Privacy-Preserving Online Task Assignment in Spatial Crowdsourcing: A Graph-based Approach

Hengzhi Wang, En Wang*, Yongjian Yang, Jie Wu, Falko Dressler

Speaker: Hengzhi Wang

IEEE INFOCOM 2022

Outlines

1. Introduction
2. Challenges and contributions
3. Problem formulation
4. Approaches
5. Experiments

1. Introduction
2. Challenges and contributions
3. Problem formulation
4. Approaches
5. Experiments

1. Introduction

Spatial Crowdsourcing

1) Background

A new problem-solving paradigm to explore the power of crowd with location-aware tasks.

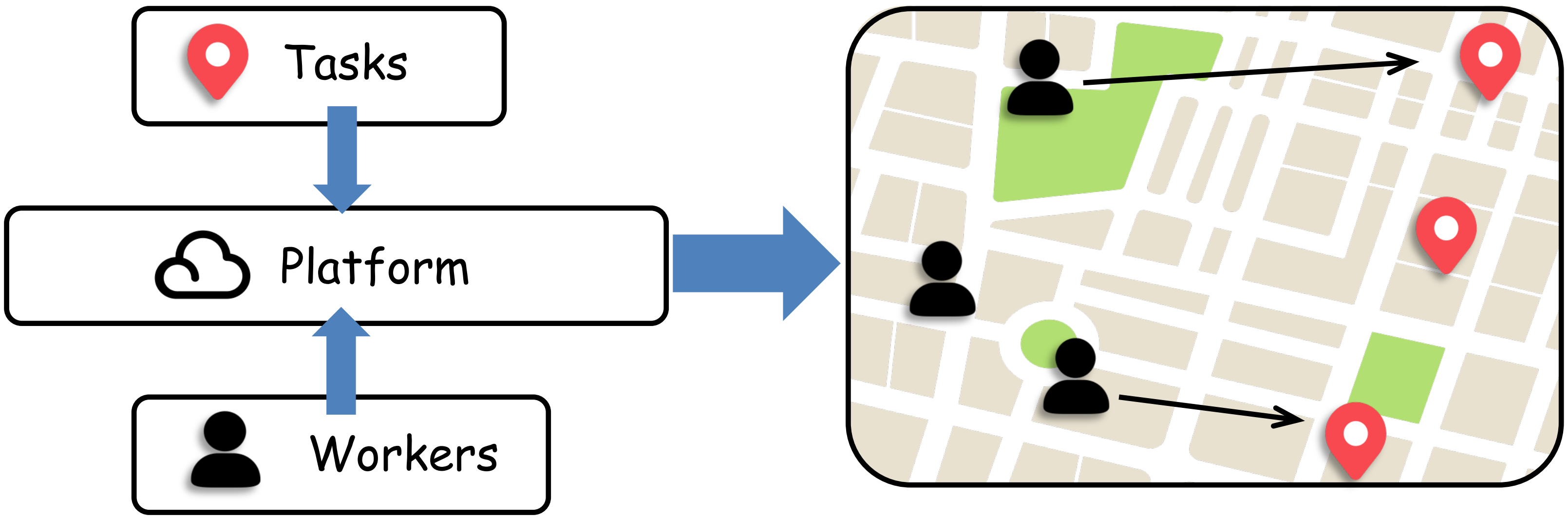
2) Applications

Uber Gigwalk TaskRabbit

1. Introduction

Spatial Crowdsourcing

3) Components



4) Issues

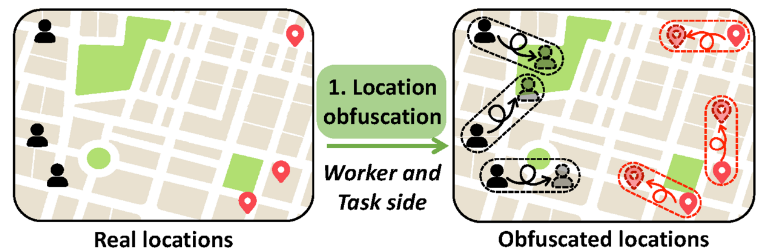
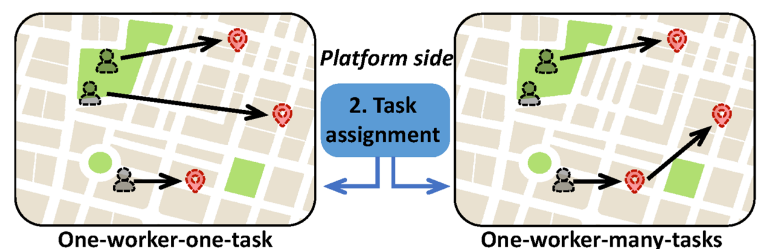
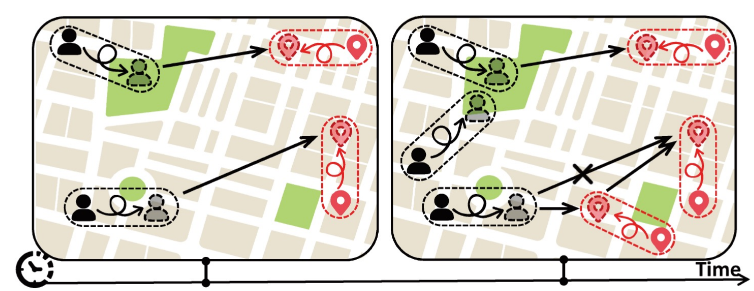
Task assignment

Incentive mechanism

Privacy protection

1. Introduction

Task assignment:

Prior works	Our work	
1) No-privacy	1) Privacy	
2) One-worker-one-task	2) One-worker-many-tasks	
3) Offline	3) Online	

Outlines

1. Introduction
2. Challenges and contributions
3. Problem formulation
4. Approaches
5. Experiments

2. Challenges and contributions

- Challenges
  - Balance the tradeoff between privacy protection and utility
  - Execute the one-worker-many-tasks assignment
  - Deal with the online task assignment

- Existing works cannot deal with these challenges

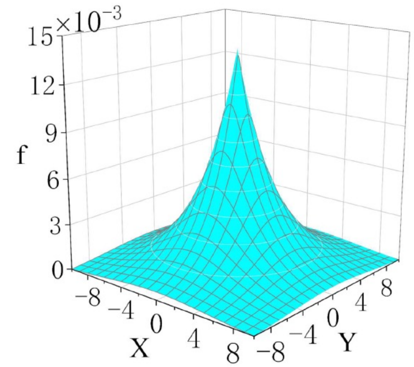
	1) Privacy	2) One-worker-many-tasks	3) Online
Work [1]	✓	✗	✗
Works [2,3]	✓	✗	✓
Works [4,5]	✗	✓	✗
Our work	✓	✓	✓

[1] Z. Wang, J. Li, J. Hu, J. Ren, Z. Li, and Y. Li, "Towards privacy- preserving incentive for mobile crowdsensing under an untrusted plat- form," in *Proc. IEEE INFOCOM*, 2019, pp. 2053–2061.
[2] H. To, C. Shahabi, and L. Xiong, "Privacy-preserving online task assignment in spatial crowdsourcing with untrusted server," in Proc. IEEE ICDE, 2018, pp. 833–844.
[3] Q. Tao, Y. Tong, Z. Zhou, Y. Shi, L. Chen, and K. Xu, "Differentially private online task assignment in spatial crowdsourcing: A tree-based approach," in Proc. IEEE ICDE, 2020.
[4] D. Deng, C. Shahabi, and U. Demiryurek, "Maximizing the number of worker's self-selected tasks in spatial crowdsourcing," in Proc. ACM SIGSPATIAL GIS, 2013, pp. 314–323.
[5] D. Deng, C. Shahabi, and L. Zhu, "Task matching and scheduling for multiple workers in spatial crowdsourcing," in Proc. ACM SIGSPATIAL GIS, 2015, pp. 21:1–21:10.

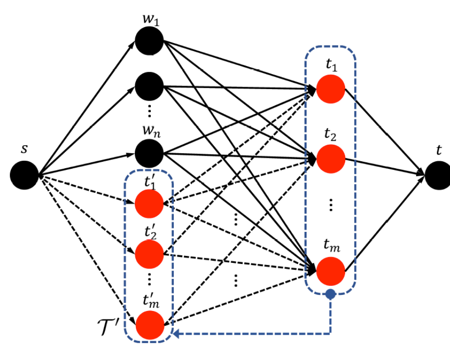
2. Challenges and contributions

- How to deal with challenges
  - tradeoff between privacy protection and utility
  - One-worker-many-tasks assignment
  - Online task assignment

Geo-Indistinguishability

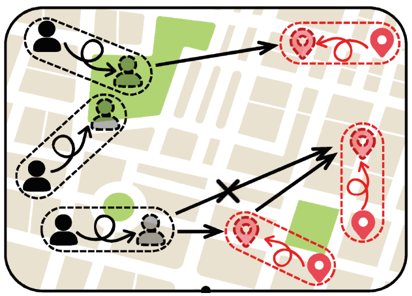


PLP



TOA

Bound of $\frac{\phi \cdot \delta}{d(O') - \sigma}$



PLP-TOA

Bound of $O(1/(\bar{d} \cdot \epsilon^2))$

2. Challenges and contributions

- Our work has the following contributions:
  - Propose a privacy mechanism to balanced the tradeoff between privacy and utility.
  - Solve the online one-worker-many-tasks assignment problem with the competitive ratio of $O(1/(\bar{d} \cdot \epsilon^2))$.
  - Evaluate the effectiveness of the proposed method using real-world datasets.

Outlines

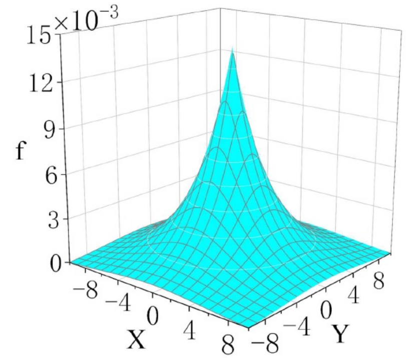
1. Introduction
2. Challenges and contributions
3. Problem formulation
4. Approaches
5. Experiments

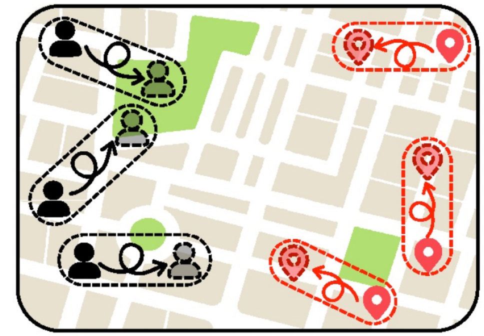
3. Problem formulation

- Privacy model

Geo-Indistinguishability [1]: A privacy mechanism M satisfies Geo-Indistinguishability iff

$$\mathcal{M}(l)(\mathcal{Z}) \le e^{\epsilon d(l,l')}\mathcal{M}(l')(\mathcal{Z})$$



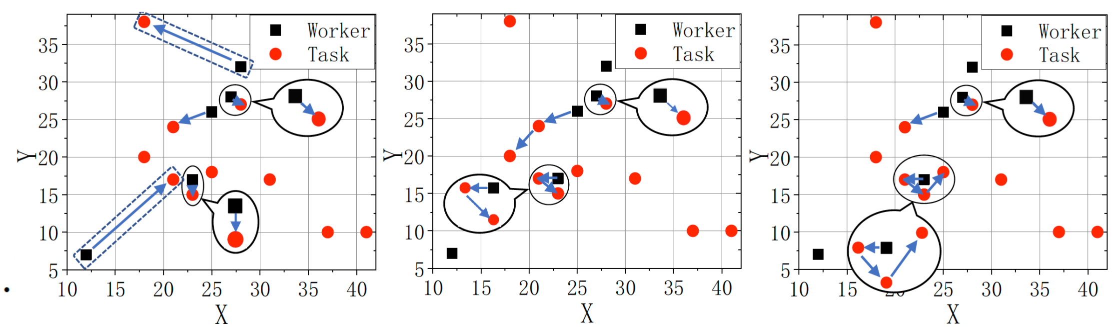


- Privacy-preserving Online Task Assignment (POTA) Problem

$$\text{Minimize} \quad \sum_{w_i \in \mathcal{W}} d_i$$

$$\text{Subject to} \quad \sum_{w_i \in \mathcal{W}} |\mathcal{T}_i| \ge \delta$$

$$|\mathcal{T}_i| \le w_i.\mathcal{C}, \ \forall w_i \in \mathcal{W}.$$



(a) Capacity=1 (b) Capacity=2 (c) Capacity=3

[1] M. E. Andr´ es, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, “Geo-indistinguishability: differential privacy for location-based systems,” in Proc. ACM CCS, 2013, pp. 901–914.

Outlines

1. Introduction
2. Challenges and contributions
3. Problem formulation
4. Approaches
5. Experiments

4. Approaches

- Planar Laplace distribution based Privacy mechanism (PLP)

Algorithm 1: PLP

Input: Privacy budget ϵ, real location $l = (l_x, l_y)$

Output: Obfuscated location $l^* = (l_x^*, l_y^*)$

Draw $p \in [0, 1]$ uniformly, $\rho = F^{-1}(p)$;

Draw $\theta \in [0, 2\pi]$ uniformly;

$l_x^* = l_x + \rho\cos(\theta)$, $l_y^* = l_y + \rho\sin(\theta)$;

return $l^* = (l_x^*, l_y^*)$

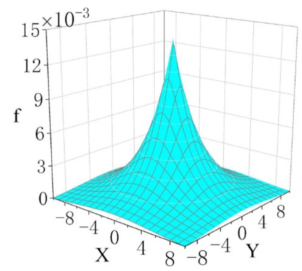


Fig. 3: The planar Laplace distribution centered at l_0.

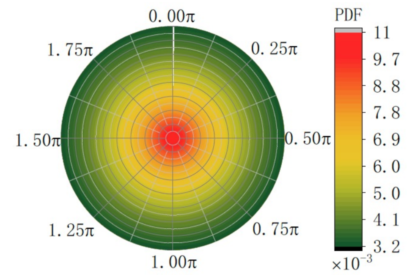


Fig. 4: The polar Laplace distribution with origin in l_0.

- Planar Laplace distribution

$$f(l, l^*, \epsilon) = \frac{\epsilon^2}{2\pi} e^{-\epsilon \cdot d(l, l^*)} \qquad F(\rho, \epsilon) = \int_0^{\rho} \int_0^{2\pi} \frac{\epsilon^2}{2\pi} e^{-\epsilon\rho} d\theta d\rho = 1 - (1 + \epsilon\rho) e^{-\epsilon\rho}$$

$$l^* = (l_x + \Delta x, l_y + \Delta y) = (l_x + \rho\cos(\theta), l_y + \rho\sin(\theta))$$

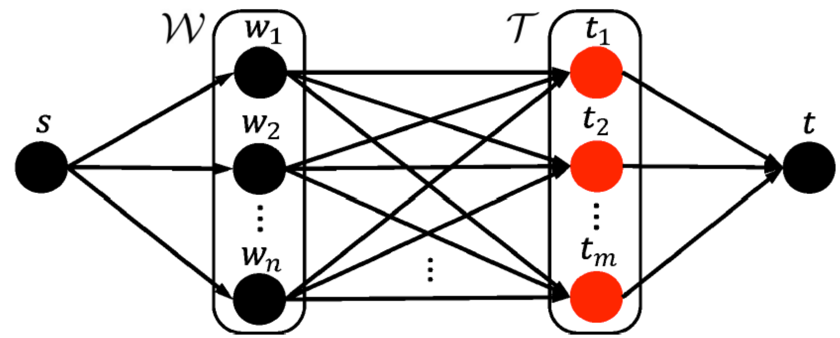
- Theorem 1. Geo-Indistinguishability

$$\mathcal{M}(l_1)(l^*)/\mathcal{M}(l_2)(l^*) = f(l_1, l^*, \epsilon)/f(l_2, l^*, \epsilon)$$
$$= e^{\epsilon \cdot (d(l_2, l^*) - d(l_1, l^*))} \leq e^{\epsilon \cdot d(l_1, d_2)}.$$

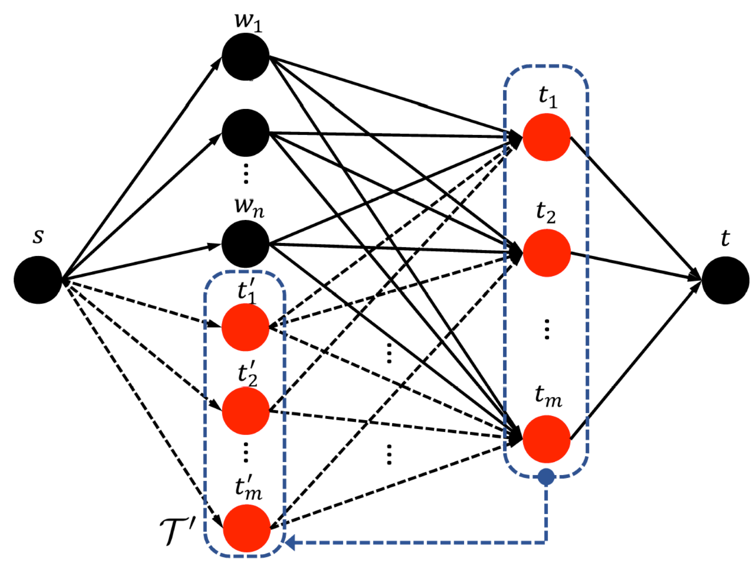
4. Approaches

- Threshold-based Online task Assignment mechanism (TOA)
  - Offline one-worker-many-tasks assignment

one-worker-one-task → Minimum Bipartite Matching problem (MBM) →



one-worker-many-tasks → Extended Minimum Bipartite Matching problem (EMBM) →



4. Approaches

- Threshold-based Online task Assignment mechanism (TOA)
  - Offline one-worker-many-tasks assignment

one-worker-one-task → Minimum Bipartite Matching problem (MBM)

one-worker-many-tasks → Extended Minimum Bipartite Matching problem (EMBM) →

Algorithm 2: Extened minimum-cost flow (EMCF)

Input: Workers $\mathcal{W}$, tasks $\mathcal{T}$, cardinality constraint δ
Output: The minimum-cost flow $\mathcal{F}$, total cost $\mathcal{D}$

Construct $G' = (V', A')$ as Eq. (10) based on $\mathcal{W}, \mathcal{T}$;
Initialize the flow $\mathcal{F} \leftarrow \emptyset$, total cost $\mathcal{D} \leftarrow 0$;
foreach $(w_i, t_j) \in A'$ **do**
  if $w_i.t_l < t_j.t_a$ *or* $w_i.t_a > t_j.t_l$ **then**
    Remove the arc (w_i, t_j) from A';
Find the minimum-cost augmenting path $\mathcal{P}(G', s, t)$;
while $\mathcal{P}(G', s, t)$ *exists and* $|\mathcal{F}| < \delta$ **do**
  $\mathcal{F} \leftarrow \mathcal{F} \cup \mathcal{P}(G', s, t)$;
  foreach *arc* $(v, v') \in \mathcal{P}(G', s, t)$ **do**
    if *arc* $(v', v) \notin A'$ **then**
      $A' \leftarrow A' \cup \{(v', v)\}$;
      $d(v', v) \leftarrow -d(v, v')$, $c(v', v) \leftarrow 0$;
    $c(v, v') \leftarrow c(v, v') - 1$, $c(v', v) \leftarrow c(v', v) + 1$;
    $\mathcal{D} \leftarrow \mathcal{D} + d(v, v')$;
    if $v' = t_j, \forall t_j \in \mathcal{T}$ **then**
      $d(s, t'_j) \leftarrow 0$;
  Find $\mathcal{P}(G', s, t)$ again based on the current G';
return $\mathcal{F}$, $\mathcal{D}$

4. Approaches

- Threshold-based Online task Assignment mechanism (TOA)
  - Online one-worker-many-tasks assignment

✓ TOA estimates a threshold

$$\kappa = min\{d(\mathcal{P})\}, \forall \mathcal{P} \in \mathcal{F}$$

✓ Theorem 2. TOA achieves a bound of $\frac{\phi \cdot \delta}{d(O') - \sigma}$

$$Pr[|d(\mathcal{O}) - E(d(\mathcal{O}'))| \leq \varepsilon] \geq 1 - \sigma^2/\varepsilon^2,$$

$$CR = E(TOA)/OPT = E(\delta \cdot \Delta)/d(\mathcal{O}) = \delta E(\Delta)/d(\mathcal{O})$$
$$\leq \phi \cdot \delta/d(\mathcal{O}) \leq \phi \cdot \delta/(1 - \sigma^2/\varepsilon^2) \cdot (d(\mathcal{O}') - \varepsilon). \quad (13)$$

✓ Theorem 3. PLP-TOA achieves a bound of $O(1/(\bar{d} \cdot \epsilon^2))$

Algorithm 3: Threshold online assignment (TOA)

```
Input: 𝒲, 𝒯, δ, T, φ
Output: ℱ
ℱ ← ∅, t ← 0 ;
while t ≤ T and |ℱ| < δ do
    𝒲_t, 𝒯_t ← the current workers and tasks in t;
    ℱ_t, 𝒟_t = EMCF(𝒲_t, 𝒯_t, δ);
    foreach 𝒫 ∈ ℱ_t do
        if d(𝒫) ≤ κ and the pre-path of 𝒫 is in ℱ then
            ℱ ← ℱ ∪ {𝒫};
    t ← t + 1;
return ℱ
```

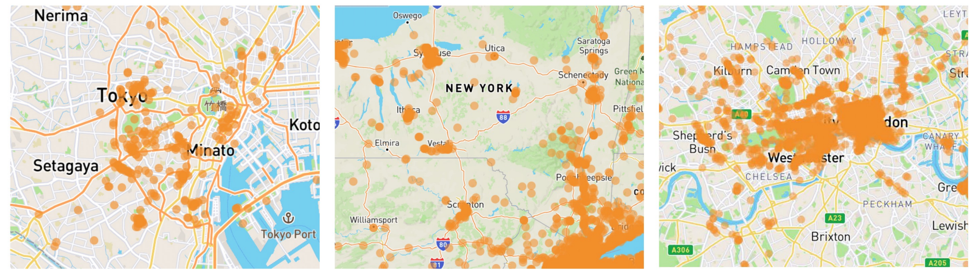
$$Pr[|d(\tilde{\mathcal{F}}) - d(\mathcal{F})| \geq \lambda] \geq \frac{6}{\lambda^2 \cdot \epsilon^2},$$

$$CR = d(\tilde{\mathcal{F}})/OPT \leq \frac{6(d(\mathcal{F}) + \lambda)}{d(\mathcal{O}) \cdot \lambda^2 \epsilon^2} \leq \frac{6(\kappa \cdot \delta + \lambda)}{\lambda^2 \epsilon^2 (d(\mathcal{O}') - \sigma)},$$

5. Experiments

5. Experiments

- Two real-world datasets, three cities
  - Gowalla: New York
  - Foursquare: Tokyo, London
- Baselines
  - PLP-TOA
  - TBF (ICDE'20 [1])
  - PLP-OA
  - OPT
  - PLP-Gre
- Settings



(a) Tokyo (b) New York (c) London

Parameters	Value
Time length T	100
Worker number $\|\mathcal{W}\|$	$10 \sim 90$
Task number $\|\mathcal{T}\|$	$50 \sim 250$
Privacy budget ϵ_1	$0.1 \sim 2.1$
Cardinality δ	$5 \sim 45$
Capacity $\mathcal{C}$	$1 \sim 10$

[1] Q. Tao, Y. Tong, Z. Zhou, Y. Shi, L. Chen, and K. Xu, "Differentially private online task assignment in spatial crowdsourcing: A tree-based approach," in Proc. IEEE ICDE, 2020.

- RQ1: Does our method deal with the online one-worker-many-tasks assignment?

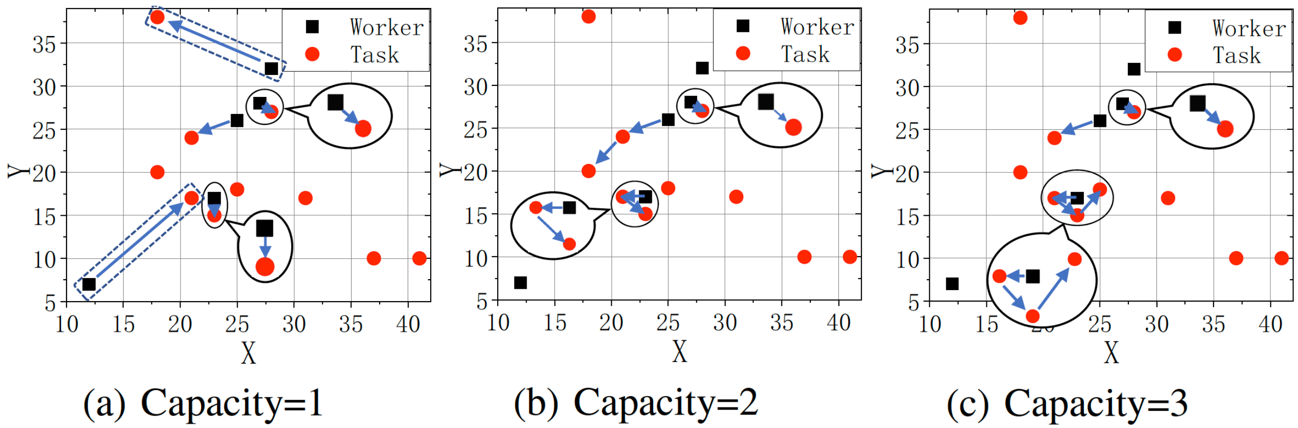


Fig. 14: Examples of the one-worker-many-tasks assignment.

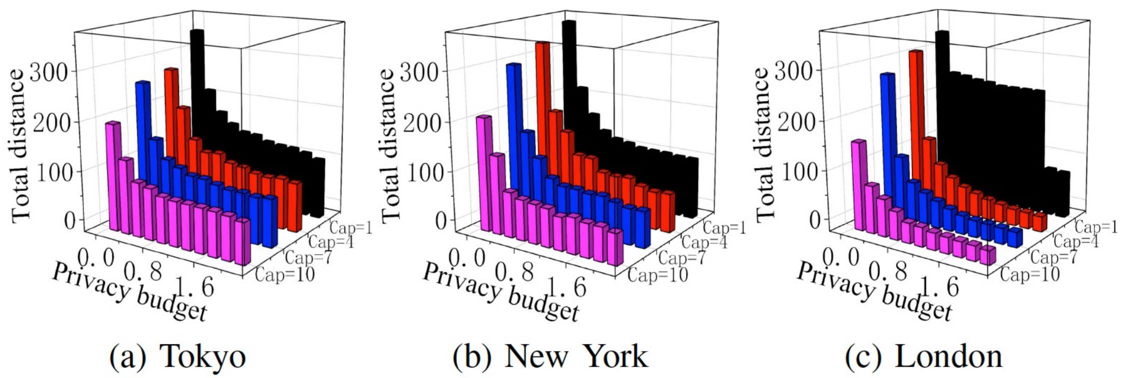
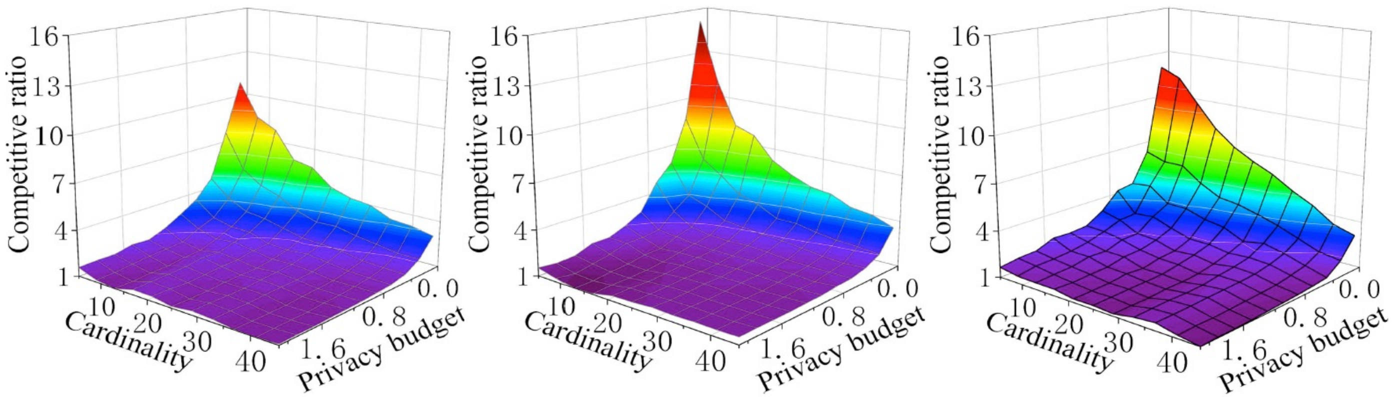


Fig. 13: Total distance vs. Capacity.

5. Experiments

- RQ2: Does our method agrees with the theoretical results?



(a) Tokyo (b) New York (c) London

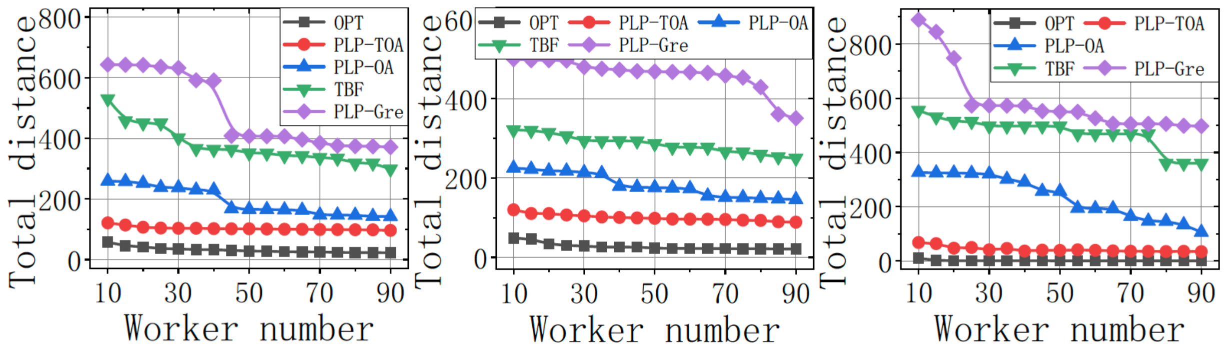
Fig. 12: Competitive ratio.

PLP-TOA achieves a bound of $O(1/(\bar{d} \cdot \epsilon^2))$

$$CR = d(\tilde{\mathcal{F}})/OPT \leq \frac{6(d(\mathcal{F}) + \lambda)}{d(\mathcal{O}) \cdot \lambda^2 \epsilon^2} \leq \frac{6(\kappa \cdot \delta + \lambda)}{\lambda^2 \epsilon^2 (d(\mathcal{O}') - \sigma)},$$

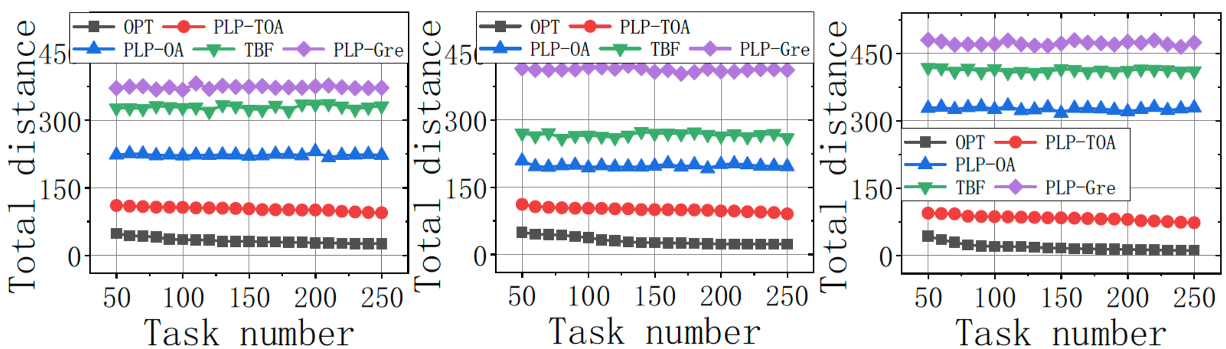
5. Experiments

❖ RQ3: How different parameters affect simulation results?



(a) Tokyo (b) New York (c) London

Fig. 8: Total distance vs. Worker number.

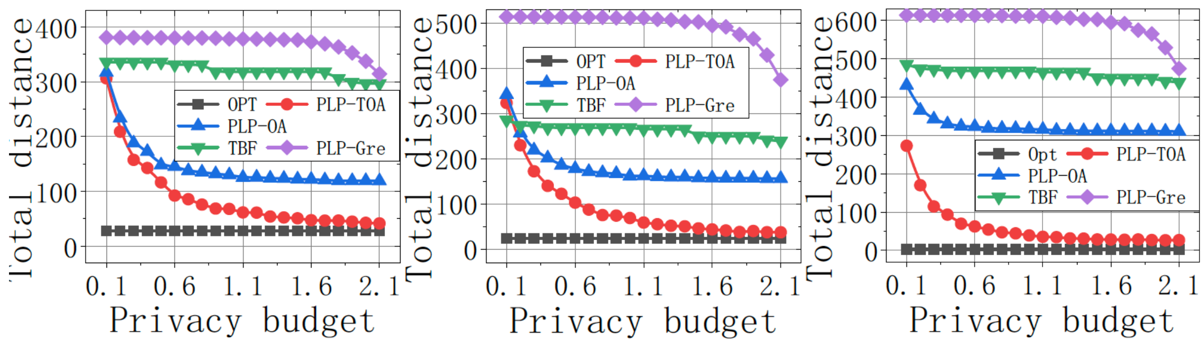


(a) Tokyo (b) New York (c) London

Fig. 9: Total distance vs. Task number.

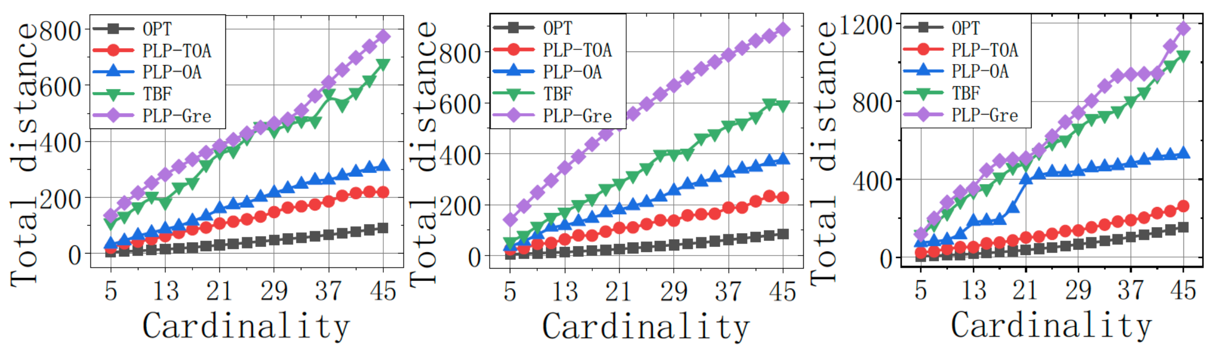
5. Experiments

- RQ3: How different parameters influence our mechanism?



(a) Tokyo (b) New York (c) London

Fig. 10: Total distance vs. Privacy budget.



(a) Tokyo (b) New York (c) London

Fig. 11: Total distance vs. Cardinality.

Thank you!

Q&A