAC-PapersOnLine.
- Zhang, J.; Lu, P.; Li, Z.; and Gan, J. 2018. Distributed trip selection game for public bike system with crowdsourcing. In IEEE INFOCOM 2018-IEEE Conference on Computer Communications, IEEE.