Learning to Price Vehicle Service with Unknown Demand

Haoran Yu\textsuperscript{1}, Ermin Wei\textsuperscript{2}, and Randall A. Berry\textsuperscript{2}

\textsuperscript{1}School of Computer Science, Beijing Institute of Technology
\textsuperscript{2}Department of Electrical and Computer Engineering, Northwestern University

Oct. 2020 @ACM MobiHoc
Problem
Location-Based Vehicle Service Pricing

- People use vehicle service offered by ride-sharing platforms.

- **Location-based pricing:** It depends on origin-destination pairs.
  - **Purpose:** Balance demand and supply.

Example of origin-based charge: 
price = standard price \( \times \) multiplier
We introduce a traffic graph to illustrate location-based pricing.

- **Node:** location, **link:** traffic demand.

Provider sets different vehicle service prices for different links. Let $p_{ij}$ be the price for link $(i, j)$ ($i$: origin; $j$: destination).

- e.g., $p_{13} = \$1/minute$.
- Can be converted to $\$/mile based on vehicle velocity.

For each link $(i, j)$, actual demand changes with $p_{ij}$. 
We introduce a traffic graph to illustrate location-based pricing.

- **Node**: location, **link**: traffic demand.

Provider sets different vehicle service prices for different links. Let \( p_{ij} \) be the price for link \((i, j)\) (\(i\): origin; \(j\): destination).

- e.g., \( p_{13} = \$1/\text{minute} \).
- Can be converted to \$/\text{mile} based on vehicle velocity.

For each link \((i, j)\), actual demand changes with \( p_{ij} \).
Optimal pricings for links are coupled due to vehicle flow balance.

Example: Suppose $p_{14}$ increases. How should provider change other prices?

- Increase $p_{46}$: to save supply at node 4.
- Decrease $p_{54}$: to increase supply at node 4.

Provider needs to jointly optimize $p_{ij}$ for different links.
Challenge of Unknown Demand

- If mapping from price to demand is known:
  - Example: If $p_{12} = 2$, demand = 100; If $p_{12} = 4$, demand = 50.
  - Given all parameters and topology, can calculate $p_{ij}^*$ for all $(i,j)$. 
If mapping from price to demand is unknown:

- **Example:** If $p_{12} = 2$, demand = ? If $p_{12} = 4$, demand = ?
- **Intuitive solution:** (i) test many prices $p^1_{ij}, p^2_{ij}, \ldots$ to learn mapping; (ii) derive optimal prices based on learned mapping.
- **Challenge:** If do not choose $p^1_{ij}, p^2_{ij}, \ldots$ carefully, the provider’s payoff at initial stage is low.
Challenge of Unknown Demand

- If mapping from price to demand is unknown:
  - **Example:** If $p_{12} = 2$, demand =? If $p_{12} = 4$, demand =?
  - **Intuitive solution:** (i) test many prices $p_{ij}^1, p_{ij}^2, \ldots$ to learn mapping; (ii) derive optimal prices based on learned mapping.
  - **Challenge:** If do not choose $p_{ij}^1, p_{ij}^2, \ldots$ carefully, the provider’s payoff at initial stage is low.
Consider a simplified model with a monopolistic provider.

Design an online pricing policy:
- (i) Can learn accurate user demand for each \((i, j)\);
- (ii) Achieve asymptotically-optimal provider long-term payoff.
Related Work

- Prior work on **vehicle service pricing**: [Banerjee et al. 2015], [Banerjee et al. 2016], [Ma et al. 2018], [Bimpikis et al. 2019], [Yu et al. 2019] etc.
  - **Our work**: Consider unknown user demand.

- Prior work on **pricing with unknown demand**: [Besbes and Zeevi 2009], [Broder and Rusmevichientong 2012], [Den Boer and Zwart 2013] [Keskin and Zeevi 2014] [Khezeli and Bitar 2017] etc.
  - **Our work**: Consider vehicle service, where prices for links are coupled due to vehicle flow balance.

- Prior work on **multi-armed bandit problem**: [Berry and Fristedt 1985], [Kleinberg 2005], [Vermorel and Mohri 2005], [Wang and Huang 2018] etc.
  - **Our work**: Consider an infinite decision space.
Related Work

- Prior work on vehicle service pricing: [Banerjee et al. 2015], [Banerjee et al. 2016], [Ma et al. 2018], [Bimpikis et al. 2019], [Yu et al. 2019] etc.
  - Our work: Consider unknown user demand.

- Prior work on pricing with unknown demand: [Besbes and Zeevi 2009], [Broder and Rusmevichientong 2012], [Den Boer and Zwart 2013] [Keskin and Zeevi 2014] [Khezeli and Bitar 2017] etc.
  - Our work: Consider vehicle service, where prices for links are coupled due to vehicle flow balance.

- Prior work on multi-armed bandit problem: [Berry and Fristedt 1985], [Kleinberg 2005], [Vermorel and Mohri 2005], [Wang and Huang 2018] etc.
  - Our work: Consider an infinite decision space.
Related Work

- Prior work on **vehicle service pricing**: [Banerjee et al. 2015], [Banerjee et al. 2016], [Ma et al. 2018], [Bimpikis et al. 2019], [Yu et al. 2019] etc.
  - **Our work**: Consider *unknown* user demand.

- Prior work on **pricing with unknown demand**: [Besbes and Zeevi 2009], [Broder and Rusmevichientong 2012], [Den Boer and Zwart 2013] [Keskin and Zeevi 2014] [Khezeli and Bitar 2017] etc.
  - **Our work**: Consider **vehicle service**, where prices for links are coupled due to vehicle flow balance.

- Prior work on **multi-armed bandit problem**: [Berry and Fristedt 1985], [Kleinberg 2005], [Vermorel and Mohri 2005], [Wang and Huang 2018] etc.
  - **Our work**: Consider an **infinite decision space**.
Model
A monopolistic provider offers service on day $d = 1, 2, \ldots$

Let $p_{ij}^d$ be service price for $(i, j)$ on $d$-th day ($\$ \text{ per time slot}$).

Realized user demand on $(i, j)$ is

$$\Psi_{ij}^d (p_{ij}^d, \epsilon_{ij}^d) = \alpha_{ij} - \beta_{ij} p_{ij}^d + \epsilon_{ij}^d.$$

- $\alpha_{ij}$ and $\beta_{ij}$ are positive parameters that are unknown to provider and need to be learned.
- $\epsilon_{ij}^d$ is a zero-mean i.i.d. random variable, capturing demand shock. Provider only knows its distribution.
- On each day, provider can only observe $\Psi_{ij}^d (p_{ij}^d, \epsilon_{ij}^d)$.

Assumptions: linear and time-invariant demand.
User Demand

- A monopolistic provider offers service on day $d = 1, 2, \ldots$
- Let $p_{ij}^d$ be service price for $(i, j)$ on $d$-th day ($\$ per time slot$).
- Realized user demand on $(i, j)$ is
  \[ \Psi_{ij}^d \left( p_{ij}^d, \epsilon_{ij}^d \right) = \alpha_{ij} - \beta_{ij} p_{ij}^d + \epsilon_{ij}^d. \]

  - $\alpha_{ij}$ and $\beta_{ij}$ are positive parameters that are unknown to provider and need to be learned.
  - $\epsilon_{ij}^d$ is a zero-mean i.i.d. random variable, capturing demand shock. Provider only knows its distribution.
  - On each day, provider can only observe $\Psi_{ij}^d \left( p_{ij}^d, \epsilon_{ij}^d \right)$.

  Assumptions: linear and time-invariant demand.
Provider Decisions and Constraints

- \( p_{ij}^d \): service price for \((i, j)\) ($ per time slot).
  - Should satisfy \( p_{ij}^d \leq p_{\text{max}} \), e.g., due to government regulation.

- \( w_{ij}^d \): vehicle supply for \((i, j)\), i.e., mass of vehicles departing from \(i\) to \(j\) per time slot.
  - Should satisfy \( w_{ij}^d \geq 0 \) and vehicle flow balance:
    \[
    \sum_j w_{ij}^d = \sum_j w_{ji}^d , \forall i.
    \]
    departure rate  
    arrival rate
  - This couples the provider’s decisions for different links.

- **Assumptions:** full control over vehicles and consideration of system’s steady state.
Provider Payoff

- Provider’s time-average payoff on day $d$ in the steady state
  $= \text{user payment per slot - operation cost per slot}$

\[
\Pi\left(p^d, w^d, \epsilon^d\right) \triangleq \sum_{(i,j)} \xi_{ij} \min \left\{ \Psi_{ij}^d(p_{ij}^d, \epsilon_{ij}^d), w_{ij}^d \right\} p_{ij}^d - \sum_{(i,j)} \xi_{ij} w_{ij}^d c.
\]

- $\xi_{ij}$: travel time from $i$ to $j$ (measured by number of slots).
- $c$: one vehicle’s operation cost per slot.
- $\Psi_{ij}^d(p_{ij}^d, \epsilon_{ij}^d)$: realized demand given price and demand shock.
Provider Target

Provider should choose $p^d$ and $w^d$ in real time to maximize
\[ \lim_{D \to \infty} \mathbb{E} \left\{ \frac{1}{D} \sum_{d=1}^{D} \prod (p^d, w^d, \epsilon^d) \right\}. \]

- **Expectation** is taken with respect to $\epsilon^1, \ldots, \epsilon^D$ and the possible randomness in the provider policy.
We design a policy under which provider’s time-average payoff converges to the optimal objective value of following problem:

\[
\max \mathbb{E}_{\epsilon^d} \left\{ \prod \left( p^d, w^d, \epsilon^d \right) \right\}
\]

s.t. \[
\sum_{j} w^d_{ij} = \sum_{j} w^d_{ji}, \forall i,
\]

\[
w^d_{ij} = \mathbb{E}_{\epsilon^d} \left\{ \Psi^d_{ij} \left( p^d_{ij}, \epsilon^d_{ij} \right) \right\}, \forall i, j
\]

\[
\text{var. } p^d_{ij} \leq p_{\text{max}}, w^d_{ij} \geq 0, \forall i, j.
\]

- **Intuition:** Optimal payoff when provider knows all \(\alpha_{ij}\) and \(\beta_{ij}\).
- **Assumption:** local supply-demand balance.
Our Policy
**Our Online Policy**

On Odd Day

1. Based on historical observations, estimate all $\alpha_{ij}$ and $\beta_{ij}$

2. Decide pricing and supply based on estimated $\alpha_{ij}$ and $\beta_{ij}$

After pricing, observe realized demand

On Even Day

3. Modify last pricing & supply decisions and implement

After pricing, observe realized demand

**Intuition:** balance exploitation and exploration.
Odd Day: Demand Parameter Estimation

- Given historical observations on demand and pricing, estimate \( \alpha_{ij} \) and \( \beta_{ij} \) for each \((i, j)\) by least squares estimation:

\[
\left( \hat{\alpha}^{d-1}_{ij}, \hat{\beta}^{d-1}_{ij} \right) = \text{arg min}_{(\bar{\alpha}_{ij}, \bar{\beta}_{ij})} \sum_{\tau=1}^{d-1} \left( \Psi^T_{ij} (p_{ij}^\tau, \epsilon_{ij}^\tau) - \left( \bar{\alpha}_{ij} - \bar{\beta}_{ij} p_{ij}^\tau \right) \right)^2.
\]

- Observed demand - demand under estimation
Odd Day: Pricing Under Estimated Parameters

(1) Based on historical observations, estimate all $\alpha_{ij}$ and $\beta_{ij}$

(2) Decide pricing and supply based on estimated $\alpha_{ij}$ and $\beta_{ij}$

After pricing, observe realized demand

(3) Modify last pricing & supply decisions and implement

After pricing, observe realized demand
Provider makes decisions based on estimated parameters $\hat{\alpha}_{ij}^{d-1}, \hat{\beta}_{ij}^{d-1}$:

$$\max \sum_{(i,j)} \xi_{ij} \mathbb{E}_{\epsilon_{ij}} \left\{ \min \left\{ \hat{\alpha}_{ij}^{d-1} - \hat{\beta}_{ij}^{d-1} p_{ij}^d + \epsilon_{ij}^d, w_{ij}^d \right\} \right\} p_{ij}^d - \sum_{(i,j)} \xi_{ij} w_{ij}^d c$$

s.t. $\sum_j w_{ij}^d = \sum_j w_{ji}^d, \forall i$, (vehicle flow balance)

$$w_{ij}^d = \hat{\alpha}_{ij}^{d-1} - \hat{\beta}_{ij}^{d-1} p_{ij}^d, \forall i, j$$, (local supply demand balance)

var. $p_{ij}^d \leq p_{\text{max}}, w_{ij}^d \geq 0, \forall i, j$.

After rearrangement, can show problem is convex.
Even Day: Pricing Under Estimated Parameters

(1) Based on historical observations, estimate all $\alpha_{ij}$ and $\beta_{ij}$

(2) Decide pricing and supply based on estimated $\alpha_{ij}$ and $\beta_{ij}$

After pricing, observe realized demand

(3) Modify last pricing & supply decisions and implement

After pricing, observe realized demand
Even Day: Modify Odd Day’s Decisions

- Let $p_{ij}^* \left( \hat{\alpha}^{d-2}, \hat{\beta}^{d-2} \right)$ and $w_{ij}^* \left( \hat{\alpha}^{d-2}, \hat{\beta}^{d-2} \right)$ be the decisions on odd day $d - 1$.

- On each even day $d$, for each $(i, j)$:
  - Implement $p_{ij}^* \left( \hat{\alpha}^{d-2}, \hat{\beta}^{d-2} \right) - \frac{\rho}{\hat{\beta}_{ij}^{d-2}} d^{-\eta}$ as the pricing decision.
  - Implement $w_{ij}^* \left( \hat{\alpha}^{d-2}, \hat{\beta}^{d-2} \right) + \rho d^{-\eta}$ as the supply decision.

- $\rho > 0$ and $0 < \eta < \frac{1}{2}$ are controllable parameters.

- **Intuition**: Adding offset terms facilitates exploring different prices and learning demand parameters.

- The offset terms decay to zero as $d$ increases.
| Performance |
|-------------|
| Performance |
Theoretical Performance: Squared Estimation Error

**Theorem**

For all $d \geq 5$ and $(i, j)$:

$$
\mathbb{E} \left\{ \left\| \left( \hat{\alpha}_{ij}^{d-1}, \hat{\beta}_{ij}^{d-1} \right) - (\alpha_{ij}, \beta_{ij}) \right\|_2^2 \right\} < \Phi_1 (\rho, \eta) \frac{\ln (d - 1)}{(d - 1)^{1-2\eta}}.
$$

The upper bound approaches zero as $d$ goes to infinity.
Theoretical Performance: Time-Average Payoff

Theorem

For all $D > 4 + e^{1-2\eta}$:

$$E\left\{ \frac{1}{D} \sum_{d=1}^{D} \left( \prod \left( p^*, w^*, \epsilon^d \right) - \prod \left( p^d, w^d, \epsilon^d \right) \right) \right\} < \Phi_2 (\rho, \eta) D^{-1} + \Phi_3 (\rho, \eta) (\ln D)^{\frac{1}{2}} D^{\eta-\frac{1}{2}} + \Phi_4 (\rho, \eta) D^{-\eta}.$$

The upper bound approaches zero as $D$ goes to infinity.
Numerical Performance

- Real-world dataset (DiDi Chuxing GAIA Open Data Initiative).
- Compare our policy with:
  - Clairvoyant policy: make decisions with complete information;
  - Myopic policy: choose decisions without adding offset terms;
  - Random policy: choose decisions based on randomly guessed parameters.
- Can see our paper for comparison with more policies (e.g., perturbed myopic policy).
Numerical Performance

Controllable parameters in our policy: $\rho = 2$ and $\eta = 0.45$. 
Conclusion

- Propose an effective online pricing and supply policy that balances exploitation and exploration.

Future directions

- Consider driver side compensation design and learn drivers’ willingness to work.
- Use closed-queueing network to model users’ stochastic demand.
THANK YOU