认知智能全国重点实验室
STATE KEY LABORATORY OF COGNITIVE INTELLIGENCE

2024第一届认知智能与大数据技术研讨会

CIBD 2024

1 USTC

2 SUDA

Crowdsensing Data Trading for Unknown Market: Privacy, Stability, and Conflicts

He Sun[1], Mingjun Xiao[1], Yin Xu[1], Guoju Gao[2], Shu Zhang[1]

IEEE INFOCOM

IEEE TRANSACTIONS ON MOBILE COMPUTING

Key Question

How to design a Crowdsensing Data Trading Framework considering **privacy** and **stability** for **unknown market** in **centralized**[1] and **decentralized**[2] settings?

Introduction

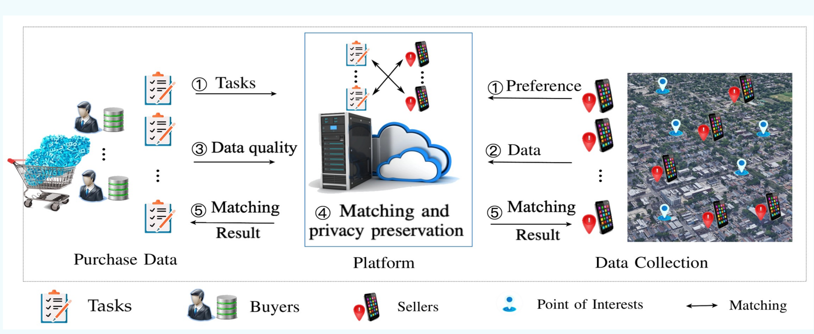
➢ Crowdsensing Data Trading (CDT)

A new data trading paradigm where the Mobile CrowdSensing (MCS) technique is adopted to provide data sources, e.g., Thingful, ThingSpeak.

➢ Concept of Matching Markets

- ✓ Both sides of the markets can' t just choose what you want even if you can afford it.
- ✓ One of them also have to be chosen.
- ✓ They choose each other according to the preferences of each other.

➢ Components of CDT systems



PS-CDT platform

Platform: As a broker, it provides credible data trading services for sellers and buyers.

Buyers: Propose and publish their data requirements to the platform to collect data.

Sellers: A crowd of mobile users to provide data collection service to buyers.

➢ Existing Problems

A. A few existing CDTs consider the stability of the Data Trading Market.

B. The Data Trading Market is unknown in practice, i.e., the preference sequences over sellers are unknown by buyers.

C. The private information of sellers needs to be preserved.

D. Decentralized CDT has more practical significance.

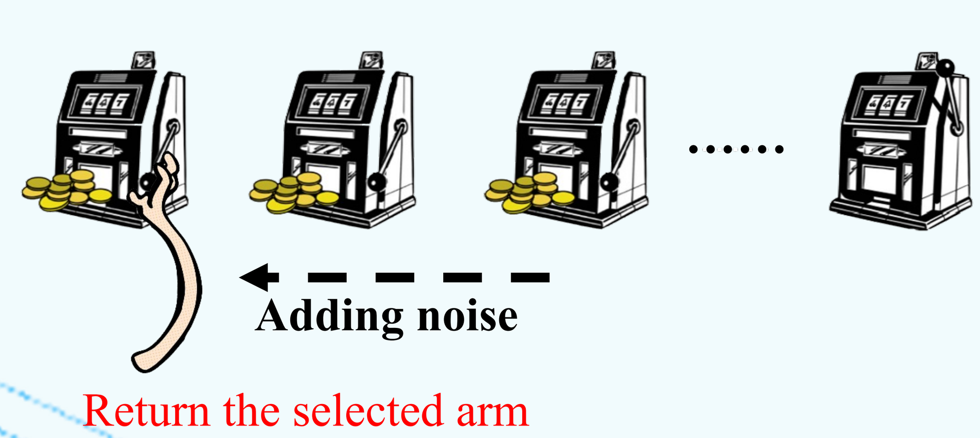
E. Two matching requests for the same seller would create a competitive matching conflict.

➢ Contributions

- ✓ To the best of our knowledge, this is the first CDT work that takes the unknown market, privacy preservation, and the stability of the data trading into consideration simultaneously in centralized and decentralized settings.
- ✓ We define a novel metric, i.e., δ-stability to measure the stability of the markets.
- ✓ We propose the DPS-CB and CDPS-CB mechanisms to solve the privacy, stability, and conflicts-avoiding problems.

System, Modeling, and Problem

➢ ϵ-Differentially private bandit model



Platform	Game players
Sellers	Arms
Select a seller	Pull an arm
Data quality	Reward
Protected Data quality	Perturbed reward

- ✓ $\mathbb{P}\{\Phi_i(q_i^{1:l-1}) \in \mathcal{X}\} \le e^{\epsilon} \cdot \mathbb{P}\{\Phi_i(q_i^{1:l-1'}) \in \mathcal{X}\}$ (1)
- ✓ where $\epsilon>0$ is a small constant that the policy provides, indicating the privacy-preserving level.
- ✓ A bandit policy Φ_i of play i is a sequence of arm-pulling decisions.
- ✓ $q_i^{1:l} = \{q_i^1, \dots, q_i^l\}$, $q_i^{1:l-1'}$ is its adjacent sequence.

➢ δ-Stable Matching Model

Definition of preference

Unknown preference sequences of the buyer

- ✓ Denoted by $\pi_k^{l'} = \{\dots, \pi_i^l, \dots\}$, $\pi_i^l = \{\dots, j, j', \dots\}$.
- ✓ $\pi_i^l(j)$ denotes the rank of seller j in π_i^l.
- ✓ $v_i = \{\dots, v_i^l(j), \dots\}$ denotes the value. (Unknown)

Preference sequence of the seller

- ✓ Denoted by $\pi_j = \{\dots, i, i', \dots\}$.
- ✓ $\pi_j(i)$ denotes the rank of task i in π_j.

Adding noise ➡ Matching is not truly stable

Definition of δ-stable: We say a market outcome M^l is δ-stable with a probability less equal than $1-\delta$ that a preference sequence is invalid, i.e., there exists two matching pairs $\langle i,j\rangle$ and $\langle i,j^*\rangle, \forall i \in T, \forall j, j^* \in S$, satisfies $\pi_i^l(j) \prec_i \pi_i^l(j^*)$, $\hat{\pi}_i^l(j) \prec_i \hat{\pi}_i^l(j^*)$ and $\hat{v}_i^l(j) - \hat{v}_i^l(j^*) > \xi_0'$, denoted by $\widehat{M}^k$. ξ_0' is a perturbed care bound and δ is a constant less than but close to 1. $\pi_i^l(j) \prec_i \pi_i^l(j^*)$: task i prefers seller j to j^* in l^{th} round.

➢ Problem formulation

Our goal is to make the optimal matching in each round according to the built perturbed preference sequences, i.e., to maximize the expected accumulative reward for each task, assuring the ϵ-differential privacy and δ-stable of market outcomes in each rounds.

$$Maximize: \sum_l q_i^l\left(m^l(i)\right)$$

$$Subject\ to: Eq.(1)\ holds$$

$$M^l\ is\ \delta - stable$$

DPS-CB and CDPS-CB mechanisms

Algorithm 1: DPS-CB mechanism

```
Input: the total rounds N, the preference sequences
       set {π̄_j | ∀j ∈ S} of sellers.
Output: {M^l | l = 1, 2, ....}
1  for l = 1, ..., N do
2    if l ≤ T then
3      m^l(l) ← j, ∀ j ∈ S;
4      Get q̂_l^l(j) as the corresponding reward
         according to Eqs. (6-8) while using ε as the
         privacy budget under the hybrid
         differentially private mechanism;
5    else if l = T + 1 then
6      Compute the DP-UCB indexes I_i^l(j), ∀i ∈ T,
         ∀j ∈ S according to Eq. (9);
7      Sort the sellers by the DP-UCB index to build
         the initial perturbed preference sequence π̂_i^l
         of each task over sellers;
8      Compute stable matching to get the market
         outcome M^l according to {π̄_j | ∀j ∈ S} and
         {π̂_i^l | ∀i ∈ T} using the Gale and Shapley
         algorithm;
9    else
10     Update I_i^l(j), ∀i ∈ T, ∀j ∈ S and
         {π̂_i^l | ∀i ∈ T} according to Eqs. (6-9).
11     Compute stable matching to get the market
         outcome M^l in the way of Step 8.
12   end
13 end
```

Algorithm 2: CDPS-CB mechanism

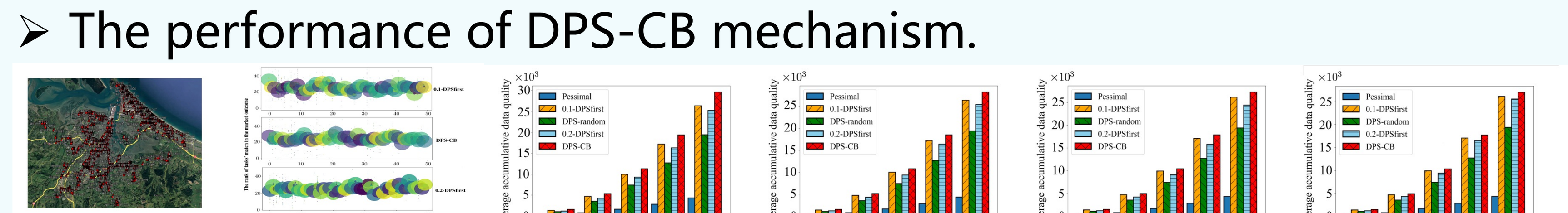
```
Input: the preference sequences set {π̄_j | ∀j ∈ S} of
       sellers, the Bernoulli mean p
Output: {M^l | l = 1, 2, ....}
Initialization:
I_i^1(j) = +∞, ∀i ∈ T, j ∈ S;
Find a matching that one-to-one maps from i to j
  randomly, ∀i ∈ T, j ∈ S;
for l = 0, ..., N do
  for i = 1, ..., T do
    Sample an random value B^l(i) from Ber(p);
    if B^l(i) = 0 then
      Update the set of feasible sellers according
        to Eqs (10);
      Task i selects the seller with the maximum
        DP-UCB index to match:
      m^l(i) = max{I_i^l(j) | ∀j ∈ F^l(i)};
    end
    else
      Task i matches the same seller as the last
        round:
      m^l(i) = m^{l-1}(i);
    end
    if i wins the conflicts then
      M^l ← ⟨i, m^l(i)⟩;
      Update I_i^l(m^l(i)) and {π̂_i^l | ∀i ∈ T};
    end
  end
end
```

➢ Theoretical Analysis

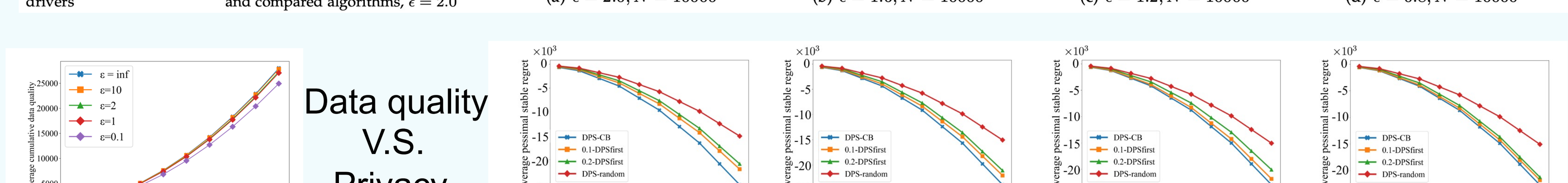
- ✓ The two mechanisms satisfies ϵ-differential privacy.
- ✓ The market outcome computed by twos mechanisms are δ-stable.
- ✓ The two mechanisms can achieve *sublinear* pessimal stable regret

Performance Analysis

➢ The performance of DPS-CB mechanism.



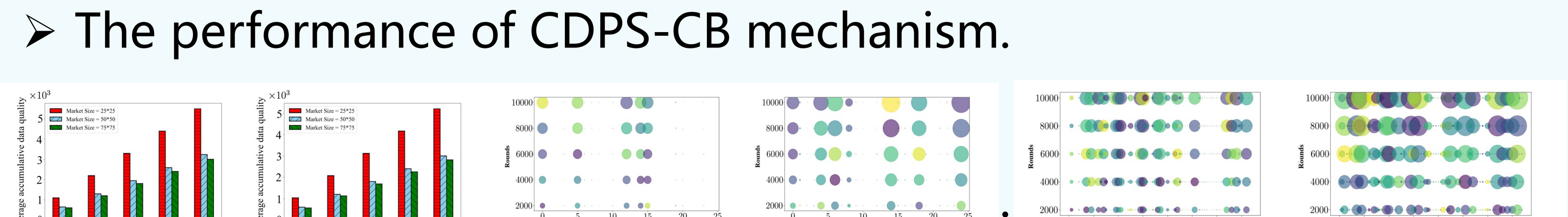
(a) The distribution of Uber drivers (b) The rank distribution of DPS-CB and compared algorithms, $\epsilon = 2.0$ (a) $\epsilon = 2.0, N = 10000$ (b) $\epsilon = 1.6, N = 10000$ (c) $\epsilon = 1.2, N = 10000$ (d) $\epsilon = 0.8, N = 10000$



Data quality V.S. Privacy

(a) $\epsilon = 2.0, N = 10000$ (b) $\epsilon = 1.6, N = 10000$ (c) $\epsilon = 1.2, N = 10000$ (d) $\epsilon = 0.8, N = 10000$

➢ The performance of CDPS-CB mechanism.



(a) $p = 0.2, \epsilon = 2.0, N = 10000$ (b) $p = 0.5, \epsilon = 2.0, N = 10000$ (a) $p = 0.2, \epsilon = 2.0, S = T = 25, N = 10000$ (b) $p = 0.5, \epsilon = 2.0, S = T = 25, N = 10000$ (c) $p = 0.2, \epsilon = 2.0, S = T = 75, N = 10000$ (d) $p = 0.5, \epsilon = 2.0, S = T = 75, N = 10000$

[1] **He Sun**, Mingjun Xiao, Yin Xu, Guoju Gao, Shu Zhang "Privacy-preserving Stable Crowdsensing Data Trading for Unknown Market", IEEE INFOCOM'23, May. 2023

[2] **He Sun**, Mingjun Xiao, Yin Xu, Guoju Gao, Shu Zhang. "Crowdsensing Data Trading for Unknown Market: Privacy, Stability, and Conflicts", IEEE Transactions on Mobile Computing, 2024.