Neurocomputing 406 (2020) 371–377

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Privacy preserving online matching on ridesharing platforms

Yi Xu, Shuyue Wei, Yansheng Wang*

State Key Laboratory of Software Development Environment, Beihang University, Beijing 100191, China

ARTICLE INFO

Article history:
Received 31 May 2019
Revised 24 August 2019
Accepted 3 September 2019
Available online 12 April 2020

Keywords:
Spatial crowdsourcing
Online bipartite matching
Privacy preserving

ABSTRACT

Ridesharing platforms, as typical applications of spatial crowdsourcing, are becoming more and more popular in the era of mobile internet and sharing economy. One of the most fundamental issues on ridesharing platforms is to assign orders to drivers, which can be naturally modeled as online bipartite matching problem. However, conventional online matching algorithms usually lack data privacy protection mechanisms. This has become a serious issue since the spatiotemporal data of passengers is often sensitive. New policies such as EU's General Data Protection Regulation (GDPR) also enforce protection of sensitive data, which further exacerbate the privacy issues. To deal with the problems, in this paper we propose a framework based on differential privacy (DP) techniques to preserve the privacy of individuals on ridesharing platforms. Specifically, we devise a novel approach to perturb locations in online minimum bipartite matching problem and theoretically show that the performance of the perturbed matching algorithm has the same magnitude with the unperturbed one. Experiments conducted on real datasets have also shown the effectiveness of proposed framework.

1. Introduction

With the emergence of mobile Internet and sharing economy, spatial crowdsourcing is becoming more and more popular in recent years. In spatial crowdsourcing, the platform assigns orders of requesters to crowd workers. For example, Gigwalk [1] and Gmission [2] recruit crowd workers to take pictures or check information at a specific location. OpenStreetMap [3] mobilizes voluntary crowd workers to edit world map together and Waze [4], another digital map application, uses crowd workers to collect real-time traffic information. Ridesharing platforms, such as the real-time taxi-calling service Uber [5], are one of the most typical applications of spatial crowdsourcing, where passengers dynamically appear and need to be assigned to drivers in minutes. We refer to such problem that requires instant response as *online assignment*. It can be naturally modeled as online bipartite graph matching problem with objectives like minimizing the total waiting time of passengers. It is essential to various services on ridesharing platforms such as on-demand taxi-hailing [6,7] and route planning [8,9]. Unfortunately, online matching often requires the personal locations of passengers or drivers, which may result in privacy leaks. The privacy issue is one of the most concerned problems in the era of Internet. Millions of personal information and sensitive data are in danger of exposure every day. To deal with the problem, new policies such as the EU's General Data Protection Regulation (GDPR) have been put forward to enforce protection of sensitive data. Anyone who violates the law will be in face of huge fines. Therefore, traditional online matching algorithms that ignore protecting the privacy of individuals will be in challenge.

To deal with the privacy issue in spatial crowdsourcing, especially on ridesharing platforms, an alternative is to design privacy-preserving mechanisms for the task assignment problem, to make the locations of individuals indistinguishable. However, such mechanisms can lead to inaccurate data, resulting in poor performance of matching algorithms. Existing research on privacy preserving focuses on decreasing the level of privacy leaks, but ignores the errors that the noise brings to the utility, which may result in a great increasement of cost for platforms.

As an improvement, this paper studies the influence of privacy preserving mechanisms on the cost of online minimum bipartite matching (OMBM) algorithms, which is essential to task assignment on ridesharing platforms. Our contributions are as follows:

1. *Privacy preserving framework for OMBM*. To our best knowledge, we are the first to address privacy protection on OMBM problem.
2. *Theoretical performance guarantee*. We first prove that the loss of cost caused by the noise can be bounded (Theorem 1), which means that the performance of the privacy-preserving mechanism on OMBM problem is guaranteed.

* Corresponding author.
E-mail addresses: xuy@buaa.edu.cn (Y. Xu), weishuyue@buaa.edu.cn (S. Wei), arthur_wang@buaa.edu.cn (Y. Wang).

https://doi.org/10.1016/j.neucom.2019.09.116
0925-2312/

3. *Experimental verification*. We conduct experiments on real datasets and verify the effectiveness of our proposed methods.

The rest of this paper is organized as follows. In Section 2 we introduce the background on privacy preserving and minimum bipartite matching problem. In Section 3 we introduce the proposed framework. In Section 4 we make theoretical analysis of our methods. In Section 5 we evaluate our methods with experiments on real datasets. Related works are shown in Section 6. Finally, we conclude this paper in Section 7.

2. Preliminaries

In this section, we first introduce the background on GEO-Indistinguishablity (GEO-I) [10], a notion of location privacy. Then we review the formal definition of Online Minimum Bipartite Matching (OMBM) problem.

2.1. Background on GEO-Indistinguishablity

Based on Differential Privacy (DP), GEO-I is a probabilistic model, making all input data indistinguishable. From the perspective of an attacker, it is impossible to distinguish any two individuals' actual locations by distribution of their reported locations. A mechanism K satisfying ϵ-GEO-I outputs similar distribution for any actual location within radius r, where ϵ indicates the protecting level. The raw input $\mathcal{X}$ is actual locations of requesters. $\mathcal{Z}$, the output of K, are locations reported to the assignment platform. Particularly, nobody but the requester and the worker has direct access to the actual location. In this setting, neither the platform nor malicious attackers can acquire requesters' actual location. The formal definition of GEO-I is as below:

Definition 1 (ϵ-GEO-Indistinguishablity[10])**.** *A mechanism K satisfies ϵ-Geo-indistinguishablility iff. $\forall x, x' \in \mathcal{X}$ such that $d(x, x') \leq r$:*

$$d_P(K(x), K(x')) \leq \epsilon d(x, x')$$

For simplification, we take the Euclidean distance between x and x' as $d(x, x')$. $d_P(K(x), K(x'))$ measures the distance between distributions produced by x and x' (i.e. distinguishablity between x and x'). For example, in [10], it is the multiplicative distance.

Laplace Mechanism naturally satisfying ϵ-GEO-I, by injecting random noise into each original location[10]. In Laplace Mechanism, the reported location z_0 which is perturbed from the original location x_0 satisfies planar Laplace distribution. When the original locations are x_0 and x_0' respectively, the mechanism ensures that the probability of distinguishing different original locations is at most proportional to a multiplicative factor $e^{\epsilon d(x,x')}$, and as a result, x_0 and x_0' are almost indistinguishable to the attacker. We refer to the probability density function of the noise as *planar Laplacian centerer* at x_0.

$$d_\epsilon(x_0, z_0) = \frac{\epsilon^2}{2\pi} e^{-\epsilon d(x_0, z_0)}$$

,where $\frac{\epsilon^2}{2\pi}$ is a normalization factor and z_0 is the output.

2.2. Online minimum bipartite matching

In this part, we review the formal definition of on online minimum bipartite matching problem and then illustrate the problem by an instance of taxi-calling.

Definition 2 (OMBM Problem [11])**.** Given a set of service providers W with specific initial locations, a set of requesters T, which appear dynamically, and a distance function $dis(.,.)$ in 2D space. The OMBM problem is to find a matching M, with minimum total distance $Cost(M) = \Sigma_{t \in T, w \in W} dis(t, w)$, between the matched pairs such that the matching M satisfies the following two constraints:

(a) Real-time constraint: When a task appears, the platform must immediately assign a service provider to the requester before the arrival of next requester.
(b) Invariable constraint: Once a service provider is allocated to a requester, the allocation cannot be revoked.

The real-time constraint is reasonable, for the requester is less willing to wait and expects to be assigned to a worker as soon as she/he arrives. We further illustrate this problem by an example.

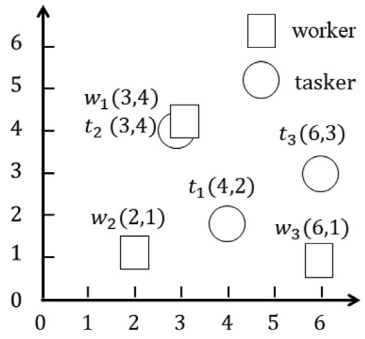


Fig. 1. Locations of Clients and Service providers.

Example 1. *We take the real-time taxi-calling service as an example. Suppose there are three service providers (w_1, w_2, w_3) in the taxi dispatching platform and three tasks (t_1, t_2, t_3) appear in a certain order. The locations of service providers and requesters on the 2D plane are shown in Fig. 1. In the offline scenario, optimal matching pairs are $\{(t_1, w_2), (t_2, w_1), (t_3, w_3)\}$ with minimum total traveling distance $\sqrt{13} + 2 \approx 5.61$. In an online scenario, each requester must be assigned to a service provider as soon as she/he arrives. The simplest way of matching, the greedy strategy, is to assign the nearest unmatched service provider to the requester. With task arrival order (t_1, t_2, t_3), the matching pairs of greedy strategy is (t_1, w_1), (t_2, w_2), (t_3, w_3), making the total travel distance equal to $\sqrt{2} + \sqrt{17} + 2 \approx 7.53$.*

3. Privacy preserving framework for OMBM

In this section, we propose a framework for unilateral privacy protection for OMBM problem. In this framework, any possible matching algorithms can be applied.

Specifically, we take three steps to conduct private matching. Firstly, the client injects noise to her/his original location and then reports the perturbed location to an untrusted third-party platform. Secondly the platform executes an online matching algorithm and assigns an unmatched service provider to the client based on the perturbed locations. Finally, the service provider establishes a direct connection with the client and provides service for the requester. More specifically, the requesters can inform the drivers their actual locations, and then the drivers will pick them up, while the platform only has access to the perturbed locations [12]. The following example in Fig. 2 shows the process of privacy preserving matching.

Example 2. *When a client t(4,1) arrives, she/he firstly adds noise to her/his original location and gets the perturbed location (5,3). Then the client reports the perturbed location to the untrusted platform and the platform assigns w_1 to t according to a matching algorithm. As a result, the travel distance is $\sqrt{17}$, while without noise, the platform assigns w_2 to t and the minimum travel distance is 2.*

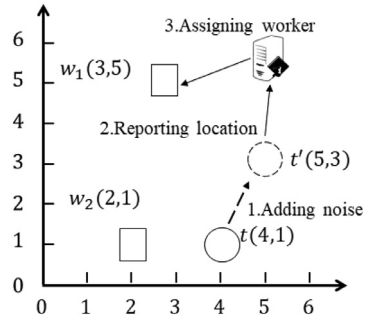


Fig. 2. Process of privacy preserving matching.

From Example 2, we can learn that the platform executes algorithm on inaccurate locations of clients, so the platform may select a further service provider and the total travel distance of matching is larger than that without privacy mechanisms. However, in Section 4 we will show that the distance loss of privacy preserving is limited.

Existing online matching algorithms can be divided into two groups. One group conducts matching on the original metric space and the other runs on HST metric space. In our framework, all of these online algorithms can be private. Algorithm 1 explicitly

Algorithm 1 Private Matching.

Input: Service providers W,Clients T
Output: a matching M
Init $M = \emptyset$
for each newly arrived client $t_i \in T$ **do**
Drawn noise d from some distribution.
Report perturbed location $z_i = (x_{t_i}, y_{t_i}) + d$
Run online algorithm on z_i and W and get a matching pair (t_i, w_i)
$M = M \cup \{(t_i, w_i)\}$
$W = W \backslash \{w_i\}$
end for
return M

describes how to apply privacy-preserving mechanisms to existing online bipartite minimum matching algorithms. We can choose the distribution of noise d, such as the Gaussian noise and the Laplacian noise, which are commonly used in DP. In this paper, we adopt Laplacian noise, and describe how to generate it in the part of experiments. Next, we will make theoretical analysis of our framework.

4. Theoretical analysis

In this section, we first review Competitive Ratio (CR) of OMBM problem and then we calculate CR of private matching, which shows that even if the locations are perturbed, CR of existing algorithms for OMBM problem does not change in magnitude.

4.1. Definitions

The arrival order of clients influences the matching result and the total utility. To evaluate the performance of an online algorithm with different arrival order, two arrival models, namely adversarial model (the worst-case analysis) and random order model (the average-case analysis) have been proposed. In [11], the authors show that an online algorithm with bad performance in adversarial model still works in practice, as the worst case rarely happens, so we adopt the random order model to analyze the performance of private matching.

The *Competitve ratio* of an online algorithm is the ratio between the cost of that algorithm and the optimal cost in offline scenario, which can indicate the performance of an algorithm. We formally define CR in the random order model as follows:

Definition 3 (CR in the Random Order Model)**.**

$$CR_{RO} = \max_{\forall G(T,W)} \frac{\mathbb{E}[Cost(I)]}{OPT(I)}$$

where G(T, W) is an arbitrary bipartite graph. The weight of an edge in G(T, W) is the distance between two objects in T and W. I is an instance of bipartite graph. $\mathbb{E}[Cost(I)]$ *is the expected total travel distance of an algorithm over all possible arrival orders of T. OPT(I) is the offline optimal total cost.*

4.2. Main results

In this subsection, we prove that the performance of existing algorithms under the privacy preserving framework will not be largely violated by the noise. In other words, existing online minimum bipartite algorithms still work under the framework of privacy preserving in theory. Our main results will be given after the following lemma.

Lemma 1. *Given* $D = (x, y)$ *with* $x \sim Lap(0, \lambda)$ *and* $y \sim Lap(0, \lambda)$*, we have the inequality* $\mathbb{E}[|D|] \le 2\lambda$*.*

Here the two components of D are drawn from Laplacian distribution, centered at zero, and independent. λ *is the scale parameter. And* $\mathbb{E}[|D|]$ *is the expected vector magnitude.*

Proof.

$$\begin{aligned}\mathbb{E}[|D|] &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{\sqrt{x^2+y^2}}{4\lambda^2} e^{-(|x|+|y|)} dxdy \\ &\le 4\int_{0}^{+\infty}\int_{0}^{+\infty} \frac{\sqrt{x^2+y^2+2xy}}{4\lambda^2} e^{-(x+y)} dxdy \\ &= 4 \times \frac{\lambda}{2} = 2\lambda\end{aligned}$$

□

By the lemma, we get the expected distance with noise drawn from Laplacian distribution, which measures the effect of noise.

Theorem 1. $\frac{\mathbb{E}[Cost(I')]}{\mathbb{E}[Cost(I)]} \le \eta$*, where* η *is a constant.*

Remark 1. The theorem shows that, in OMBM problem, the expected total distance between perturbed unilateral locations, is at most η times worse than that on original locations. The assumption of expected distance is reasonable in real life, because the clients are unlikely to appear very close to workers meanwhile a client will not be assigned to a worker that is too far away from her.

Proof. Let I denote any instance of OMBM problem and let I' denote the instance in the view of the platform. t_i' denotes a virtual client whose location is the perturbed location reported by t_i. From the perspective of the platform, the purpose is to assign service providers to these virtual clients with minimum total distance. $u(t_i)$ is the service provider matched to t_i.

According to the property of the metric space, we have the triangular inequality:

$$dis(t_i', u(t_i')) \le dis(t_i', t_i) + dis(t_i, u(t_i))$$

We can calculate the upper bound of $\mathbb{E}[Cost(I')]$ by the inequality:

$$\mathbb{E}[Cost(I')] = \mathbb{E}[\sum_i dis(t_i', u(t_i'))]$$

$$\leq \sum_i \mathbb{E}[dis(t_i', t_i) + dis(t_i, u(t_i))]$$
$$= \sum_i \mathbb{E}[dis(t_i', t_i)] + \sum_i \mathbb{E}[dis(t_i, u(t_i))] \quad (1)$$

Then we can calculate the private CR as follows.

$$\eta = \frac{\mathbb{E}[Cost(I')]}{\mathbb{E}[Cost(I)]} = \frac{\mathbb{E}[\sum_i dis(t_i', u(t_i'))]}{\mathbb{E}[\sum_j dis(t_j, u(t_j)]}$$
$$\leq \frac{\sum_i \mathbb{E}[dis(t_i', t_i)] + \sum_i \mathbb{E}[dis(t_i, u(t_i))]}{\sum_j \mathbb{E}[dis(t_j, u(t_j)]}$$
$$= \frac{c_u}{c_l} + \frac{2\lambda}{c_l}$$

□

Here c_u is the upper bound of the mathcing distance, while c_l is the lower bound. Since in practice the real matching distance is within a certain range, c_u and c_l are both constants.

Eventually, we have proved our main results.

5. Experimental results

In this section, we study the performance of four representative algorithms (two on original metric space and two on HST metric space) on OMBM problem, all with or without the privacy preserving framework. The good performance also verifies the theory in Section 4.

5.1. Experimental settings

Real Dataset We use the taxi-calling data on ShenZhou real-time taxi-calling platform [13], in four weeks in May 2015 in Beijing, as the real dataset. It contains 15,802 taxi-calling requests and 1263 private taxis. We take the private taxis as service providers. To satisfy online settings, we assume that once a task is finished, the taxi and the requester in the task disappear from the platform. So when a taxi finishes a request and re-appears on the platform, she/he will be regarded as a new service provider. A taxi serves 10–15 requests a day in the dataset, making it in average 15,364 service providers each day. The number of requesters and service providers is roughly equal in our setting. Specifically, we divide 24 hours of a day into four time periods, 12AM-6AM, 6AM-12PM, 12PM-6PM, 6PM-12AM, and conduct experiments on each period. In this four periods, the number of taxis in early morning is relatively small and the data size of the 12PM-6PM time period is more than 5 times that of the 12AM-6AM time period. In order to evaluate the performance of algorithms under different levels of privacy protection, we choose typical values of ϵ, where $\epsilon \in \{0.1, 0.4, 0.7, 1.0\}$. As analyzed in the lemma in Section 4, we add one-dimensional Laplace noise to two components of the requesters' coordinate respectively to get the perturbed locations. The experimental settings are inspired by [14,15].

Compared Algorithms A metric space can be denoted as (V, d), where V is a set and d is a function of $V \times V \rightarrow [0, \infty)$. The metric space satisfies three properties: (1) $d(u, v) = 0$ iff. $u = v (u, v \in V)$, (2) $d(u, v) = d(v, u)$, and (3) $d(u, w) + d(w, v) \leq d(u, v)$. In this paper, we take the Euclidean Space as the original metric space. HST metric space is also commonly applied. An HST metric space (V', d_T) provides several properties to make better theoretical guarantees [16]. For example, if an HST metric space (V', d_T) is converted from the original (V, d), it guarantees $\mathbb{E}[d_T(u, v)] \geq \mathcal{O}(log|V|)d(u, v)$. Algorithms executed on HST metric space with appropriate metric space projection are similar to those on original metric space. We run the following algorithms including optimal solution in offline scenario and implement each algorithm varying privacy budgets and time period. The details of compared algorithms are shown as follow.

(1) *Optimal Algorithm*. If we remove the real-time constraint of the OMBM problem, every requester can tolerate infinity waiting time. Under this condition, an offline algorithm, such as *Hungarian algorithm*, can be directly executed on the arrival of last request. Obviously, the result of such offline algorithm is better than any online algorithm. Therefore we take the cost of offline minimum bipartite matching as the optimal cost.
(2) *Algorithms executed on HST Metric Space*. HST-Greedy and HST-Reassignment are proved to have upper bound on total distance in random model and we take them as the baseline in HST metric pace.
(3) *Algorithms executed on Original Metric Space*. Greedy has been shown to be the closest to the optimal solution, which has better performance than HST-Greedy and HST-Reassignment in practice [11]. We take deterministic Greedy and randomized Greedy as the baselines on original metric space.

5.2. Experimental results

The results of algorithms are shown in the following figures. The figures in each line (e.g Fig(a), Fig(b), and Fig(c)) show the cost of optimal algorithm, algorithms executed on original metric space and algorithms executed on HST metric space with the same level of privacy protection. The figures in each column (e.g Fig. 3(a), Fig. 4(a), Fig. 5(a) and Fig. 6(a)) show the cost of a certain type of algorithm with different privacy budget.

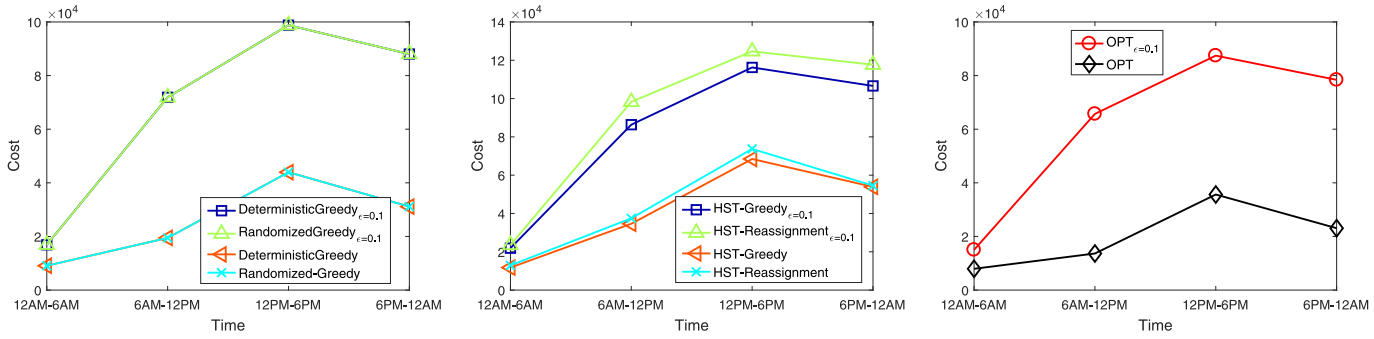
The maximum cost ratio in our experiments is **3.64**, which means the total distance under privacy preserving framework is at most 3.64 times of that without privacy-preserving approaches. *Impact on different baseline algorithms* From the figures in each line, the cost of different algorithms has a similar trend, both in perturbed locations and original locations, which means that the loss of cost caused by privacy-preserving method is slightly influenced by the algorithm itself, but largely influenced by the dataset. The observation of the trend also confirms our theoretical derivation in Chapter 4, i.e. the loss of cost is determined by the noise on the dataset. The number of requests and service providers varies over time, but the cost ratio of the privacy-preserving algorithm and the non privacy-preserving algorithm are almost the same. *Impact on privacy budgets* We observe that with the privacy budget decreasing, the cost of the private algorithm gradually approaches the cost of non-private algorithms. When ϵ is 0.4, the cost ratios of each algorithm in different periods are close to 1. That is, under the privacy protection framework, the performance of the OMBM algorithm is almost unaffected, which demonstrates the effectiveness of our framework.

6. Related work

In this section, we review literatures from two fields, online task assignment and location privacy.

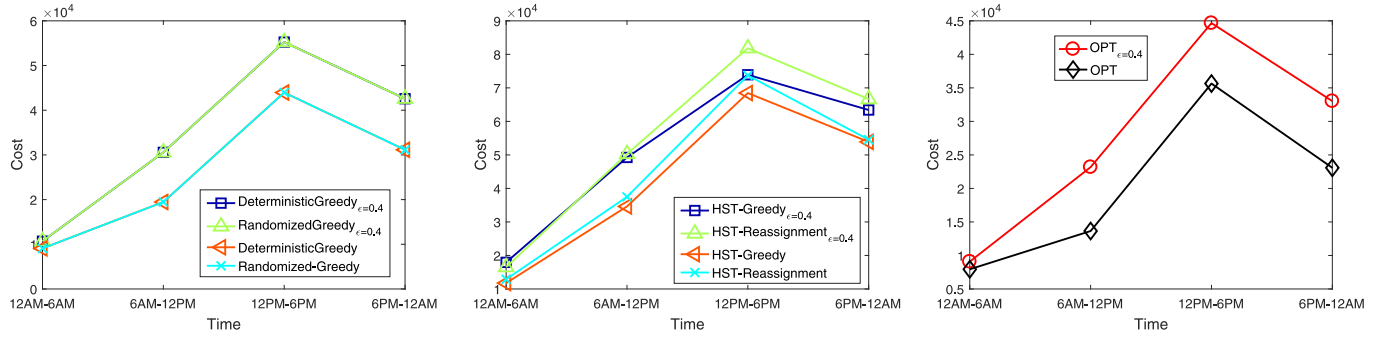
Online task assignment. Online task assignment is one of the key issues in spatial crowdsourcing [17–19]. Apart from minimizing the cost, maximizing matching number is another important issue in online assignment. Based on the two important issues, some works [12,20–27] focus on online matching problems with more specific constraints.

In [12], the authors discuss online assignment with the distance constraint that a service provider can only match requesters within a certain distance. An extension of online assignment is introduced in [20], where service providers master various skills and



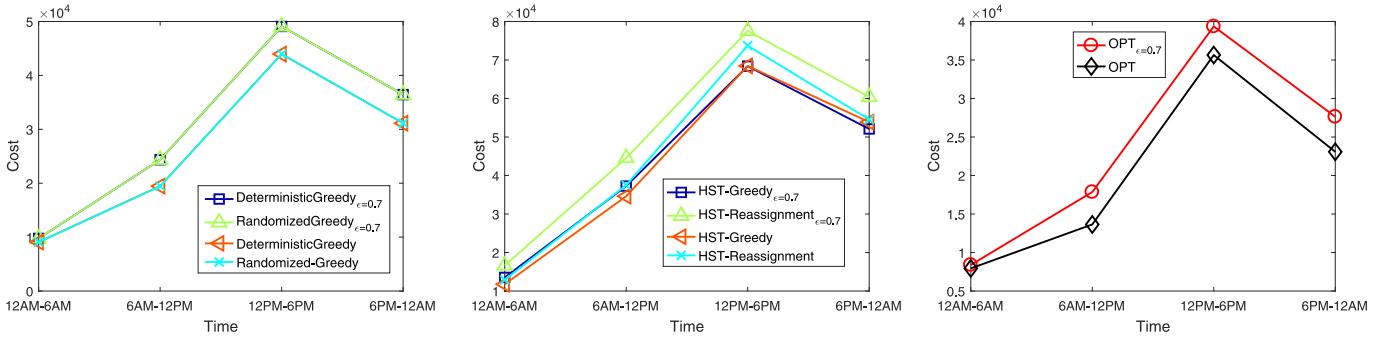
(a) Original Metric Space (b) HST Metric Space (c) OPT

Fig. 3. Performance of algorithms with privacy budget $\epsilon = 0.1$.



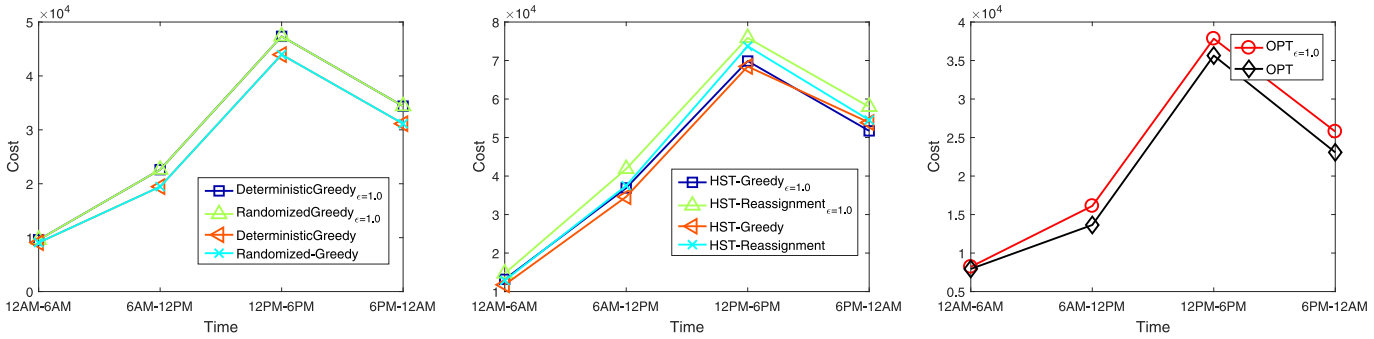
(a) Original Metric Space (b) HST Metric Space (c) OPT

Fig. 4. Performance of algorithms with privacy budget $\epsilon = 0.4$.



(a) Original Metric Space (b) HST Metric Space (c) OPT

Fig. 5. Performance of algorithms with privacy budget $\epsilon = 0.7$.



(a) Original Metric Space (b) HST Metric Space (c) OPT

Fig. 6. Performance of algorithms with privacy budget $\epsilon = 1.0$.

the platform assigned them to different types of tasks. In [21,22], the crowdworkers are restricted by the capacity constraint, which allows a worker to serve limited times. With only the real-time constraint, our work focuses on the basic online assignment problem, which is more fundamental and general. In terms of algorithms analyzing, it has been proved in [23] that the competitive ratio of online algorithms, in random model, is at least $\mathcal{O}(H_n)$, where H_n is the nth *Harmonic number*. Our work can apply all these algorithms in privacy protection framework and the corresponding competitive ratio can also be calculated.

Location privacy. Spatiotemporal data applications such as spatiotemporal crowdsourcing [11,12], route planning [28–30] and spatial keyword search [31–33], all need to protect location privacy. Privacy-preserving techniques can be used in these applications. For example, many spatial crowdsourcing applications require individuals' personal information [34–36]. To protect location privacy, existing research has proposed many techniques, such as cloaking [37], perturbation (by adding noise) [12,38,39] and encryption [40,41]. We adopt perturbation methods like GEO-I in this paper.

Compared with GEO-I, cloaking [37] needs assumption of adversary's prior knowledge. However, GEO-I provides stronger privacy guarantee regardless of adversary's knowledge while cloaking is sensitive to the prior assumption. In some work based on encryption [40], the approaches often lack efficiency, for it usually takes time in both encryption and decryption. In contrast, perturbation is easy to implement in practice without spending too much time.

Local difference privacy (LDP). Although location privacy has the similar guarantee to LDP [42], their objectives are different. LDP focuses on reconstructing the statistical information from individuals' reported data such as mean value or variance, while our work based on GEO-I directly uses individuals' data rather than statistics. Since in online assignment, statistics are not so important, GEO-I is more suitable to location privacy protection.

7. Conclusion

This paper proposes a privacy-preserving framework for online task assignment on ridesharing platforms and proves that the loss of cost caused by protecting privacy is limited. The competitive ratio of existing algorithms for OMBM problem is proved to be of the same magnitude with the original one. Moreover, we conduct experiments on a real dataset and verify that the total cost is at most **3.64** times larger than the original algorithm. We theoretically and experimentally prove that existing algorithms for OMBM problem still work under privacy protection, which enable third-party platforms successfully perform online task assignments while the privacy of individuals is preserved.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

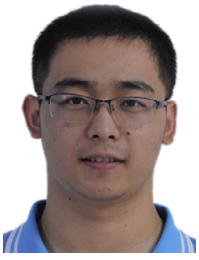
We are grateful to anonymous reviewers for their constructive comments. This work is partially supported by the National Key Research and Development Program of China under Grant No. 2018AAA0101100, National Science Foundation of China (NSFC) under Grant No. 61822201, U1811463 and 71531001, Science and Technology Major Project of Beijing under Grant No. Z171100005117001 and Didi Gaia Collaborative Research Funds for Young Scholars.

References

[1] Gigwalk, http://www.gigwalk.com,
[2] Gmission, http://gmissionhkust.com,
[3] Openstreetmap, https://www.openstreetmap.org,
[4] Waze, http://www.waze.com,
[5] Uber, https://www.uber.com,
[6] Y. Tong, Z. Zhou, Dynamic task assignment in spatial crowdsourcing, SIGSPATIAL Spec. 10 (2) (2018) 18–25.
[7] Y. Tong, L. Wang, Z. Zhou, B. Ding, L. Chen, J. Ye, K. Xu, Flexible online task assignment in real-time spatial data, PVLDB 10 (11) (2017) 1334–1345.
[8] Y. Tong, Y. Zeng, Z. Zhou, L. Chen, J. Ye, K. Xu, A unified approach to route planning for shared mobility, PVLDB 11 (11) (2018) 1633–1646.
[9] Y. Xu, Y. Tong, Y. Shi, Q. Tao, K. Xu, W. Li, An efficient insertion operator in dynamic ridesharing services, in: 35th IEEE International Conference on Data Engineering, ICDE 2019, Macau SAR, China, April 8–11, 2019, 2019, pp. 1022–1033.
[10] M.E. Andrés, N.E. Bordenabe, K. Chatzikokolakis, C. Palamidessi, Geo-indistinguishability: differential privacy for location-based systems, in: 2013 ACM SIGSAC Conference on Computer and Communications Security, CCS'13, Berlin, Germany, November 4–8, 2013, 2013, pp. 901–914.
[11] Y. Tong, J. She, B. Ding, L. Chen, T. Wo, K. Xu, Online minimum matching in real-time spatial data: experiments and analysis, PVLDB 9 (12) (2016) 1053–1064.
[12] H. To, C. Shahabi, L. Xiong, Privacy-preserving online task assignment in spatial crowdsourcing with untrusted server, in: 34th IEEE International Conference on Data Engineering, ICDE 2018, Paris, France, April 16–19, 2018, 2018, pp. 833–844.
[13] Shenzhou private cars, http://zhuanche.zuche.com,
[14] J. She, Y. Tong, L. Chen, C.C. Cao, Conflict-aware event-participant arrangement and its variant for online setting, IEEE Trans. Knowl. Data Eng. 28 (9) (2016) 2281–2295.
[15] Y. Tong, J. She, R. Meng, Bottleneck-aware arrangement over event-based social networks: the max-min approach, World Wide Web 19 (6) (2016) 1151–1177.
[16] J. Fakcharoenphol, S. Rao, K. Talwar, A tight bound on approximating arbitrary metrics by tree metrics, in: Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9–11, 2003, San Diego, CA, USA, 2003, pp. 448–455.
[17] Y. Tong, L. Chen, C. Shahabi, Spatial crowdsourcing: challenges, techniques, and applications, PVLDB 10 (12) (2017) 1988–1991.
[18] Y. Tong, L. Chen, Z. Zhou, H.V. Jagadish, L. Shou, W. Lv, SLADE: a smart large-scale task decomposer in crowdsourcing, IEEE Trans. Knowl. Data Eng. 30 (8) (2018) 1588–1601.
[19] Y. Tong, L. Wang, Z. Zhou, L. Chen, B. Du, J. Ye, Dynamic pricing in spatial crowdsourcing: A matching-based approach, in: Proceedings of the 2018 International Conference on Management of Data, SIGMOD Conference 2018, Houston, TX, USA, June 10–15, 2018, 2018, pp. 773–788.
[20] H. To, C. Shahabi, L. Kazemi, A server-assigned spatial crowdsourcing framework, ACM Trans. Spat. Algorithms Syst. 1 (1) (2015) 2:1–2:28.
[21] U. Leong Hou, M.L. Yiu, K. Mouratidis, N. Mamoulis, Capacity constrained assignment in spatial databases, in: Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2008, Vancouver, BC, Canada, June 10–12, 2008, 2008, pp. 15–28.
[22] R.C.-W. Wong, Y. Tao, A.W.-C. Fu, X. Xiao, On efficient spatial matching, in: Proceedings of the 33rd International Conference on Very Large Data Bases, University of Vienna, Austria, September 23–27, 2007, 2007, pp. 579–590.
[23] S. Raghvendra, A robust and optimal online algorithm for minimum metric bipartite matching, in: Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2016, September 7–9, 2016, Paris, France, 2016, pp. 18:1–18:16.
[24] R. Reiffenhäuser, An optimal truthful mechanism for the online weighted bipartite matching problem, in: Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6–9, 2019, 2019, pp. 1982–1993.
[25] T. Song, Y. Tong, L. Wang, J. She, B. Yao, L. Chen, K. Xu, Trichromatic online matching in real-time spatial crowdsourcing, in: 33rd IEEE International Conference on Data Engineering, ICDE 2017, San Diego, CA, USA, April 19–22, 2017, 2017, pp. 1009–1020.
[26] Y. Zeng, Y. Tong, L. Chen, Z. Zhou, Latency-oriented task completion via spatial crowdsourcing, in: 34th IEEE International Conference on Data Engineering, ICDE 2018, Paris, France, April 16–19, 2018, 2018, pp. 317–328.
[27] Y. Wang, Y. Tong, C. Long, P. Xu, K. Xu, W. Lv, Adaptive dynamic bipartite graph matching: areinforcement learning approach, in: 35th IEEE International Conference on Data Engineering, ICDE 2019, Macau SAR, China, April 8–11, 2019, 2019, pp. 1478–1489.
[28] J.M. Dibbelt, Engineering Algorithms for Route Planning in Multimodal Transportation Networks, Karlsruhe Institute of Technology, 2016 Ph.D. thesis.
[29] S. Wang, W. Lin, Y. Yang, X. Xiao, S. Zhou, Efficient route planning on public transportation networks: a labelling approach, in: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, Melbourne, Victoria, Australia, May 31, - June 4, 2015, 2015, pp. 967–982.
[30] X. Zhang, H. Duan, An improved constrained differential evolution algorithm for unmanned aerial vehicle global route planning, Appl. Soft Comput. 26 (2015) 270–284.
[31] C. Zhang, Y. Zhang, W. Zhang, X. Lin, Inverted linear quadtree: efficient top k spatial keyword search, IEEE Trans. Knowl. Data Eng. 28 (7) (2016) 1706–1721.

[32] K. Zheng, B. Zheng, J. Xu, G. Liu, A. Liu, Z. Li, Popularity-aware spatial keyword search on activity trajectories, World Wide Web 20 (4) (2017) 749–773.
[33] K. Zheng, H. Su, B. Zheng, S. Shang, J. Xu, J. Liu, X. Zhou, Interactive top-k spatial keyword queries, in: 31st IEEE International Conference on Data Engineering, ICDE 2015, Seoul, South Korea, April 13–17, 2015, 2015, pp. 423–434.
[34] Y. Tong, Y. Chen, Z. Zhou, L. Chen, J. Wang, Q. Yang, J. Ye, W. Lv, The simpler the better: a unified approach to predicting original taxi demands based on large-scale online platforms, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13, - 17, 2017, 2017, pp. 1653–1662.
[35] D. Gao, Y. Tong, J. She, T. Song, L. Chen, K. Xu, Top-k team recommendation in spatial crowdsourcing, in: Web-Age Information Management – 17th International Conference, WAIM 2016, Nanchang, China, June 3–5, 2016, Proceedings, Part I, 2016, pp. 191–204.
[36] D. Gao, Y. Tong, J. She, T. Song, L. Chen, K. Xu, Top-k team recommendation and its variants in spatial crowdsourcing, Data Sci. Eng. 2 (2) (2017) 136–150.
[37] L. Pournajaf, L. Xiong, V.S. Sunderam, S. Goryczka, Spatial task assignment for crowd sensing with cloaked locations, in: IEEE 15th International Conference on Mobile Data Management, MDM 2014, Brisbane, Australia, July 14–18, 2014 –Volume 1, 2014, pp. 73–82.
[38] L. Wang, D. Yang, X. Han, T. Wang, D. Zhang, X. Ma, Location privacy-preserving task allocation for mobile crowdsensing with differential geo-obfuscation, in: Proceedings of the 26th International Conference on World Wide Web, WWW 2017, Perth, Australia, April 3–7, 2017, 2017, pp. 627–636.
[39] X. Jin, R. Zhang, Y. Chen, T. Li, Y. Zhang, Dpsense: differentially private crowdsourced spectrum sensing, in: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24–28, 2016, 2016, pp. 296–307.
[40] B. Liu, L. Chen, X. Zhu, Y. Zhang, C. Zhang, W. Qiu, Protecting location privacy in spatial crowdsourcing using encrypted data, in: Proceedings of the 20th International Conference on Extending Database Technology, EDBT 2017, Venice, Italy, March 21–24, 2017., 2017, pp. 478–481.
[41] A. Pham, I. Dacosta, B. Jacot-Guillarmod, K. Huguenin, T. Hajar, F. Tramèr, V.D. Gligor, J.-P. Hubaux, Privateride: a privacy-enhanced ride-hailing service, PoPETs 2017 (2) (2017) 38–56.
[42] P. Kairouz, S. Oh, P. Viswanath, Extremal mechanisms for local differential privacy, Journal of Machine Learning Research 17 (2016) 17:1–17:51.

Yi Xu is currently working toward the Ph.D. degree in the School of Computer Science and Engineering, Beihang University. His research interests include crowd intelligence, ridesharing, crowdsourcing, spatio-temporal data management and data mining.



Shuyue Wei is currently an undergraduate student in SAE at Beihang University. His current research interests include spatio-temporal data processing, differential privacy and federated learning.

Yansheng Wang received the B.E degree from Beihang University, Beijing, China in 2017. He is currently working toward the Ph.D. degree in the School of Computer Science and Engineering, Beihang University. His research interests include spatial crowdsourcing, federated learning and data privacy.