

# Stable Task Assignment with Range Partition under Differential Privacy

Leilei Du<sup>1</sup>, Peng Cheng<sup>1(⊠)</sup>, Lei Chen<sup>2</sup>, Wangze Ni<sup>2</sup>, Jing Zhao<sup>3,4</sup>, and Xuemin Lin<sup>5</sup>

<sup>1</sup> East China Normal University, Shanghai, China leileidu@stu.ecnu.edu.cn, pcheng@sei.ecnu.edu.cn <sup>2</sup> Hong Kong University of Science and Technology, Hong Kong, China leichen@cse.ust.hk, wniab@cse.ust.hk <sup>3</sup> State Key Laboratory of Precision Blasting, Jianghan University, Wuhan, China jingzhao@jhun.edu.cn

School of Artificial Intelligence, Jianghan University, Wuhan, China Shanghai Jiaotong University, Shanghai, China

Abstract. With the development of cloud computing, spatial crowd-sourcing (SC) has become a significant concern in data processing, including food delivery and online car-hailing. However, privacy leakage presents a challenge for requesters who need to share their task information with the server. Differential privacy (DP) is a robust privacy protection paradigm that allows the release of useful information while safeguarding requesters' privacy. However, task assignment under DP often results in ineffective utility. In this paper, we propose a stable task assignment scheme that enables requesters to apply for workers and achieve effective stable matching while preserving the privacy of the requests (tasks). Specifically, we introduce an approach called ECM that achieves stable matching while protecting the preference of requesters. We demonstrate the efficiency and effectiveness of our ECM on synthetic and real datasets.

## 1 Introduction

With the increasing popularity of cloud computing, spatial crowdsourcing has emerged as a computing paradigm for solving spatial tasks that involve human participation. To access content services, individuals are required to provide certain personal information to the server. This enables the server to efficiently execute task computing based on the provided information.

However, the location information of users may be sensitive and can compromise their privacy. To address this issue, differential privacy [4] can be used to protect individual data while still ensuring accurate statistical results for the entire data set.

Most schemes [13,14] disturb task positions under different privacy levels and then send the noisy data to the server. The server has to match tasks and workers using the disturbed data, which increases the error in the matching result. In

<sup>©</sup> The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024 M. Onizuka et al. (Eds.): DASFAA 2024, LNCS 14853, pp. 243–253, 2024. https://doi.org/10.1007/978-981-97-5562-2\_16

this paper, we propose a new model where tasks compete for workers based on their knowledge of workers' information. In this model, the location of tasks is not made public to others. The tasks compete for workers while using differential privacy counters.

In this paper, we study the problem of spatial crowdsourcing, specifically the Private Stable Task Assignment Problem (PSTAP). In this context, tasks can be assigned to workers based on their proximity without revealing the tasks' locations. We assume that the workers' information is publicly available, the server is honest but curious, and the tasks are rational. To address this problem, we propose a private stable matching scheme called Enhanced Concentric Mechanism (ECM). This scheme divides the plane into multiple areas centered around each worker and transforms distances into preferences. It leverages the binary mechanism [2] and extends the Private Gale-Shapley Mechanism [15] to achieve a stable matching between tasks and workers. ECM ensures the privacy of workers' preferences under joint differential privacy. The contributions of this paper are as follows:

- (1) We formally define the Private Stable Task Assignment Problem (PSTAP) in Section 3.
- (2) We propose the Enhanced Concentric Mechanism (ECM) in Section 4 to solve PSTAP.
- (3) We evaluate the efficiency and effectiveness of ECM on synthetic and real datasets in Section 5.

### 2 Related Work

We introduce the related work on stable allocation and private data stream publication under differential privacy.

Stable Allocation. Stable allocation can be seen as a variation of the stable marriage matching problem, which can be solved using the Gale-Shapley algorithm [6]. However, the Gale-Shapley algorithm can only handle matching scenarios where both sides have strict preference tables. And it does not consider the privacy of individual agents.

Golle [7] proposes Matching Authorities to perform private matching, which provides privacy and correctness as long as the majority of Matching Authorities are honest. Keller and Scholl [9] and Doerner et al. [3] use RAM-based secure computation to implement stable matching. However, all of the above methods are time-consuming during the matching progress. Hsu et al. [8] have declared that private matching and allocation problems cannot be solved under differential privacy. They relax the privacy constraint as joint differential privacy (JDP) and design a private billboard to achieve JDP.

Private Data Stream Publication under Differential Privacy. Differential privacy [4] is an efficient approach to protect location privacy. To et.al [12] first decompose the spatial using Private Spatiotemporal Decomposition (PSD) [10] and then perform task assignment with obscure task location geocasted to those near regions with workers with high probability.

Dwork et al. [5] focus on private data stream publication and propose a binary tree technique for finite streams. Chan et al. [2] formalize the differentially private continual counter and design it to achieve  $\epsilon$ -differential privacy with  $\mathcal{O}(\frac{1}{\epsilon} \cdot (\log t)^{1.5} \cdot \log \frac{1}{\delta})$  error with probability  $1 - \delta$ . Zhang et al. [15] propose the Joint Differentially Private Gale-Shapley Mechanism for Location Privacy Protection. They use the differentially private continual counter to protect preferences. However, it does not consider the distance between different entities.

### 3 Problem Definition

We define a spatial task as  $t_i$  and its location as  $loc_i$ . We also define a spatial worker as  $w_j$  and its location and capacity as  $loc_j$  and  $c_j$  respectively. The notations are summarized in Table 1. We assume that workers are honest and public. The server is honest-but-curious, which means it will execute required algorithms correctly but may attempt to obtain information from data and requests. The tasks are rational, as they seek to maximize their benefits.

**Definition 1** ( $(\alpha, \beta)$ -stable matching). Given a set P with size |P| and a set Q with total capacity |Q.c| satisfying  $|P| \leq |Q.c|$ . A matching  $M = (P_m, Q_m)$  is a  $(\alpha, \beta)$ -stable matching, if there exists a stable matching M' such that with probability  $(1 - \beta)$ ,

$$\sum_{i \in M.P_m \cup M'.P_m, (i,j) \in M, (i,j') \in M'} |rank(i,j) - rank(i,j')| \le \alpha$$

where rank(i, j) is the rank of j in i's preference list.

**Definition 2.** (Private Stable Task Assignment Problem, PSTAP) Given a set of spatial tasks  $T = \{t_1, t_2, ..., t_m\}$ , and a set of spatial workers  $W = \{w_1, w_2, ..., w_n\}$ , the Private Stable Task Assignment Problem aims to find a  $(\alpha, \beta)$ -stable matching [11] M that ensures the tasks' location privacy while satisfies the following conditions:

$$\begin{aligned} & \min & \sum_{t_{i} \in T} \sum_{w_{j} \in W} x_{i,j} \cdot d_{i,j} \\ & s.t. & x_{i,j} \in \{0,1\} \\ & \sum_{t_{i} \in T} x_{i,j} \leq c_{j}, & \forall j = 1,2,...,n \\ & \sum_{w_{i} \in W} x_{i,j} \leq 1, & \forall i = 1,2,...,m \end{aligned}$$

### 4 Enhanced Concentric Mechanism

### 4.1 The Overview of ECM

The main idea of ECM is using the Private Gale-Shapley algorithm [15] to obtain a stable matching between spatial tasks and spatial workers. This is done

Table 1. Notations

t the ittle	
$t_i$ the $i$ -th	task
$w_j$ the $j$ -th	worker
T the task	set
$\overline{W}$ the wor	ker set
$d_{i,j}$ the real	distance from $t_i$ to $w_j$
(0)	nt-element in $B$ recording the applying of the $k$ -th serving area of $w_j$
$\overline{PR_{i,W}}$ the dist	ance set from $t_i$ to all workers
$AP_i$ the area	set of $t_i$ in all workers' serving range
$RP_i^{(k)}$ the rand	lom area set of $t_i$ in all workers' serving range at $k$ -th round
$b_i$ the count	nt vector held by $t_i$
$c_j$ the capa	acity of $w_j$
$\mu_i$ the inde	$\mathbf{x}$ of worker proposed by $t_i$ in the current round
$a_{j,k}$ the $k$ -th	serving area of $w_j$
$o_i(j)$ the dist	ance rank of $w_j$ in $t_i$ 's preference list
$o_i^{-1}(k)$ the inde	$\mathbf{x}$ of the worker who is in the $k$ -th position in $t_i$ 's preference list
$r_j(i)$ the area	rank of $t_i$ in the serving range of $w_j$
$Sum(C_i)$ the nois	e sum of counters not ranking behind $t_i$ 's current applying area

under a differential private streaming counter [2] structure, with preference lists constructed based on distance. ECM can be divided into the following 4 steps.

Task Preference Construction. At first, Each task  $t_i$  calculates the distance  $d_{i,j}$  between itself and each worker  $w_j$ . Then  $t_i$  sorts these distances in ascending order to construct the preference rank list  $PR_{i,W} = \{d_{i,o_i^{-1}(1)},...,d_{i,o_i^{-1}(n)}\}$ . All of the  $PR_{i,W}$  (for  $i \in [m]$ ) together form the tasks' preference list  $PR_{T,W}$ .

Worker Preference Construction. Each worker  $w_j$  first selects a series of distance values  $PD_{j,A} = \{dr_{j,1}, dr_{j,2}, ..., dr_{j,U_j}\}$  in ascending order.  $w_j$  then publishes  $PD_{j,A}$  to S. Centered at  $loc_j$ ,  $PD_{j,A}$  divides the plane into  $U_j + 1$  serving range areas  $PR_{j,A} = \{a_{j,1}, a_{j,2}, ..., a_{j,U_j+1}\}$ . The collection of all  $PR_{j,A}$  (for  $j \in [n]$ ) forms the workers' preference list  $PR_{W,A}$ . An example of  $w_j$  with U = 4 is illustrated in Figure 1(a).

**Area Positioning.** After obtaining the ranges of each worker,  $t_i$  looks up the range  $a_{j,k}$  for  $w_j$  that satisfies  $dr_{j,k-1} < d_{i,j} \le dr_{j,k}$  (recorded as  $r_j(i) = k$ ).  $t_i$  keeps track of all workers' area positions in a list  $AP_i = \{a_{j,r_j(i)}\}$  for each  $j \in [n]$ . We provide an instance in Figure 1(b).  $t_1$  falls within the serving range of  $w_1, w_2$ , and  $w_3$ . It is in  $a_{1,3}$  for  $w_1, a_{2,2}$  for  $w_2$ , and  $a_{3,3}$  for  $w_3$ . Therefore, the area positions for  $t_1$  are  $AP_1 = \{a_{1,3}, a_{2,2}, a_{3,3}\}$ .

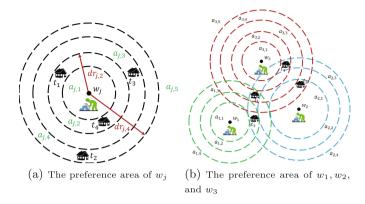


Fig. 1. Area division by workers.

Matching Construction. In this step, all tasks help the server S perform task assignment using the private Gale-Shapley algorithm. They hold a billboard [2,8] to record the application process of the tasks. Each task will iteratively apply to a worker in its worker rank list when the task is available. Once the application process is complete, S assigns the tasks to workers according to the billboard.

Next, we present the construction of our Enhanced Concentric Mechanism (ECM) algorithm which is shown in Algorithm 1.

```
Algorithm 1: ECM
   Input: Tasks: t_i's location loc_i for each i \in [m], Workers: w_i's locations loc_i,
            range preferences PR_{i,A} and capacity c_i
    Output: Matching results
 1 S initializes the billboard B by setting B(j,s) = BM(\epsilon/2mn, mn^2) for each
   j \in [n] and s \in [U_i];
 2 Each task t_i gets its rank of workers PR_{i,W};
 3 Set R_T as the tasks with candidate workers in PR_{T,W};
 4 while R_T is not empty do
       for t_i \in R_T do
 5
            Initializes current applying worker w_a as the next element in PR_{i,W};
 6
            Set b_i(a) = B(a, r_a(i));
 7
           Set b_i(j) = B(j, random(AP_i)) for w_i \in PR_{i,W} \setminus w_a;
 8
            Update b_i by executing Private Gale-Shapley algorithm [15];
 9
10
            t_i sends b_i to S;
            S updates B by b_i;
11
12 Each t_i sends its \mu(i) to S;
13 S gets the matching M by \mu(i) for i \in [m];
14 return B;
```

We provide an example in Figure 1(b) and Figure 2. Suppose there are 4 tasks:  $t_1$ ,  $t_2$ ,  $t_3$ , and 3 workers:  $w_1$ ,  $w_2$ ,  $w_3$ . At the beginning,  $w_1$ ,  $w_2$ , and  $w_3$  publish their serving distances, locations and capacities (The first table in Figure 2). The serving distances of the workers divide the plane into multiple areas (as shown in the second table in Figure 2) which form the preference list of all workers. After that, all tasks construct their preference tables for workers. S holds a billboard B. In B, each element is a counter and initialized as 0.

									S	erver S	S												
Worker	r ID	Location		Serving distance rank			Capacity			$PR_{W,A}$		1	Area 1		Area	Area 2 A		A	rea 4	Aı	Area 5		
$w_1$		$loc(w_1)$		10m, 20m, 30m					1		$w_1$			a <sub>1,1</sub>		$a_{1,2}$		a <sub>1,3</sub>		a <sub>1,4</sub>			
$w_2$		$loc(w_2)$		20m, 30m, 40			n		1		$w_2$			$a_{2,1}$		a <sub>2,2</sub>		a <sub>2,3</sub>		a <sub>2,4</sub>			
$w_3$		$loc(w_3)$		10m, 20m, 30m		, 40m		2			$w_3$			a <sub>3,1</sub>		a <sub>3,2</sub>		a <sub>3,3</sub>		$a_{3,4}$	$a_{3,5}$		
$t_1$	$w_1$	$w_3$	$w_2$		$t_2$	w	2	$w_3$	$w_1$		t <sub>3</sub>		w	1	$w_2$	$w_3$		t.	4	$w_3$	$w_2$	$w_1$	
$PR_{1,W}$	$d_{1,1}$	$d_{1,3}$	$d_{1,2}$		$PR_{2,W}$	$d_2$	,2	$d_{2,3}$	$d_{2,1}$		$PR_3$	3,W	$d_3$	$d_{3,1}$ $d_{3,2}$		$d_{3,3}$	1	$PR_{4,W}$		$d_{4,3}$	$d_{4,2}$	$d_{4,1}$	
$AP_1$	a <sub>1,3</sub>	a <sub>3,3</sub>	a <sub>2,3</sub>		$AP_2$	$a_2$	,1	a <sub>3,4</sub>	a <sub>1,4</sub>		AI	3	a <sub>1,</sub>	,2	a <sub>2,3</sub>	a <sub>3,5</sub>		$AP_4$		a <sub>3,2</sub>	a <sub>2,3</sub>	a <sub>1,4</sub>	
$RP_{1}^{(1)}$	a <sub>1,2</sub>	a <sub>3,1</sub>	a <sub>2,2</sub>		$RP_{2}^{(1)}$	$a_2$	,2	a <sub>3,2</sub>	a <sub>1,1</sub>		RP	(1)	a <sub>1,</sub>	3	a <sub>2,2</sub>	a <sub>3,2</sub>		RP,	(1) i	a <sub>3,5</sub>	$a_{2,1}$	a <sub>1,2</sub>	
$RP_{1}^{(2)}$	a <sub>1,1</sub>	a <sub>3,2</sub>	a <sub>2,3</sub>		$RP_{2}^{(2)}$	$a_2$	,3	a <sub>3,1</sub>	a <sub>1,2</sub>		RP	(2)	$a_1$	1	$a_{2,1}$	a <sub>3,4</sub>		$RP_{i}$	(2)	a <sub>3,4</sub>	a <sub>2,2</sub>	a <sub>1,1</sub>	
$RP_1^{(3)}  a_{1,3}  a_{3,1}  a_{2,1}$																							
B (Round 1)		Area 1 Area 2		Area 2	Area 3	A	rea 4	l A	rea 5	[	B (Ro		nd :	d 2) Area		1 A	rea 2	2 Area 3		Area 4 Area		rea 5	
$w_1$		0		0.7 0.8		0					$w_1$			0 0.		0.7	0.8		0				
$w_2$		0.6		0 0		0					$w_2$			0.6		5	0	0		0			
$w_3$		0 (		0.9	0 0		0	0			$w_3$			0 (		0.9	1.1		0 0		0		
Round	1 1		$a_{1,3}$ $b_1(1)$ $0.8$	$a_{3,1}$ $b_1(3)$ $0$	$a_{2,2}$ $b_1(2)$ $0$	<b>t</b> <sub>2</sub>	L <sub>2</sub>	$a_{2,1}$ $b_2(2)$ $0.6$	) b <sub>2</sub> (3	$a_{1,1}$ $b_2(1)$ $0$		$t_3$	L <sub>3</sub>	$b_3$	100		1 <sub>3,2</sub> 3(3)	$t_4$	L <sub>4</sub>	a <sub>3,2</sub> b <sub>4</sub> (3) 0.9	$a_{2,1}$ $b_4(2)$ $0.6$	$a_{1,2}$ $b_4(1)$ $0.7$	
	Ļ	$ \mu_1 $		1			$\mu_2$		2				$\mu_3$	L		1		L	$\mu_4$		3		
Round	2	-	$a_{1,1}$ $b_1(1)$ 0	a <sub>3,3</sub> b <sub>1</sub> (3) 1.1	a <sub>2,3</sub> b <sub>1</sub> (2) 0	$t_2$	$L_2$ $b_2$ $\mu_2$	a <sub>2,2</sub> b <sub>2</sub> (2		a <sub>1,1</sub> ) b <sub>2</sub> (1) 0		$t_3$	$\frac{L_3}{b_3}$	$b_3$	1,3 a (1) b <sub>3</sub> ).8	(2) b	ι <sub>3,2</sub> <sub>3</sub> (3) 0.9	$t_4$	$L_4$ $b_4$ $\mu_4$	b <sub>4</sub> (3)	a <sub>2,1</sub> b <sub>4</sub> (2) 0.6	a <sub>1,2</sub> b <sub>4</sub> (1) 0.7	
	L	μ <sub>1</sub>		1			$\mu_2$	100.20					$\mu_3$				85999		$\mu_4$	388888			

Fig. 2. Task assignment under ECM.

In the first round,  $t_1$ ,  $t_2$ , and  $t_3$  first set their counter vectors  $b_1$ ,  $b_2$ , and  $b_3$  as the billboard values in the related areas. For example,  $t_1$  want to apply to  $w_1$ , it sets  $b_1$  as B(1,3).  $t_1$  randomly chooses areas for  $w_3$  and  $w_2$  (suppose they are  $a_{3,1}$  and  $a_{2,2}$ ), and set  $b_1(3)$  as B(3,1),  $b_1(2)$  as B(2,2). Then,  $t_1$  feeds private 1 (e.g., 0.8) to  $b_1(1)$ , which means it sets B(1,3) = 0.8, and feeds private 0 to  $b_1(3)$  and  $b_1(2)$  (i.e., B(3,1) = 0.B(2,2) = 0). After that, it calculates the sum  $C_1 = B(1,1) + B(1,2) + B(1,3) = 0.8$ .  $t_1$  checks if  $C_1 \le c_1$  ( $c_1 = 1$ ). Since this condition is satisfied,  $t_1$  sets its assignment worker index  $\mu_1$  as 1. In the same way,  $t_2$ ,  $t_3$ , and  $t_4$  set  $\mu_2$ ,  $\mu_3$ , and  $\mu_4$  as 2, 1, and 3 respectively. In the second round,  $t_1$  calculates the sum  $C_1'' = B(1,1) + B(1,2) + B(1,3) = 1.5 > c_1$ . It then

applies to the next worker  $w_3$ .  $t_1$  feeds private 1 (e.g., 1.2) to  $b_1(3)$ , which means it sets B(3,3)=1.2, and feeds private 0 to  $b_1(1)$  and  $b_1(2)$  (i.e., B(1,1)=0, B(2,3)=0). After that, it calculates the sum  $C_1=B(3,1)+B(3,2)+B(3,3)=2 \le c_3$  ( $c_3=2$ ). Thus,  $t_1$  sets  $\mu_1$  as 3.  $t_2$ ,  $t_3$ , and  $t_4$  calculate their sums  $C_2 \le c_2$ ,  $C_3 \le c_1$ , and  $C_4 \le c_3$  respectively. Therefore, they feed private 0 to all elements in  $b_2$ ,  $b_3$ , and  $b_4$ . In the third round, there are no assignment changes, and thus the applying process terminates.

### 4.2 Performance Analysis

**Time cost.** In our ECM, each task applies to at most n workers, and there are m tasks. Therefore, there will be an application time complexity of O(mn). For each application, a task feeds n data, resulting in a computation complexity of  $O(mn^2)$ .

Accuracy. Let  $Z_A = \max_{j \in [n]} U_j + 1$  represent the maximum area size among all workers. Let  $G(a_{j,k}) = \sum_{i \in [m], AP_i \ni a_{j,k}} \mathbf{1}$  denote the task size in area  $a_{j,k}$ . Let  $G_T = \max_{j \in [n], a_{j,k} \in PR_{j,A}} G(a_{j,k})$  denote the maximal task size among all areas. Let  $F_T = \max_{j \in [n], a_{j,k} \in PR_{j,A}} \left( \frac{n-1}{U_j} \cdot (m - G(a_{j,k})) + G(a_{j,k}) \right)$  be the maximal average applying times among all areas. We present Theorem 1 below.

**Theorem 1.** ECM satisfies  $(\alpha, \beta)$ -useful and  $(\alpha, \beta)$ -stable matching where  $\alpha = \frac{4\sqrt{2} \cdot Z_A \cdot mn}{\epsilon} \cdot \ln\left(\frac{2}{1-\frac{Z_A}{\sqrt{1-\beta}}}\right) \cdot (\sqrt{\log(G_T \cdot n)})^5$ . Furthermore, ECM satisfies  $(\alpha_A, \beta)$ -useful where  $\exp(\alpha_A) = \frac{4\sqrt{2} \cdot Z_A \cdot mn}{\epsilon} \cdot \ln\left(\frac{2}{1-\frac{Z_A}{\sqrt{1-\beta}}}\right) \cdot (\sqrt{\log F_T})^5$  with  $\alpha_A \leq \alpha$ .

**Privacy.** Next, we provide the privacy and security guarantees of our ECM.

**Theorem 2.** ECM is  $\epsilon$ -joint differentially private.

# 5 Experiment

### 5.1 Data Sets

Real Data Set. We conducted our experiments on Didi Chuxing [1] in Chengdu, China (referred to as *chengdu*). We utilized the node file *chengdu.node*, which contains the coordinates of 214,440 vertices in Chengdu's road network. We randomly divide these vertices into two parts to represent tasks and workers. Synthetic Data Sets. We generate a synthetic data sets: uniform data nodes (denoted as *uniform*). *uniform* consists of 100,000 points extracted from a 2-dimensional uniform distribution. We randomly divide these points into two data sets to represent tasks and workers.

### 5.2 Experimental Setup

We randomly extract points from data sets to create data scales of different sizes. Let  $S_T$  and  $S_W$  represent the sets of tasks and workers, respectively. We define the value  $p_t = \frac{|S_T|}{|S_W|}$  as the task ratio. We costruct the Basic Concentric Mechanism (BCM) by replace choosing randomly area as choosing real area (Set the step in Line 7 and 8 as  $b_i(j) = B(j, r_j(i))$  for all  $w_j \in PR_{i,W}$ ). We construct a non-private scheme of BCM by using non-private counters in the billboard. We refer to this scheme as Non-private Basic Concentric Mechanism (NP-BCM). For the worker preferences in the Joint Differentially Private Gale-Sharley Mechanism [15], we use random preferences and denote it as Joint Differentially Private Gale-Shapley with Random Worker Preferences (JDP-GS-RWP). We then compare BCM and ECM with NP-BCM and JDP-GS-RWP.

We examine the impact of various factors such as data scale, task ratio  $(p_t)$ , worker capacity (c), ring radius (the distance between neighboring serving circles), serving circle number, and total privacy budget  $(\epsilon)$  on the Average Running Time  $(Time_{AVG})$ , Average Receiving Distance  $(RD_{AVG})$ , and Average Task Rank Value  $(TRV_{AVG})$ . The parameter settings are presented in Table 2.

Parameters	Value range	Default value
data scale $( D )$	200, 400, 600, 800, 1000	0 200
task ratio $(p_t)$	4, 5, 6, 7, 8	6
worker capacity $(c)$	1, 2, 3, 4, 5	4
ring radius	600, 700, 800, 900, 1000	01000
serving circle number	60, 80, 100, 120, 140	100
total privacy budget (e	100, 200, 300, 400, 500	300

Table 2. Settings for different factors

### 5.3 Experiment Results

- 1) Average Time Cost: Figure 3(a) and Figure 3(b) display the average time cost for data scales ranging from 200 to 1000. We observe that the running time increases quadratically with the data scale. Figure 3(c) and Figure 3(d) illustrate the average time cost for task ratios ranging from 4 to 8. We observe that as the task ratio increases, the running time decreases. This is because with more tasks, the smaller workers can complete the tasks more quickly. Additionally, the running time of the non-private mechanism is significantly smaller than that of the private mechanisms.
- 2) Average Receiving Distance: Figure 4 illustrates the impact of various factors on average distance. We observe that the receiving distance decreases as the task ratio increases. This is because workers can receive more tasks with a lower

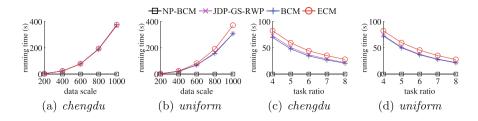


Fig. 3. The impact of data scale and task ratio on the time cost.

distance. Additionally, the receiving distance increases as the worker capacity increases. This is because a larger worker capacity slows down the competition in smaller ranges and reduces the overall distance. Furthermore, BCM and ECM consistently outperform the JDP-GS-RWP in these parameters for both the *chengdu* and *uniform* data sets.

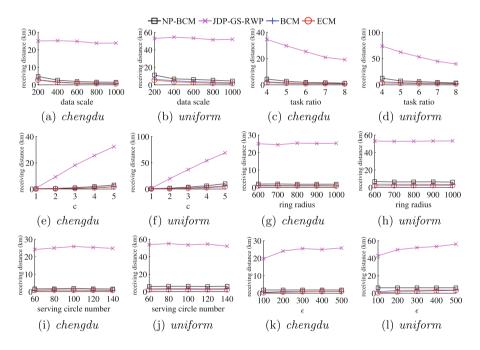


Fig. 4. The impact on average receiving distance.

3) Average Task Rank Value: Figure 5 shows the average task rank value for different ring radius and circle numbers. We can observe that both BCM and ECM consistently outperform JDP-GS-RWP in the *chengdu* and *uniform* data sets. From Figure 5(a) and Figure 5(b), we can see that there is a slight impact on the average task rank value as the ring radius varies. The average task

rank value generally increases and then decreases as the ring radius increases in BCM, ECM, and NP-BCM. Similarly, from Figure 5(c) and Figure 5(d), we can see that the serving circle number also has a slight influence on the task rank value. The average task rank value generally increases and then decreases as the serving circle number increases in BCM, ECM, and NP-BCM.

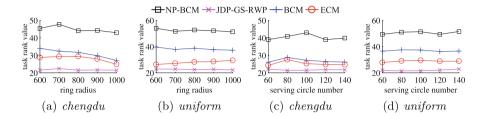


Fig. 5. The impact of data scale and task ratio on the average task rank value.

# 6 Conclusion

In this paper, we investigate the problem of stable allocation with location protection. To safeguard the location information of tasks, we transform it into preference information and present the ECM. Our experiments demonstrate that ECM exhibit improved efficiency and accuracy compared to the state-of-the-art work (Joint Differentially Private Gale-Shapley Mechanisms) with random preferences for workers and do not significantly decrease compared to the non-privacy scheme.

Acknowledgements. Peng Cheng's work is supported by the National Natural Science Foundation of China under Grant No. 62102149. Lei Chen's work is partially supported by National Key Research and Development Program of China Grant No. 2023YFF0725100, National Science Foundation of China (NSFC) under Grant No. U22B2060, the Hong Kong RGC GRF Project 16213620, RIF Project R6020-19, AOE Project AoE/E-603/18, Theme-based project TRS T41-603/20R, CRF Project C2004-21G, Hong Kong ITC ITF grants MHX/078/21 and PRP/004/22FX, Microsoft Research Asia Collaborative Research Grant and HKUST-Webank joint research lab grants. Jing Zhao's work is supported by the Knowledge Innovation Program of Wuhan-Shugung Project (No.2023020201010140). Xuemin Lin is supported by NSFC U2241211, NSFC U20B2046, and 23H020101910.

# References

- 1. Didi Chuxing. http://www.didichuxing.com/.
- T. H. Chan, E. Shi, and D. Song. Private and continual release of statistics. ACM Trans. Inf. Syst. Secur., 2011.

- J. Doerner, D. Evans, and A. Shelat. Secure stable matching at scale. In SIGSAC, 2016.
- 4. C. Dwork. Differential privacy. In ICALP, 2006.
- C. Dwork, M. Naor, T. Pitassi, and G. N. Rothblum. Differential privacy under continual observation. In STOC, 2010.
- D. Gale and L. S. Shapley. College admissions and the stability of marriage. Am. Math. Mon., 2013.
- 7. P. Golle. A private stable matching algorithm. In FC, 2006.
- J. Hsu, Z. Huang, A. Roth, T. Roughgarden, and Z. S. Wu. Private matchings and allocations. In STOC, 2014.
- M. Keller and P. Scholl. Efficient, oblivious data structures for MPC. In ASI-ACRYPT, 2014.
- J. S. Kim, Y. D. Chung, and J. W. Kim. Differentially private and skew-aware spatial decompositions for mobile crowdsensing. Sensors, 2018.
- D. G. McVitie and L. B. Wilson. The stable marriage problem. Commun. ACM, 1971.
- H. To, G. Ghinita, and C. Shahabi. A framework for protecting worker location privacy in spatial crowdsourcing. Proc. VLDB Endow., 2014.
- H. To, C. Shahabi, and L. Xiong. Privacy-preserving online task assignment in spatial crowdsourcing with untrusted server. In ICDE, 2018.
- 14. L. Wang, D. Zhang, D. Yang, B. Y. Lim, and X. Ma. Differential location privacy for sparse mobile crowdsensing. In *ICDM*, 2016.
- Y. Zhang, Y. Mao, and S. Zhong. Joint differentially private gale-shapley mechanisms for location privacy protection in mobile traffic offloading systems. *IEEE J. Sel. Areas Commun.*, 2016.