Privacy-Preserving Personalized Pricing and Matching of Ride-Hailing Platforms

Bing Song[a], Sisi Jian[a,*]

[a] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR
bing.song@connect.ust.hk, cesjian@ust.hk
[*] Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 14, 2024

Keywords: Privacy-preserving; Personalized pricing; Ride-hailing; Differential privacy

1 INTRODUCTION

The proliferation of digital technologies has enabled online service providers to amass unprecedented access to user data, which companies across industries have leveraged to implement personalized offerings and pricing (Lei *et al.*, 2023). In ride-hailing, platforms optimize personalized decisions to enhance revenue, but this raises significant privacy concerns.

Specifically, the risk of third-party impersonation and de-anonymization attacks threatens both platforms and users by enabling the inference of sensitive parameters and personal details. Despite growing consumer demand for data privacy protection, existing techniques like anonymization suffer limitations in providing rigorous guarantees and maintaining utility (Chen *et al.*, 2022).

To address these shortcomings, this study proposes a mathematically rigorous privacy preserving approach for personalized pricing and matching optimization within ride-hailing platforms. Leveraging differential privacy, the proposed method utilizes bounded Laplacian mechanisms and parallel composition to safeguard traveler privacy and platform's sensitive operational data with limited impact on platform revenue optimization. Extensive numerical experiments are conducted to validate the efficacy of the approach in preserving privacy without significantly compromising expected revenue.

2 Personalized Pricing and Matching Model

This section presents the personalized pricing and matching problem formulation for a ride-hailing platform aiming to maximize revenue. The platform receives personal information $\mathbf{x_i}$ about each traveler i, which can include factors such as origin, destination, ride history, credit card information, and other relevant information. These factors influence the traveler's decision to use the ride-hailing service. The platform offers each traveler i a personalized price p_{ij} and expected waiting time w_{ij} for the service matched to driver j. The traveler's utility u_{ij} depends on $\mathbf{x_i}$, p_{ij}, and w_{ij}, captured by parameters $\boldsymbol{\alpha_i}$, $\boldsymbol{\beta_i}$, and $\boldsymbol{\gamma_i}$, which respectively characterizes the intrinsic value of traveler i obtained from using the service, traveler's sensitivity to the price and sensitivity to the waiting time. The platform's decision-making involves batch processing of ride requests, solving an optimization problem to match travelers with drivers, determining the personalized service price. The platform's revenue maximization problem is given as below.

$$max \ \sum_{i=1}^{M}\sum_{j=1}^{N} \ p_{ij}y_{ij}(\frac{\mathrm{e}^{\sum_{j=1}^{N}((\boldsymbol{\alpha}_{i}^{T}\mathbf{x_i}-\boldsymbol{\beta}_{i}^{T}\mathbf{x_i}p_i-\boldsymbol{\gamma}_{i}^{T}\mathbf{x_i}w_{ij})y_{ij})}}{1+\mathrm{e}^{\sum_{j=1}^{N}((\boldsymbol{\alpha}_{i}^{T}\mathbf{x_i}-\boldsymbol{\beta}_{i}^{T}\mathbf{x_i}p_i-\boldsymbol{\gamma}_{i}^{T}\mathbf{x_i}w_{ij})y_{ij})}}) \tag{1}$$

$subject \ to$

$$\sum_{j=1}^{N} y_{ij} \leq 1 \ \ for \ \ i = 1, 2 \ldots M \tag{2}$$

$$\sum_{i=1}^{M} y_{ij} \leq 1 \ \ for \ \ j = 1, 2 \ldots N \tag{3}$$

$$y_{ij} \in \{0, 1\} \tag{4}$$

$$p_{ij} \leq p_{ij}^{u} \tag{5}$$

$$p_{ij} \geq p_{ij}^{l} \tag{6}$$

The objective function of the optimization model is to maximize the platform's total expected revenue. Constraints (3) and (4) ensure that each passenger is matched with at most one driver, and each driver is matched with at most one passenger. Constraint (5) specifies the binary nature of the matching decision variable. Constraints (6) and (7) require the price charged to each rider to be within an acceptable range.

It is important to note that this model relies on travelers' personal information to enable personalized pricing and matching, which poses a risk of exposing traveler privacy and platform decision-making strategy. This study proposes a novel method that not only achieves robust user privacy preservation, but also has a limited impact on the platform's expected revenue.

3 Methodology

3.1 Preliminaries on Differential Privacy

Differential privacy is a mathematically rigorous framework for quantifying and preserving privacy, which ensures the practical infeasibility of distinguishing between databases that differ by only a single entry, through the careful injection of calibrated noise into the outputs of differentially private algorithms (Dwork *et al.*, 2006).

Definition 1 ϵ-Differential Privacy. For $\epsilon > 0$, a randomized algorithm $\mathcal{F}$ satisfies ϵ-differential privacy if for every pair of neighboring databases $\mathcal{X}, \mathcal{X}^{'}$ and all subsets $\mathcal{S} \subseteq Range(\mathcal{F})$, it holds that

$$Pr[\mathcal{F}(\mathcal{X}) \subseteq \mathcal{S}] \leq e^{\epsilon} Pr[\mathcal{F}(\mathcal{X}^{'}) \subseteq \mathcal{S}]$$

The privacy parameter ϵ controls the degree of privacy provided by the differential privacy definition, with smaller values requiring higher similarity in algorithm outputs and thus stronger privacy guarantees. Differential privacy is commonly employed to address specific queries. The most straightforward approach to achieving differential privacy for such queries is by introducing random noise, such as the Unbounded Laplace mechanism, to their answers.

Definition 2 Unbounded Laplace Mechanism. For a function $\mathcal{F}$ mapping datasets $\mathcal{X}$ to real numbers, the following definition of $\hat{\mathcal{F}}(\mathcal{X})$ satisfies ϵ-differential privacy:

$$\hat{\mathcal{F}}(\mathcal{X}) = \mathcal{F}(\mathcal{X}) + Lap(\frac{s}{\epsilon})$$

3.2 Privacy-Preserving Personalized Pricing and Matching Approach

To address the limitations of standard differential privacy techniques, this paper proposes a privacy-preserving personalized pricing and matching approach that utilizes a bounded Laplace mechanism combined with parallel composition (Holohan *et al.*, 2018). A key limitation of standard approaches is that they can introduce excessive noise (Tan & Yang, 2024), which reduces data usability and ultimately decreases the platform's expected revenue. The idea of our proposed method is to design a query mechanism $\mathcal{F}$ that maps the database $\mathcal{X}$ to a finite codomain $\mathcal{Y}$, ensuring the noisy outputs can be meaningfully mapped back to the original domain, in contrast to the unbounded noise of the traditional Laplace mechanism.

Definition 3 Bounded Laplace(BL) Mechanism. Formally, for any parameters $y \in [y^l, y^u], b > 0$, a random variable is a BL mechanism ($\mathcal{Z} \sim BL(y, [y^l, y^u]; b)$) if its probability density function is

$$f_{BL}(z) = \begin{cases} exp(-\frac{|z-y|}{b}) / \int_{y_l}^{y_u} exp(-\frac{|v-y|}{b}) dv, & \text{if } z \in [y^l, y^u] \\ 0, & \text{otherwise} \end{cases}$$

In this study, we employ the Bounded Laplace mechanism to preserve privacy for the personalized pricing p_{ij} and expected waiting time w_{ij}, which are bounded within specific intervals. For pricing, we introduce the parallel composition approach due to disparities in price bounds, partitioning the personalized prices into homogeneous chunks and applying the Bounded Laplace mechanism separately within each chunk to achieve ϵ_1-differential privacy. For waiting time, we can directly apply the Bounded Laplace mechanism to achieve ϵ_2-differential privacy, as we only consider drivers within a certain distance during matching. The detailed algorithm is provided in Algorithm 1.

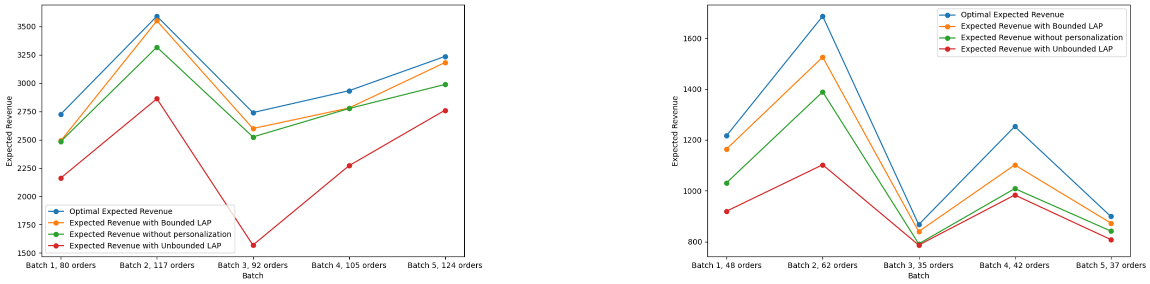
Theorem 1 Parallel Composition. Formally, if the privacy-preserving personalized mechanism $\mathcal{F}_i$ satisfies ϵ_1-differential privacy and if the personalized prices dataset $\mathcal{P}$ is split into k chunks such that $\mathcal{P}_1 \cup \mathcal{P}_2 \cup \ldots \cup \mathcal{P}_k = \mathcal{P}$, then the sequence of $\mathcal{F}_i(\mathcal{X} \cap \mathcal{P}_i)$ satisfies ϵ_1-differential privacy.

```
Algorithm 1: Personalized Pricing and matching
  Input  : (i)personal information set $\mathcal{X}$, (ii) privacy budget for price $\epsilon_1$ and privacy
           budget for waiting time $\epsilon_2$, (iii) parameter set $\mathcal{C}$
  Output: (i)privacy-preserving personalized price $\hat{p}_i$ for traveler $i$, (ii)privacy-preserving
           waiting time $\hat{w}_{ij}$ for traveler $i$ matching driver $j$
Step 1: Optimization: Compute the personalized price $p_i^*$ for all traveler and obtain the
  matching result with waiting time $w_{ij}^*$ if traveler $i$ is matched with driver $j$;
Step 2: Privatization of price: for $All\ p_i^* \in \mathcal{P}$ do
    Step 3: Partition the obtained set of personalized price set $\mathcal{P}$ into different chunks,
      where prices $p_i^* \in \mathcal{P}$ with the same lower and upper bounds will be assigned to the
      same chunk;
    Step 4: Calculate the global sensitivity of each chunk;
    Step 5: For each chunk, a bounded Laplace mechanism with a privacy budget of $\epsilon_1$ is
      applied separately, resulting in a privatized personalized price $\hat{p}_i$.;
end
Step 6: Privatization of waiting time: for $All\ w_{ij}^* \in \mathcal{W}$ do
    Step 7: Calculate the global sensitivity of $\mathcal{W}$;
    Step 8: On the set of waiting times $\mathcal{W}$, the bounded Laplace mechanism with a
      privacy budget of $\epsilon_2$ is employed, resulting in a privatized personalized waiting time
      $\hat{w}_{ij}$ ;
end
```

4 Results and Discussion

A series of numerical experiments are conducted to demonstrate the effectiveness of the proposed algorithm. Four different methods are compared: (1) without any privacy protection, (2) using the traditional unbounded Laplace mechanism for privacy protection, (3) using the bounded Laplace mechanism and parallel composition for privacy protection, and (4) without personalized pricing and waiting time.

The numerical experiments are conducted on the 2017 Haikou City ride-hailing platform dataset. Since the dataset does not include user personal information, we employed a setup similar to prior work (Zha *et al.*, 2018), where user personal information is a two-dimensional vector with each dimension randomly generated. We conduct the experiments during both peak (July) and off-peak (December) seasons (note that Haikou is a tourism city), with different parameter settings for the traveler utility function. The final experimental results are shown in the figure below.



(a) *Peak Season*

(b) *Off-peak Season*

Figure 1 – *Expected Revenue of Haikou*

As illustrated in Figure 1, the adoption of the traditional unbounded Laplace mechanism would introduce excessive noise, significantly impacting the company's revenue, which could even fall below the expected revenue without personalized pricing. In this scenario, the company would be unlikely to prioritize user privacy protection. In contrast, the use of the bounded Laplace mechanism does not seriously affect the company's optimal expected revenue, which remains higher than the revenue without personalized pricing. Given the limited impact on expected revenue, the ride-hailing platform, considering the potential user attrition, business strategy loss, and ethical and legal risks associated with user privacy risk, can effectively address these issues by adopting the approach proposed in this study.

References

Chen, Xi, Simchi-Levi, David, & Wang, Yining. 2022. Privacy-preserving dynamic personalized pricing with demand learning. *Management Science*, **68**(7), 4878–4898.

Dwork, Cynthia, McSherry, Frank, Nissim, Kobbi, & Smith, Adam. 2006. Calibrating noise to sensitivity in private data analysis. *Pages 265–284 of: Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*. Springer.

Holohan, Naoise, Antonatos, Spiros, Braghin, Stefano, & Mac Aonghusa, Pól. 2018. The bounded laplace mechanism in differential privacy. *arXiv preprint arXiv:1808.10410*.

Lei, Yanzhe, Miao, Sentao, & Momot, Ruslan. 2023. Privacy-preserving personalized revenue management. *Management Science*.

Tan, Chaopeng, & Yang, Kaidi. 2024. Privacy-preserving adaptive traffic signal control in a connected vehicle environment. *Transportation research part C: emerging technologies*, **158**, 104453.

Zha, Liteng, Yin, Yafeng, & Du, Yuchuan. 2018. Surge pricing and labor supply in the ride-sourcing market. *Transportation Research Part B: Methodological*, **117**, 708–722.