Toward Distribution Estimation under Local Differential Privacy with Small Samples

Takao Murakami (AIST)
Hideitsu Hino (University of Tsukuba)
Jun Sakuma (University of Tsukuba)

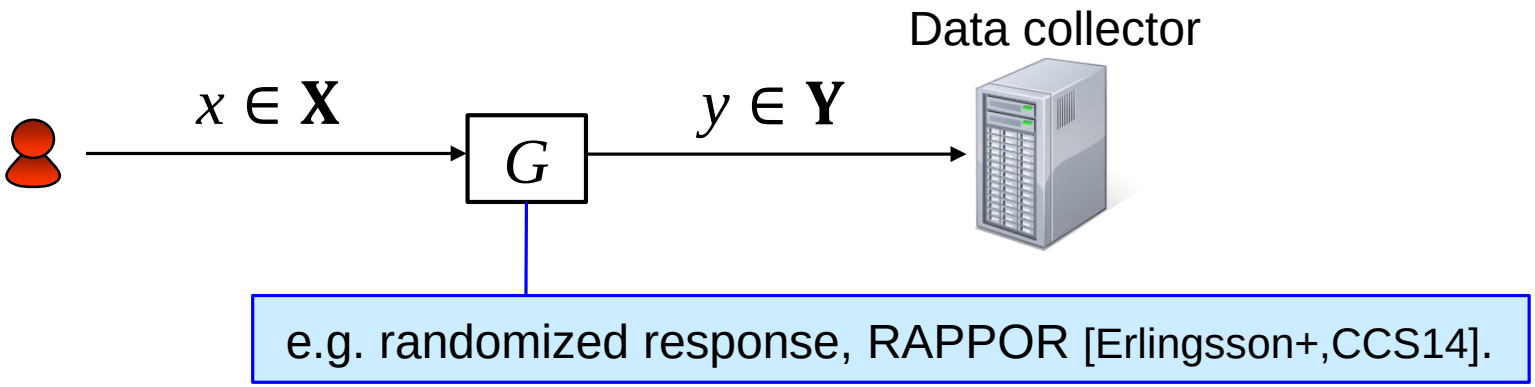
*AIST: National Institute of Advanced Industrial Science & Technology

Outline

- Local Differential Privacy (LDP) [Duchi+, FOCS13]
  - Variant of differential privacy [Dwork, ICALP06] in the “local” model.
  - User obfuscates her personal data (e.g., location, response in a survey) by herself (i.e., we do not assume a trusted third party).

Obfuscation mechanism G provides ε-LDP if for all $x, x' \in \mathbf{X}$ and all $y \in \mathbf{Y}$,

$$\Pr(y|x) \leq e^{\varepsilon} \Pr(y|x').$$

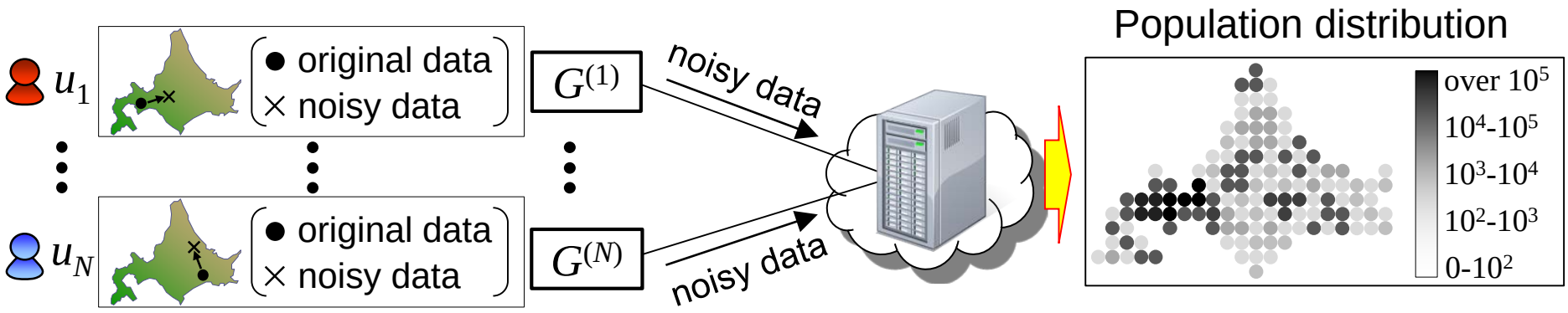


Strong Privacy

(1) Privacy is protected against attackers with any background knowledge.
(2) Original data are not leaked from DB (unlike the centralized DP).

Outline

- Distribution Estimation under LDP [Duchi+, FOCS13] [Erlingsson+,CCS14] [Kairouz+,ICML16] [Sei+,TIFS17] etc.
  - Data collector estimates a distribution of original data from “noisy” data.
  - We assume each user u_n obfuscates one sample using mechanism $G^{(n)}$.
  - Privacy budget of $G^{(n)}$ is $\varepsilon^{(n)}$ (like the personalized DP [Jorgensen+, ICDE15]).

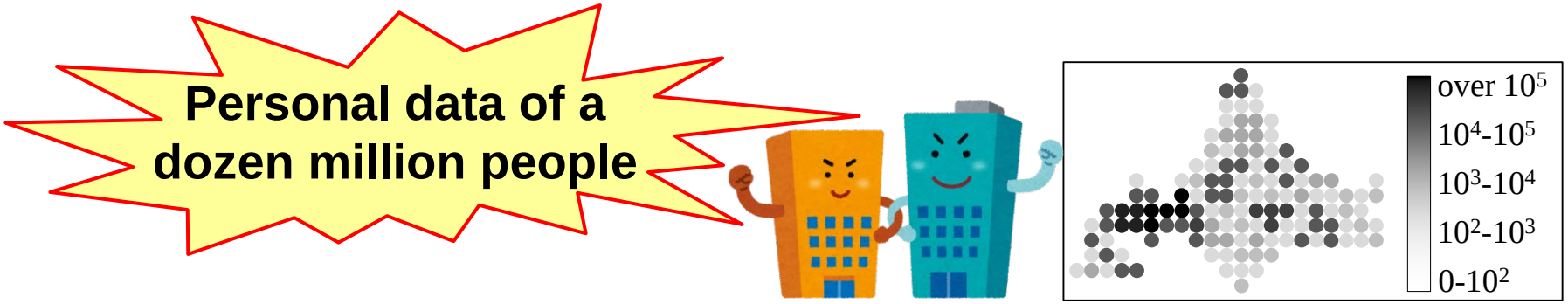


Strong Privacy

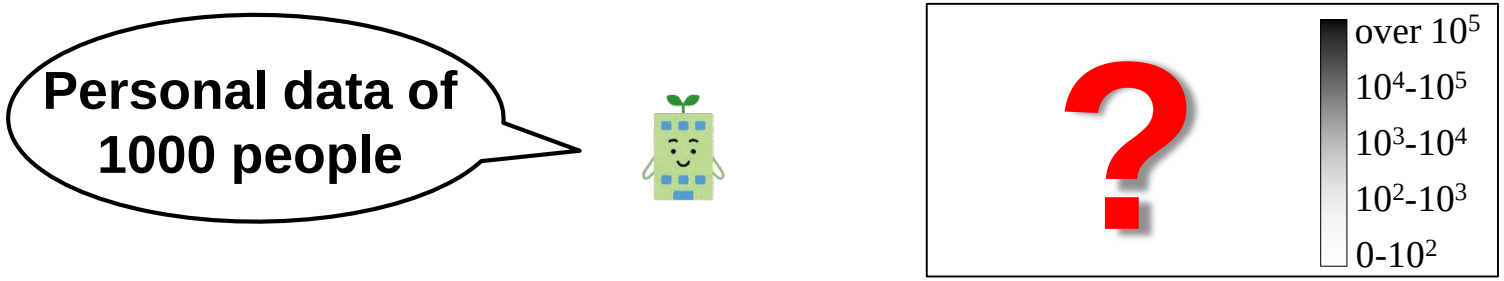
(1) Privacy is protected against attackers with any background knowledge.
(2) Original data are not leaked from DB (unlike the centralized DP).

Outline

- How about Data Utility?
  - When the sample size N is large, we can accurately estimate distribution.
  - E.g. Google collected a dozen million samples [Erlingsson+, CCS14].



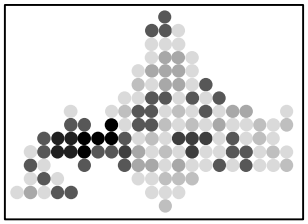
  - When N is small, distribution estimation is very challenging…



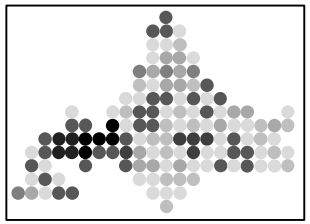
Q. How do we make LDP work for "small-scale" enterprises?

Outline

- Our Contributions
  - We propose a method to correct the estimation error of **the EM reconstruction method** based on **the theory of [Rilstone+, JE96]**.
  - We show, both theoretically & experimentally, that **the MSE (Mean Square Error) can be reduced when N is small (e.g., N=1000)**.

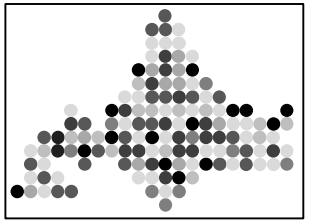


True distribution



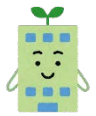
Proposal

Estimation error correction



EM reconstruction method
[Agrawal+, SIGMOD05]

Distribution estimation is possible for small-scale enterprises (to some extent).



Contents

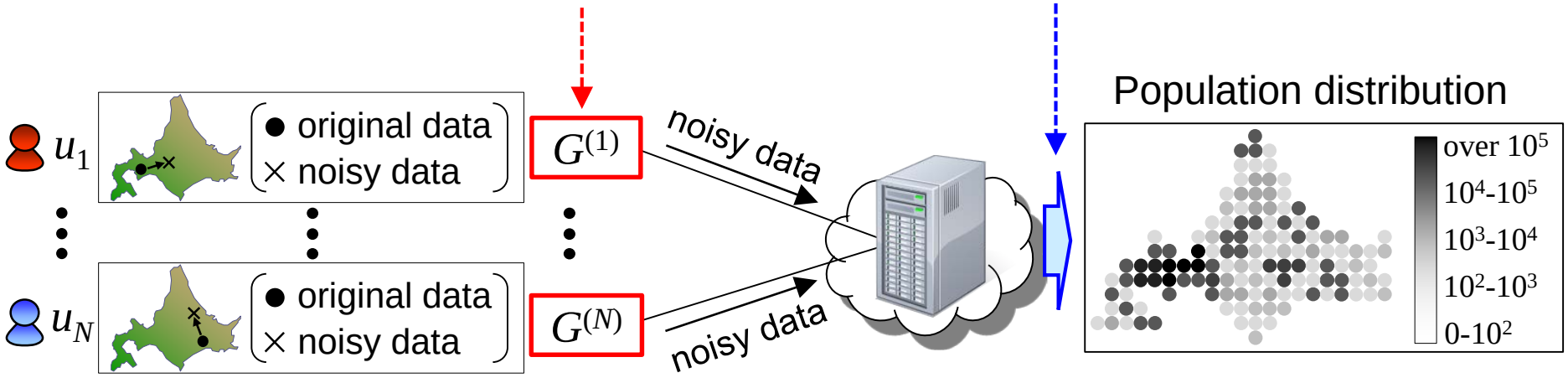
Related Work
(Obfuscation Mechanism, Distribution Estimation)

Our Proposal

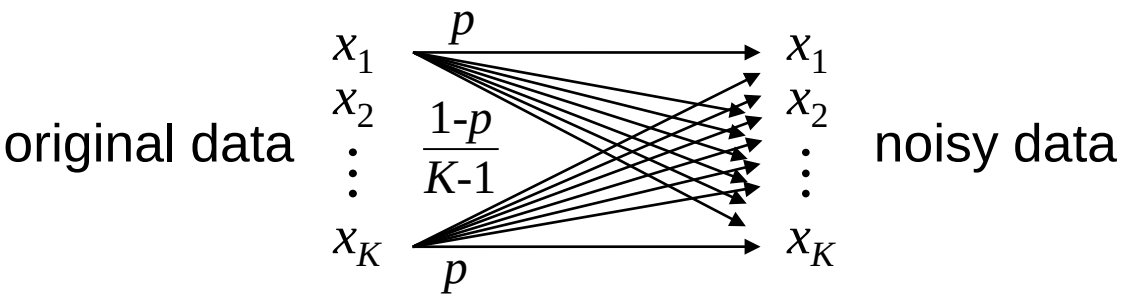
Experiments

Obfuscation Mechanism

- Distribution Estimation under LDP
  - Is composed of ***obfuscation phase*** and ***distribution estimation phase***.

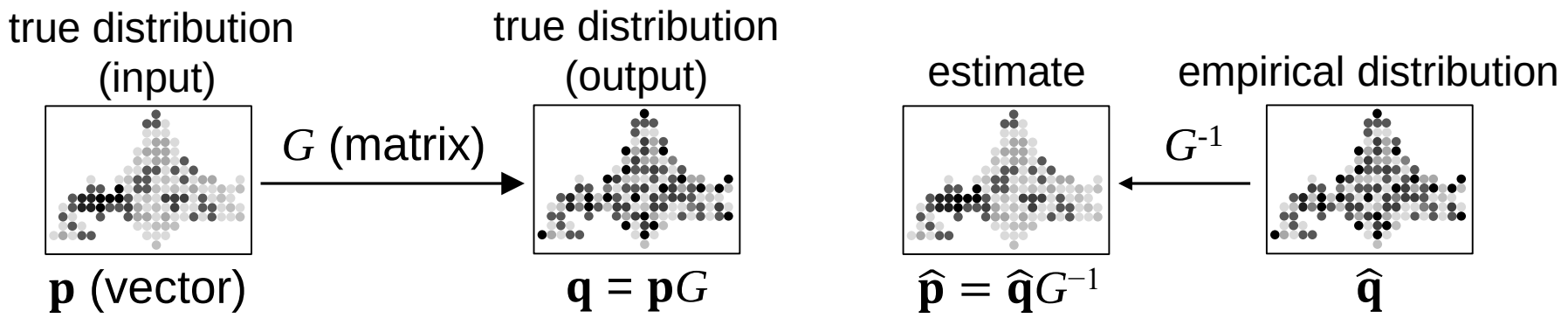


- K-RR (K-ary Randomized Response) Mechanism [Kairouz+,ICML16]
  - Generalization of Warner's binary RR to K-alphabets.
  - Sends a true value with probability $p = e^{\varepsilon} / (k - 1 + e^{\varepsilon})$. Satisfies ε-LDP.

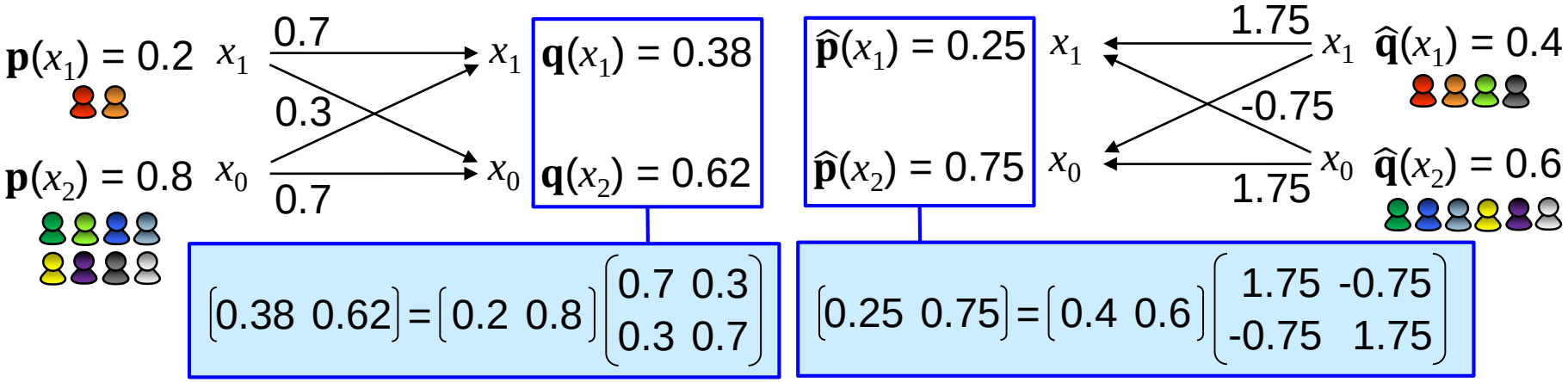


Distribution Estimation Method

- Matrix Inversion Method
  - Multiplies an empirical distribution $\hat{\mathbf{q}}$ of obfuscated data by G^{-1}: $\hat{\mathbf{p}} = \hat{\mathbf{q}}G^{-1}$.

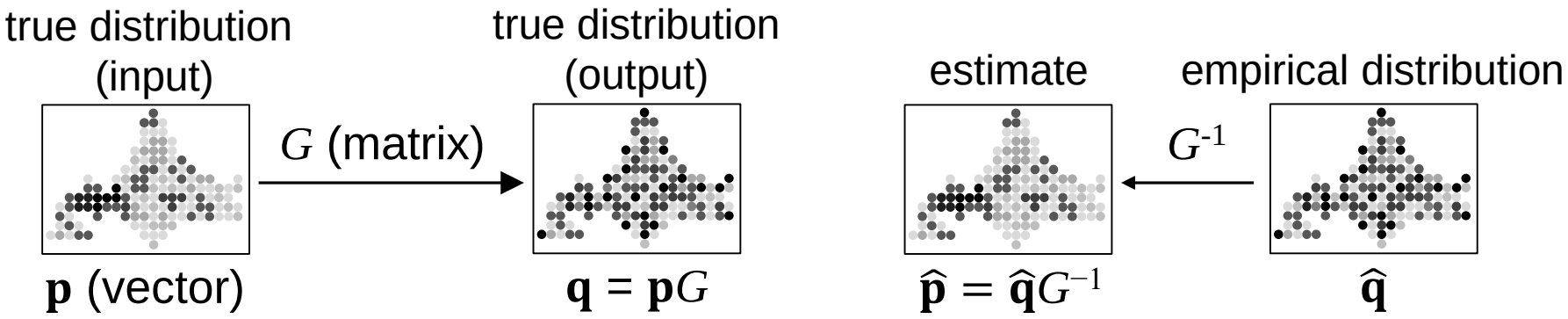


- Example:
  - Q. “Have you ever cheated in an exam?” (x_1: yes, x_0: no)

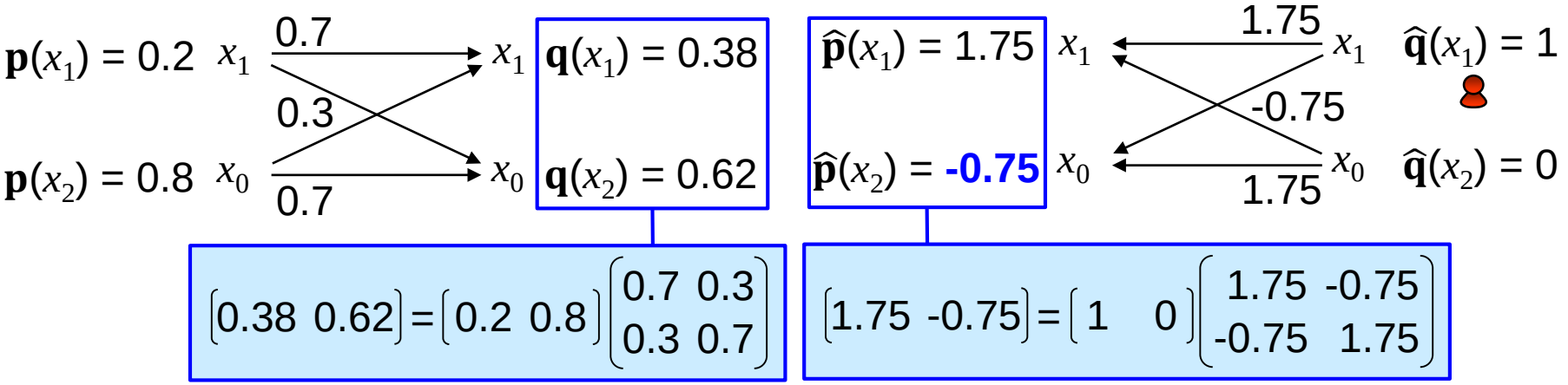


Distribution Estimation Method

- Limitation of the Matrix Inversion Method
  - If the sample size N is small, elements in $\hat{\mathbf{p}}$ can be **negative**.☹

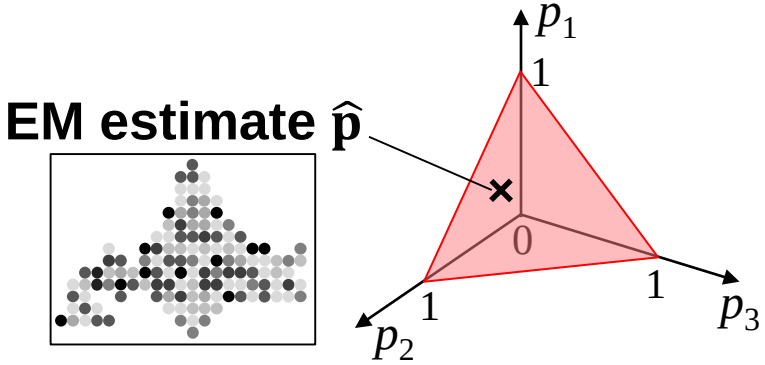


- Example (N=1):
  - Q. "Have you ever cheated in an exam?" (x_1: yes, x_0: no)

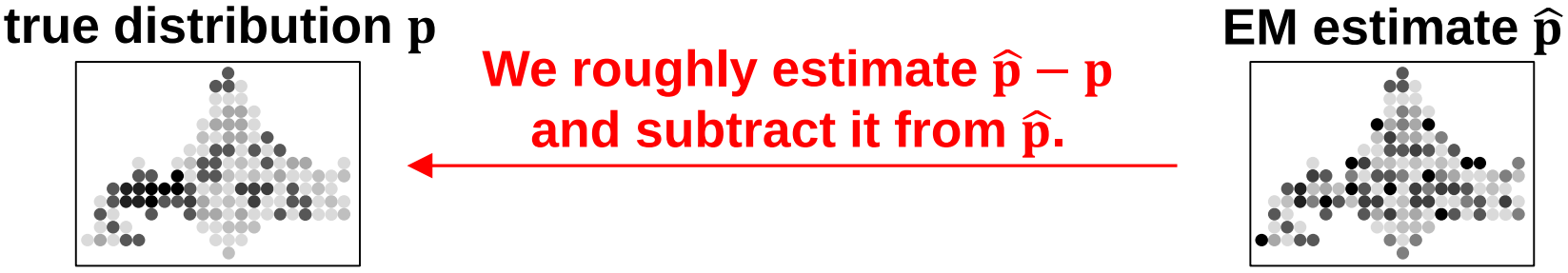


Distribution Estimation Method

- EM (Expectation-Maximization) Reconstruction Method [Agrawal+, SIGMOD05]
  - Finds the ML (maximum likelihood) estimate $\hat{\mathbf{p}}$ in the "probability simplex".
  - → Elements in $\hat{\mathbf{p}}$ are always **non-negative** (even if N is small). ☺



- Limitation of the EM Reconstruction Method
  - When the sample size N is small, the ML estimate is not accurate.



Contents

Preliminaries
(Obfuscation Mechanism, Distribution Estimation)

Our Proposal

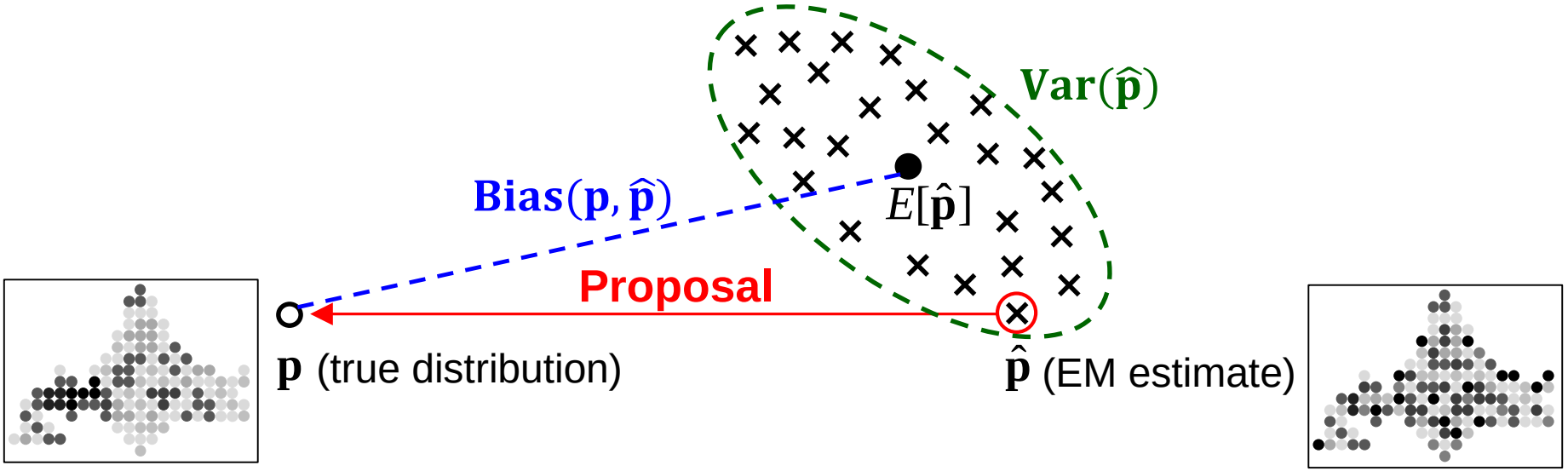
Experiments

Overview

- Proposed Method
  - Corrects the estimation error $\hat{\mathbf{p}} - \mathbf{p}$ of the EM reconstruction method.
  - Formalizes the **bias** of the EM estimate to roughly estimate $\hat{\mathbf{p}} - \mathbf{p}$.

- MSE (Mean Square Error)
  - $\mathbf{MSE} = E[||\hat{\mathbf{p}} - \mathbf{p}||_2^2]$. ($\hat{\mathbf{p}}$ depends on the realization of N noisy samples.)

$$\underbrace{\mathbf{MSE}}_{E[||\hat{\mathbf{p}} - \mathbf{p}||_2^2]} = \underbrace{||\mathbf{Bias}(\mathbf{p}, \hat{\mathbf{p}})||_2^2}_{E[\hat{\mathbf{p}}] - \mathbf{p}} + \underbrace{\mathbf{Var}(\hat{\mathbf{p}})}_{E[||\hat{\mathbf{p}} - E[\hat{\mathbf{p}}]||_2^2]}.$$



Algorithm

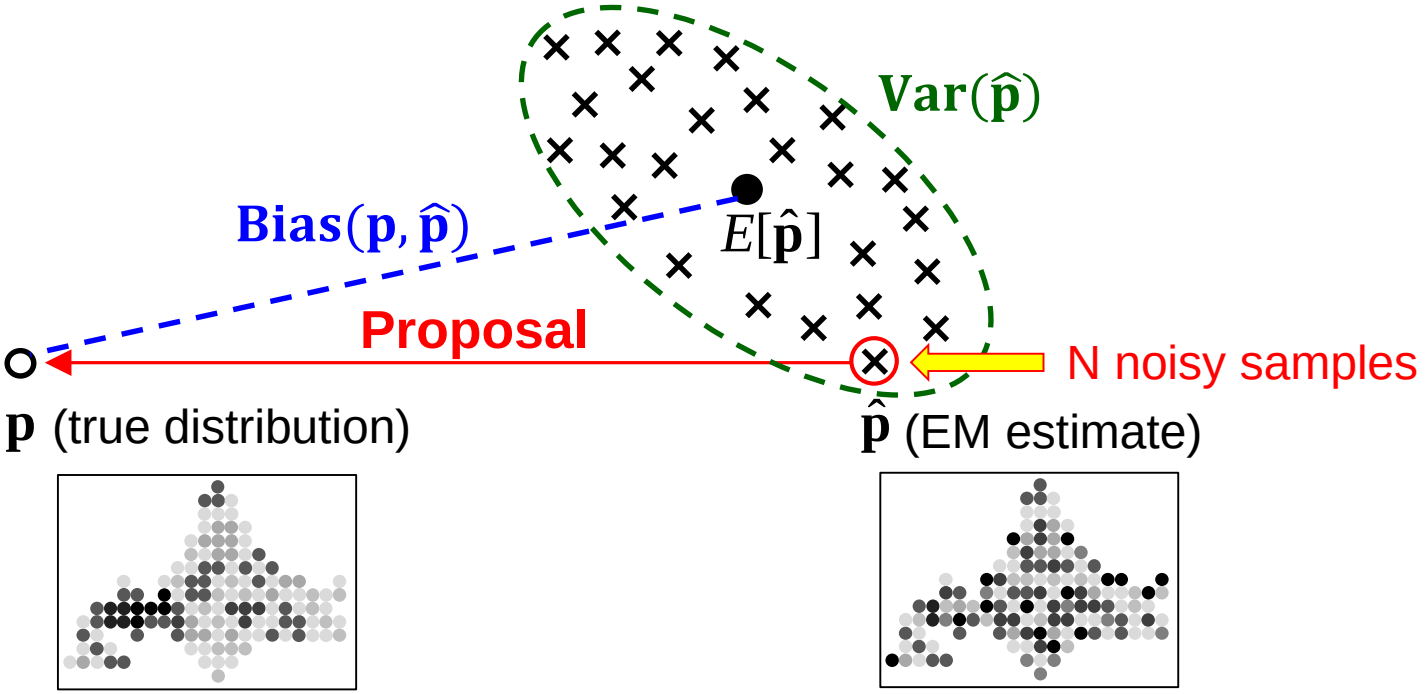
- Proposed Algorithm (for more details, please see our paper)

(1) Formalize the **bias** $(= \boldsymbol{E}[\hat{\mathbf{p}}] - \mathbf{p})$ of EM estimate based on [Rilstone+, JE96].

(2) Replace the expectation "E" with the empirical mean over N noisy samples.

- Note that $\hat{\mathbf{p}}$ is also computed based on the N noisy samples.
- → Roughly, $\boldsymbol{E}[\hat{\mathbf{p}}] - \mathbf{p}$ becomes $\hat{\mathbf{p}} - \mathbf{p}$ ("E" is removed).

(3) Subtract the rough estimate of $\hat{\mathbf{p}} - \mathbf{p}$ from $\hat{\mathbf{p}}$. → Get an approximation of $\mathbf{p}$.

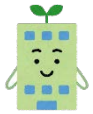


Theoretical Analysis

- MSE of the Proposed Method
  - b_1: term of $O(N^{-1})$, b_2: term of $O(N^{-3/2})$.
  - b_1 (resp. b_2) is called the first-order (resp. second-order) MSE.
  - We proved that b_2 is reduced to 0 (under some assumptions).

Method	MSE
EM Reconstruction Method	$\mathbf{MSE_{EM}} = b_1 + b_2 + O(N^{-2})$
Proposed Method	$\mathbf{MSE_{Proposal}} = \boldsymbol{b_1} + \boldsymbol{O(N^{-2})}$

- What Does It Mean?
  - Since b_1 is a dominant term, $\mathbf{MSE_{EM}} \approx \mathbf{MSE_{Proposal}}$ when N is large.
  - MSE is reduced when N is small. → We show this in our experiments.



Contents

Preliminaries
(Obfuscation Mechanism, Distribution Estimation)

Our Proposal

Experiments

Experiment 1: People-flow Dataset

- People-flow Dataset
  - Mobility traces of 303916 people around Tokyo.
  - We used one location sample per person.
  - We divided area into 20x20 regions, excluded 119 sea regions (K=281).

- True Distribution
  - We used a frequency distribution of all people as a true distribution $\mathbf{p}$.

Special wards of Tokyo

0.00E+00	0.00E+00	6.38E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00			
0.00E+00	1.12E-04	3.95E-05	1.70E-03	0.00E+00	3.29E-05	3.06E-04	1.21E-03	0.00E+00	0.00E+00	3.62E-05	0.00E+00	0.00E+00	5.03E-04	1.97E-05	0.00E+00	0.00E+00			
0.00E+00	6.91E-05	5.26E-04	1.32E-04	2.16E-03	1.22E-03	7.01E-04	1.42E-03	0.00E+00	0.00E+00	1.25E-04	0.00E+00	0.00E+00	2.86E-04	0.00E+00	0.00E+00	0.00E+00			
0.00E+00	1.65E-04	3.95E-05	8.42E-04	1.37E-03	1.97E-03	1.67E-03	4.62E-03	0.00E+00	0.00E+00	6.25E-05	2.07E-04	1.73E-03	0.00E+00	0.00E+00	0.00E+00	1.65E-05	1.97E-05		
3.29E-05	6.35E-04	0.00E+00	3.13E-04	1.54E-03	0.00E+00	7.34E-03	4.58E-03	5.57E-03	8.75E-04	4.47E-04	6.58E-06	1.86E-03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.38E-04		
3.62E-05	0.00E+00	2.63E-05	7.53E-04	4.86E-03	6.23E-03	1.06E-02	6.03E-03	6.04E-03	3.25E-03	7.27E-04	3.73E-03	1.48E-03	0.00E+00	1.41E-04	2.07E-04	1.32E-04	0.00E+00		
0.00E+00	0.00E+00	0.00E+00	1.89E-03	6.22E-03	5.43E-03	7.06E-03	1.93E-02	1.18E-02	8.33E-03	1.52E-02	1.93E-03	9.41E-04	3.52E-04	1.48E-04	6.91E-05	9.54E-05	7.57E-05		
0.00E+00	1.02E-04	4.57E-04	4.40E-03	2.82E-03	1.62E-02	1.05E-02	2.84E-02	1.33E-02	1.32E-02	1.00E-02	2.59E-03	2.73E-04	1.22E-03	1.10E-03	0.00E+00	0.00E+00	9.21E-05	1.05E-04	
0.00E+00	0.00E+00	2.67E-04	2.70E-03	1.17E-02	2.04E-02	3.21E-02	4.48E-02	4.21E-02	2.18E-02	2.30E-02	1.36E-02	5.04E-03	2.18E-03	1.97E-05	0.00E+00	1.65E-04	3.09E-04	1.05E-04	1.07E-03
0.00E+00	0.00E+00	4.28E-05	4.45E-03	1.49E-02	1.53E-02	2.36E-02	3.94E-02	1.97E-02	1.71E-02	2.09E-03	1.64E-02	4.88E-03	3.36E-04	5.30E-04	1.94E-04	3.95E-05			
0.00E+00	0.00E+00	4.05E-04	4.80E-04	9.65E-03	1.15E-02	2.55E-02	3.76E-02	8.98E-03			3.06E-03	5.02E-03	0.00E+00	7.70E-04	0.00E+00				
0.00E+00	0.00E+00	0.00E+00	0.00E+00	4.25E-03	2.32E-02	1.17E-02	2.51E-02	1.46E-03			1.34E-03	2.23E-03	1.96E-03	5.92E-05					
0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.01E-02	1.46E-02	2.89E-02	6.80E-03			8.19E-04	1.10E-03	6.25E-05	1.01E-03	0.00E+00					
1.65E-05	1.78E-04	2.86E-03	3.76E-03	1.14E-03	7.17E-03	2.83E-02	2.14E-03			1.23E-03	1.61E-04	3.52E-04	0.00E+00	4.08E-04					
0.00E+00	1.26E-03	2.26E-03	1.81E-03	7.09E-03	1.20E-02	1.20E-02	3.41E-03		8.16E-04	0.00E+00	1.18E-04	2.63E-05	1.35E-04	1.58E-04					
0.00E+00	3.88E-04	2.51E-03					1.01E-02		9.54E-05	0.00E+00	8.23E-05	2.63E-05	6.58E-06	2.47E-04					
0.00E+00	5.69E-04	3.29E-06					2.11E-03	1.32E-05	2.30E-05	0.00E+00	0.00E+00	0.00E+00	3.39E-04	6.91E-05					
0.00E+00									2.93E-04	0.00E+00	3.26E-04	1.02E-04	2.63E-05						
0.00E+00									1.58E-04	9.87E-05	3.29E-05								
0.00E+00	0.00E+00								2.83E-04	1.02E-04									

Experiment 1: People-flow Dataset

- Obfuscation Mechanism
  - We used K-RR. We set the sample size to be N = 1070.
  - 500, 500, 50, 20 users set privacy budget to be $\varepsilon = 0.1, 2, \ln(\mathrm{K}), \infty$.
    - 0.1: high privacy, 2 : middle privacy, $\ln(\mathrm{K})$: low privacy (false value with 50%).
  - We attempted 100 cases to randomly select N users from all people.

- Results
  - $\mathbf{MSE_{EM}} = \mathbf{1.68 \times 10^{-2}}$, $\mathbf{MSE_{Proposal}} = \mathbf{1.20 \times 10^{-2}}$.
  - The proposed method corrected over/underestimated values.

True distribution in Tokyo

1.05E-02	2.84E-02	1.33E-02	1.32E-02
3.21E-02	4.48E-02	4.21E-02	2.18E-02
2.36E-02	3.94E-02	1.97E-02	1.71E-02
2.55E-02	3.76E-02	8.98E-03	

EM

(1)	1.89E-06	3.04E-02	1.49E-02	2.40E-06
	1.64E-02	4.20E-02	**6.47E-02**	1.49E-02
	1.63E-02	4.52E-02	4.04E-07	9.01E-06
	1.64E-02	1.89E-06	1.89E-06	
(2)	1.29E-02	3.07E-06	6.46E-07	9.73E-07
	2.90E-02	**5.17E-02**	4.08E-02	2.62E-02
	3.36E-06	2.40E-02	1.50E-02	3.90E-02
	2.81E-02	1.65E-02	3.24E-02	
(3)	1.66E-02	1.84E-02	3.21E-07	1.17E-02
	1.46E-05	2.04E-02	4.12E-02	2.93E-02
	8.91E-05	5.21E-06	3.12E-02	**5.25E-02**
	3.68E-07	1.40E-02	3.68E-07	

Under 0.1%

Over 5%

Proposal

(1)	1.16E-03	2.16E-02	1.07E-02	1.16E-03
	1.17E-02	2.96E-02	4.55E-02	1.07E-02
	1.17E-02	3.19E-02	2.37E-03	7.56E-04
	1.18E-02	1.16E-03	1.16E-03	
(2)	9.36E-03	1.19E-03	2.41E-03	1.19E-03
	2.07E-02	3.67E-02	2.90E-02	1.87E-02
	7.78E-04	1.72E-02	1.08E-02	2.77E-02
	2.00E-02	1.19E-02	2.31E-02	
(3)	1.22E-02	1.35E-02	0.00E+00	8.64E-03
	6.11E-04	1.49E-02	2.98E-02	2.12E-02
	4.50E-04	8.19E-04	2.26E-02	3.78E-02
	2.50E-03	1.03E-02	2.50E-03	

Experiment 2: USCensus Dataset

- USCensus (1990) Dataset
  - Contains responses from 2458285 people to the US Census questions.
  - A sequence of category IDs → a "single" ID (K=8x2x5x5=400).
  - Other parameters are the same as People-flow dataset (e.g., N=1070).

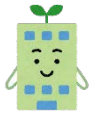
Attribute	Category ID (Value)
Age	0 (0), 1 (1-12), 2 (13-19), 3 (20-29), 4 (30-39), 5 (40-49), 6 (50-64), or 7 (65-)
Sex	0 (male) or 1 (female)
Income	0 ($0), 1 ($1-$14999), 2 ($15000-$29999), 3 ($30000-$59999), or 4 ($60000-)
Marital Status	0 (now married, except separated), 1 (widowed), 2 (divorced), 3 (separated) or 4 (never married)

- Results
  - $\mathbf{MSE_{EM}} = \mathbf{1.74 \times 10^{-2}}$, $\mathbf{MSE_{Proposal}} = \mathbf{1.42 \times 10^{-2}}$.

We also showed our proposal is effective for various values of small N and ε.

Conclusion

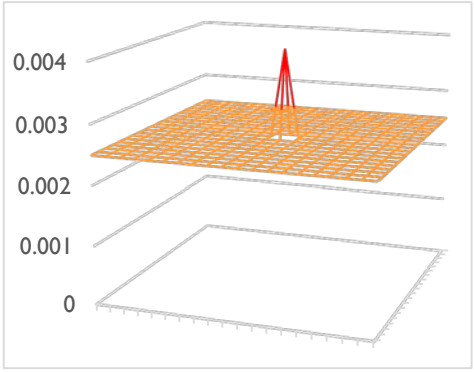
- Proposed Method
  - Reduces MSE of the EM reconstruction method when N is small.
  - Can be applied to various privacy metrics (other than LDP).

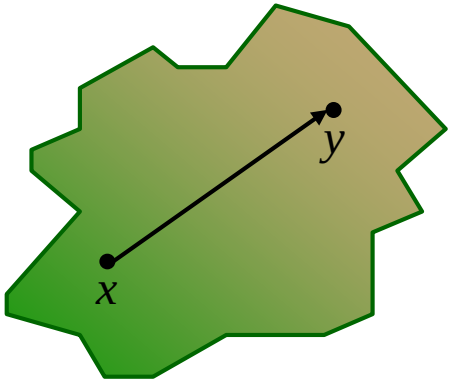


- Future Work
  - In location privacy, Geo-indistinguishability [Andres+, CCS13] is widely used.
  - We would like to use Geo-IND to significantly improve data utility.

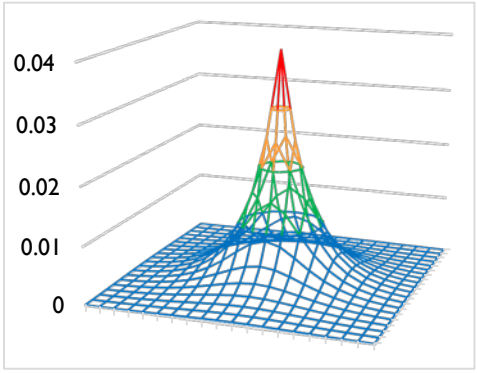
ε-LDP: $\Pr(y|x) \leq e^{\varepsilon} \Pr(y|x')$

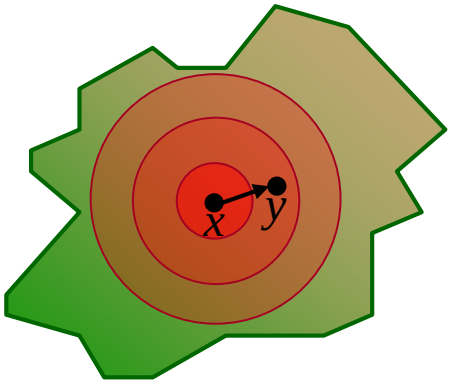
ε-Geo-IND: $\Pr(y|x) \leq e^{\varepsilon d(x,x')} \Pr(y|x')$ (d: Euclidean distance)





K-RR (ε-LDP)





2D Laplace (ε-Geo-IND)

Thank you for your attention.