Production of Categorical Data Verifying Differential Privacy: Conception and Applications to Machine Learning

Héber HWANG ARCOLEZI

Reviewer:	MCF, HDR	M. CUNCHE	INSA Lyon
Reviewer:	Pr.	B. NGUYEN	INSA Centre Val de Loire
Examiner:	Assist. Pr.	M. S. ALVIM	Federal University of Minas Gerais
Examiner :	Pr.	S. CHRÉTIEN	Université Lyon 2
Supervisor:	Pr.	J.-F. COUCHOT	Université Bourgogne Franche-Comté
Co-supervisor:	Pr.	B. AL BOUNA	Université Antonine
Co-supervisor:	Assoc. Pr.	X. XIAO	National University of Singapore

Introduction

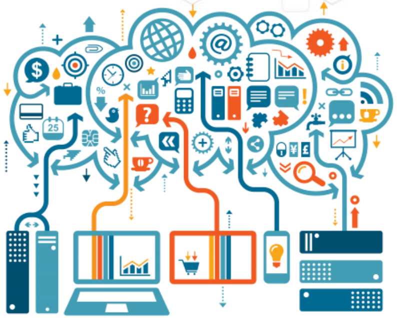
Privacy and Why Do We Need It?

Privacy:

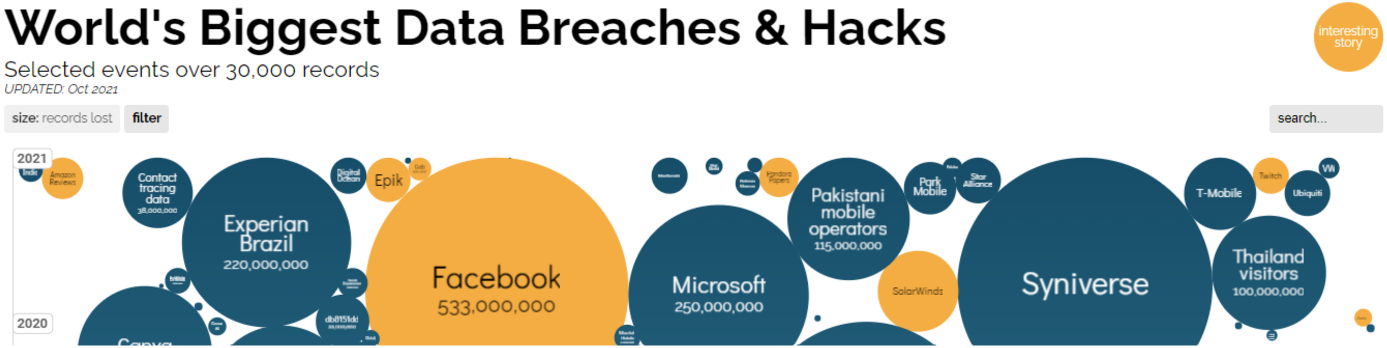
- Human right*;
- Not a new issue, aggravated by Big Data;
- Legitimate but harmful use of users' information**;
- Illegitimate access or massive data breaches***;

Societal Impact:

- Public health;
- National security;
- Development;
- Governance...



Cambridge Analytica



* https://www.un.org/en/about-us/universal-declaration-of-human-rights
** https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge_Analytica_data_scandal
*** https://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

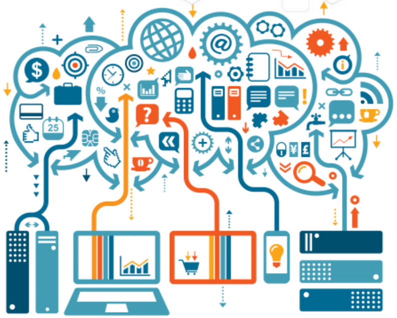
Privacy and Why Do We Need It?

Privacy:

- Human right*;
- Not a new issue, aggravated by Big Data;
- Legitimate but harmful use of users' information**;
- Illegitimate access or massive data breaches***;
- There is a need for privacy-preserving systems;
- A balance needs to be found between privacy and utility.

Societal Impact:

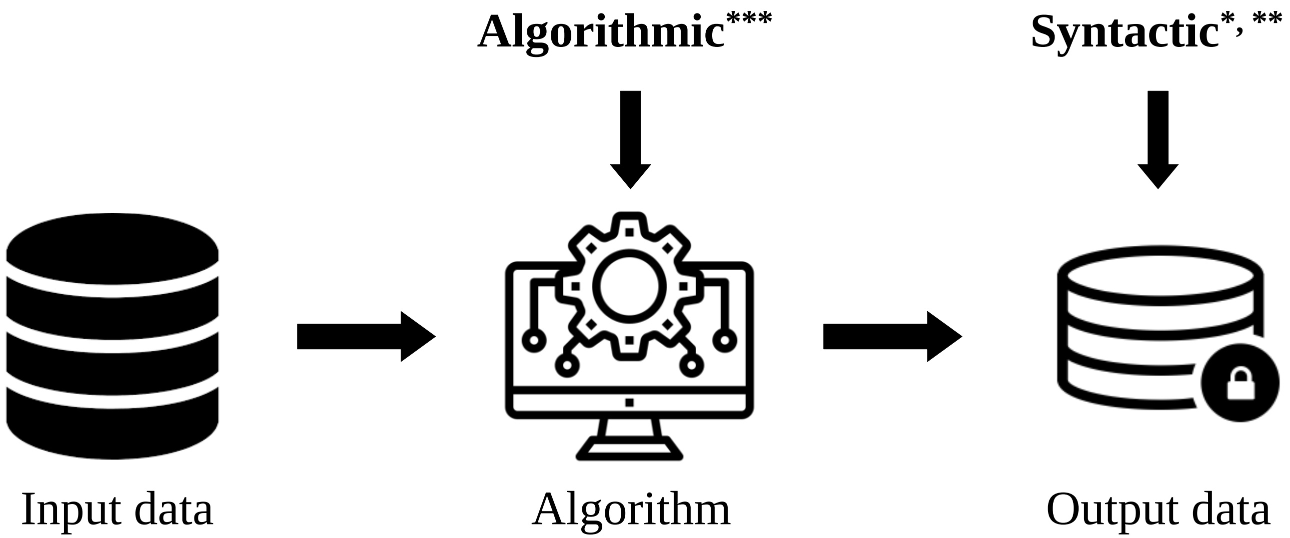
- Public health;
- National security;
- Development;
- Governance...



Privacy  Utility

* https://www.un.org/en/about-us/universal-declaration-of-human-rights
** https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge_Analytica_data_scandal
*** https://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

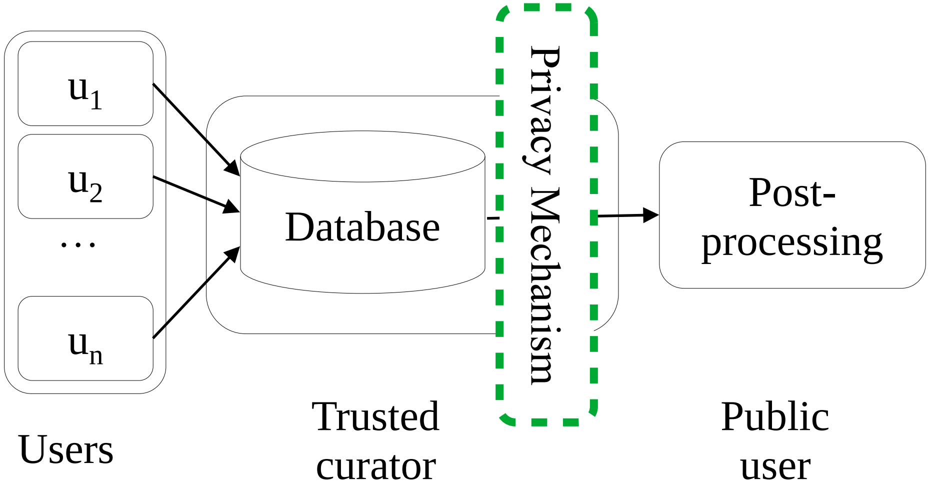
Privacy Notions: Syntactic *vs* Algorithmic



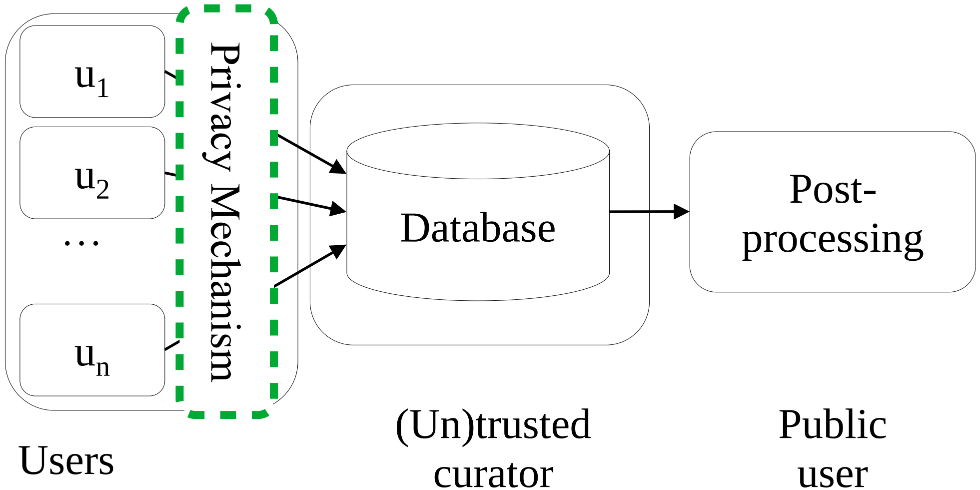
Privacy may break down if the attacker has supplementary knowledge.

[*] Sweeney, L. k-anonymity: A model for protecting privacy. In: International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (2002).
[**] Machanavajjhala, A., Kifer, D., Gehrke, J., Venkitasubramaniam, M. l-diversity: Privacy beyond k-anonymity. In: ACM Transactions on Knowledge Discovery from Data (2007).
[***] Dwork, C., Roth, A. The algorithmic foundations of differential privacy. In: Foundations and Trends in Theoretical Computer Science (2014).

The Trust Model: Centralized *vs* Local



Centralized setting

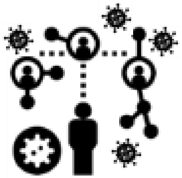


Local setting

Use of Big Data for Mobility Analytics

- Human mobility analysis through cell phone data (call detail record – CDR);

- Some motivations →



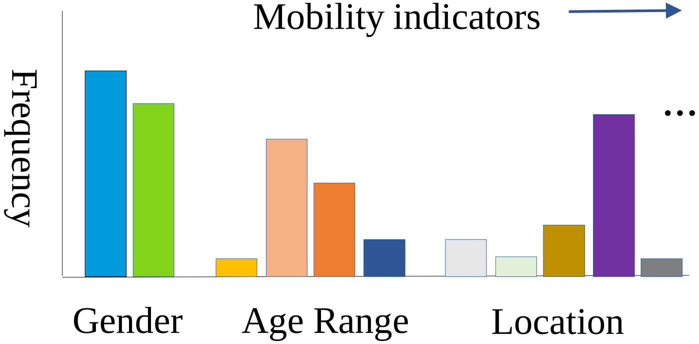
Flux Vision Mobility models

orange

Geographic area

Mobility indicators →

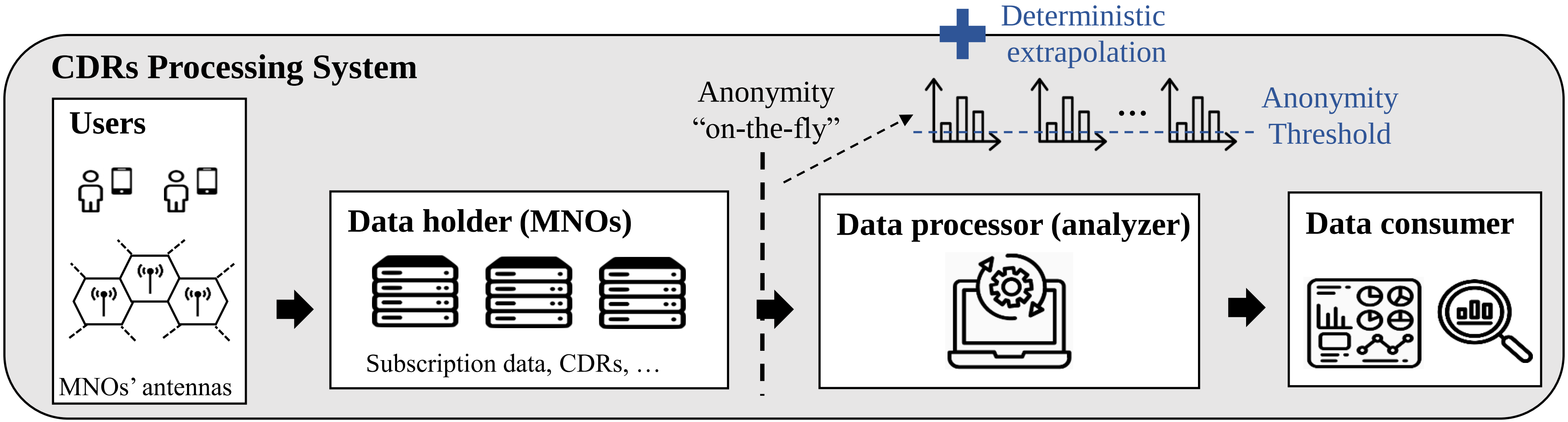
By hour;
By day;
By cumulative days...



* Flux Vision System: https://www.orange-business.com/fr/produits/flux-vision

Anonymity-Based Mobility Reports

- Human mobility is quite unique* → Mobile network operators (MNOs) must respect users' privacy;
- Users cannot sanitize their data → CDRs are automatically generated on MNOs' servers;



* De Montjoye, Y.A., Hidalgo, C.A., Verleysen, M., Blondel, V.D. Unique in the crowd: The privacy bounds of human mobility. In: Scientific reports (2013).

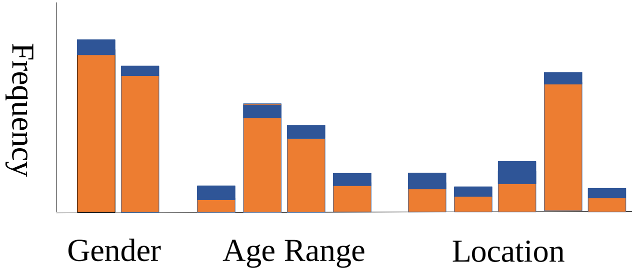
Anonymity-Based Mobility Reports

Anonymity-based solution:

- Not robust to supplementary knowledge of attackers;
- One cannot account for the privacy leak of individuals;
- Releasing raw aggregates may still be subject to privacy attacks[*, **];

Differential privacy[***]-based solution:

- Release histograms with differential privacy guarantees;
- Ex. of industry application: Google Mobility Reports[****]…



[*] Pyrgelis, A.,Troncoso, C., De Cristofaro, E. What Does The Crowd Say About You? Evaluating Aggregation-based Location Privacy. In: PoPETS (2017).
[**] Tu, Z., Xu, F., Li, Y., Zhang, P. and Jin, D., 2018. A new privacy breach: User trajectory recovery from aggregated mobility data. In: IEEE/ACM Transactions on Networking (2018).
[***] Dwork, C., Roth, A. The algorithmic foundations of differential privacy. In: Foundations and Trends in Theoretical Computer Science (2014).
[****] Google COVID-19 Community Mobility Reports: https://www.google.com/covid19/mobility/

Differential Privacy (DP)*: DP → Local DP

A randomized algorithm $\mathcal{A}$ satisfies ϵ-DP, if for **any two neighbouring databases *D* and *D*′** and for any output O of $\mathcal{A}$:

Intuitively: Any output should be about as likely regardless of whether I am in the database or not.

$$\Pr[\mathcal{A}(D) = O] \leq \underbrace{e^{\epsilon}}_{\text{Privacy loss}} \cdot \Pr[\mathcal{A}(D') = O]$$

Run by a trusted server

A randomized algorithm $\mathcal{A}$ satisfies ϵ-local-differential-privacy (ϵ-LDP), if for **any two inputs x and x'** and for any output y of $\mathcal{A}$:

Intuitively: Any output should be about as likely regardless of my secret.

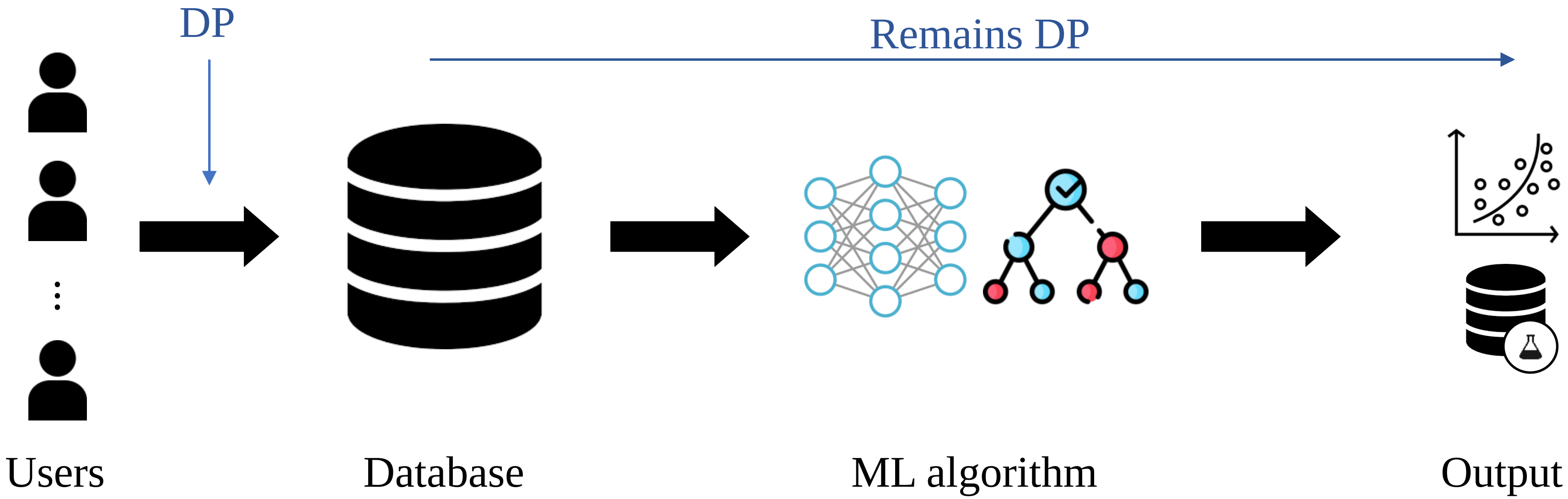
$$\Pr[\mathcal{A}(x) = y] \leq \underbrace{e^{\epsilon}}_{\text{Privacy loss}} \cdot \Pr[\mathcal{A}(x') = y]$$

Run by each user

* Dwork, C., Roth, A. The algorithmic foundations of differential privacy. In: Foundations and Trends in Theoretical Computer Science (2014).

Properties of DP [*]: Post-Processing

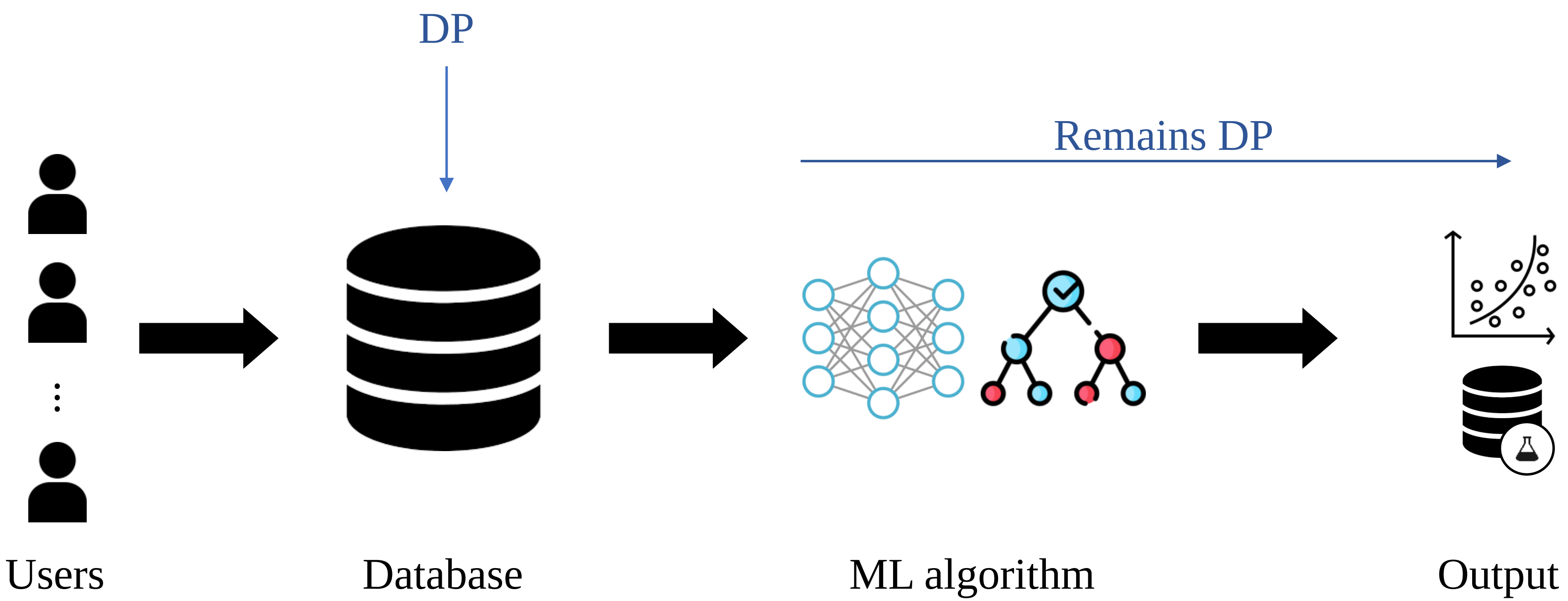
- **Robust to post-processing** → if $\mathcal{A}$ is ϵ-DP, then $f(\mathcal{A})$ is also ϵ-DP for any f.



[*] Dwork, C., Roth, A. The algorithmic foundations of differential privacy. In: Foundations and Trends in Theoretical Computer Science (2014).

Properties of DP*: Post-Processing

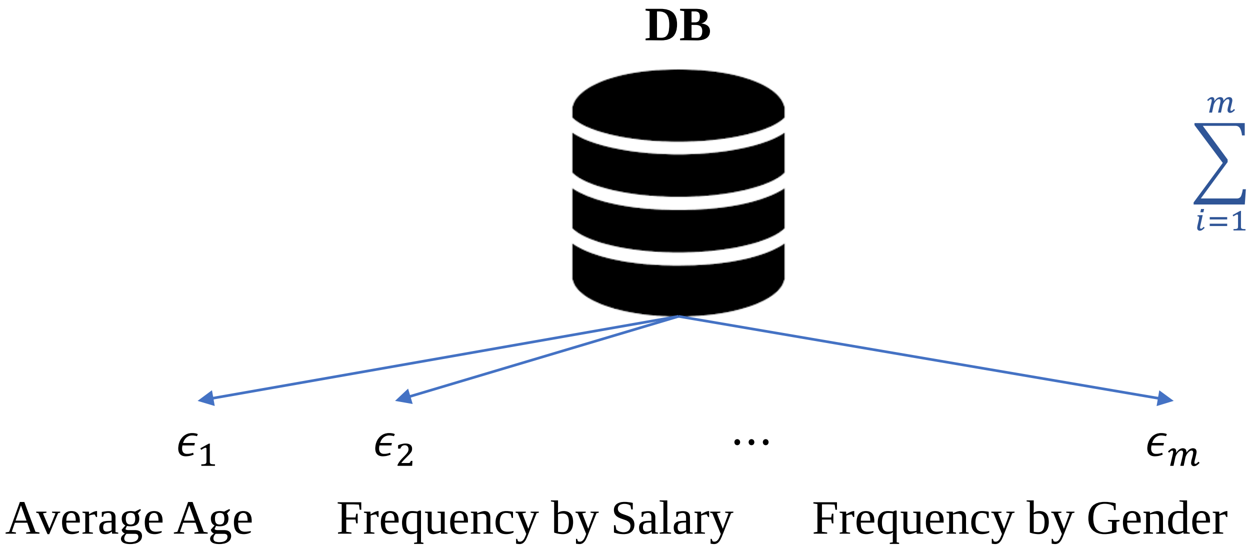
- **Robust to post-processing** $\rightarrow$ if $\mathcal{A}$ is ϵ-DP, then $f(\mathcal{A})$ is also ϵ-DP for any f.



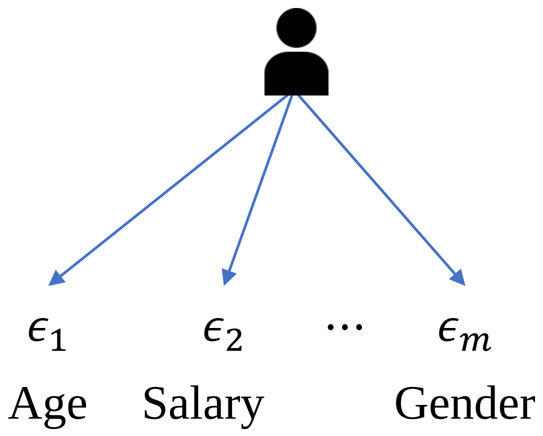
* Dwork, C., Roth, A. The algorithmic foundations of differential privacy. In: Foundations and Trends in Theoretical Computer Science (2014).

Properties of DP *: Composition

- **Composition** → DP allows to accounting for the overall privacy loss when several DP algorithms are applied to the same database (DB).



$$\sum_{i=1}^{m} \epsilon_i$$



* Dwork, C., Roth, A. The algorithmic foundations of differential privacy. In: Foundations and Trends in Theoretical Computer Science (2014).

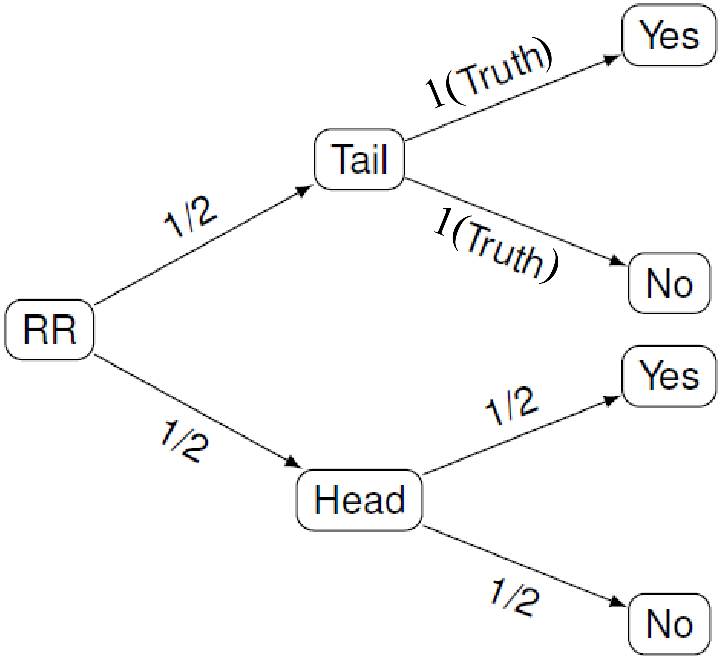
LDP: Ex. of Randomized Response (RR)*

- Motivated by surveying people on sensitive/embarrassing topics.
- Main idea → Providing **deniability** to users' answer (yes/no → binary).
- Ask: "Did you test positive for HIV (human immunodeficiency virus)?"
- Each person:
  - Throw a secret unbiased coin:
    - If tail, throw the coin again (ignoring the outcome) and answer the question honestly.
    - If head, then throw the coin again and answer "Yes" if head, "No" if tail.

RR: Seeing answer, still not certain about the secret.

* Warner, S.L. Randomized response: A survey technique for eliminating evasive answer bias. In: Journal of the American Statistical Association (1965).

Frequency Estimation and ϵ Study of RR

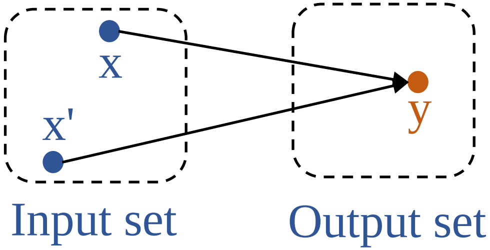


- $f(v_Y) \rightarrow$ frequency of *true Yes* (or *No* – v_N)
- $\approx \hat{f}(v_i) = \frac{N_i - nq}{(p-q)}, \forall_{i \in \{Y,N\}}$ ⇢ Estimated frequency
- Satisfies ϵ-LDP w/:

$$\frac{\Pr(y|x)}{\Pr(y|x')} \leq e^{\epsilon} \Rightarrow e^{\epsilon} = \frac{0.75}{0.25}, \epsilon = \ln(3)$$

prob. p of 'being honest'

prob. q of 'lying'



$p = \Pr[RR(Yes) = Yes] = \Pr[RR(No) = No] = 0.75$

$q = \Pr[RR(No) = Yes] = \Pr[RR(Yes) = No] = 0.25$

LDP Implem. of Big Tech Companies

RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response

Úlfar Erlingsson
Google, Inc.
ulfar@google.com

Vasyl Pihur
Google, Inc.
vpihur@google.com

Aleksandra Korolova
University of Southern California
korolova@usc.edu

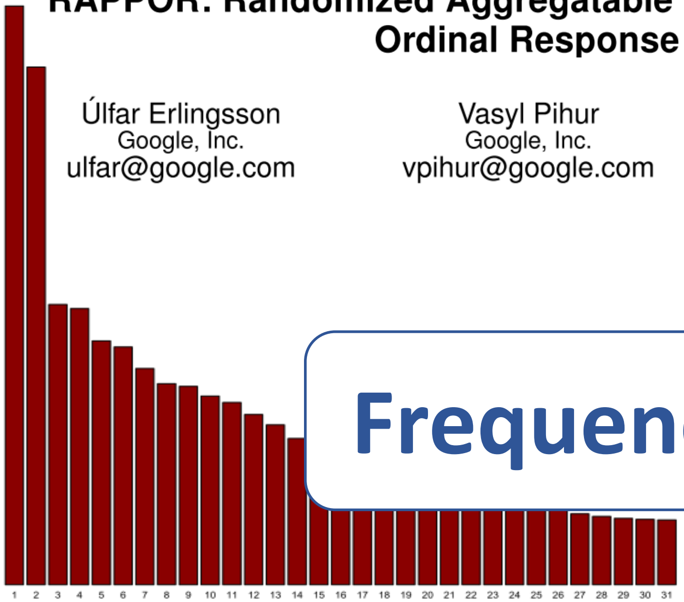
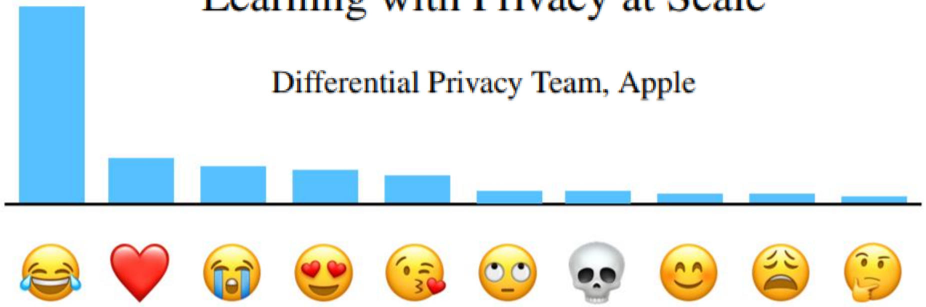


Figure 6: Relative frequencies of the top 31 unexpected Chrome homepage domains found by analyzing ~14 million RAPPOR reports, excluding expected domains (the homepage "google.com", etc.).

Learning with Privacy at Scale

Differential Privacy Team, Apple



nost popular emoji to help
i for US English speakers

Frequency (histogram) estimation

Collecting Telemetry Data Privately

Bolin Ding, Janardhan Kulkarni, Sergey Yekhanin
Microsoft Research
{bolind, jakul, yekhanin}@microsoft.com

Windows Insiders in Windows 10 Fall Creators Update to protect users' privacy while collecting application usage statistics.

LDP Protocols for Frequency Estimation

- **Generalized RR (GRR)**[*]: Extends RR to the case of $k_j \geq 2$.

$$\forall_y \in A_j \ \Pr[\mathcal{A}_{GRR(\epsilon)}(v) = y] = \begin{cases} p = \frac{e^{\epsilon}}{e^{\epsilon}+k_j-1}, if\ y = v \\ q = \frac{1}{e^{\epsilon}+k_j-1}, if\ y \neq v \end{cases} \qquad \epsilon = \ln\left(\frac{p}{q}\right)$$

- **Unary Encoding (UE)**[**]: Encode as a bit-vector B and perturb each bit independently into a new bit-vector B'. More specifically:

$$\Pr[B'_i = 1] = \begin{cases} p,\ if\ B_i = 1 \\ q,\ if\ B_i = 0 \end{cases} \qquad \epsilon = \ln\left(\frac{p(1-q)}{q(1-p)}\right)$$

Symmetric UE (SUE): $p = \frac{e^{\epsilon/2}}{e^{\epsilon/2}+1},\ q = \frac{1}{e^{\epsilon/2}+1},$ **Optimized UE (OUE)**[***]: $p = \frac{1}{2},\ q = \frac{1}{e^{\epsilon}+1}$

[*] Kairouz, P., Oh, S., Viswanath, P. Extremal mechanisms for local differential privacy. In: NeurIPS (2014).
[**] Erlingsson, Ú., Pihur, V. and Korolova, A. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. In: SIGSAC (2014).
[***] Wang, T., Blocki, J., Li, N. and Jha, S. Locally differentially private protocols for frequency estimation. In: USENIX Security Symposium (2017).

LDP Protocols for Frequency Estimation

- Unbiased* normalized frequency estimation $f(v_i)$ for $v_i \in A_j$:

$$\hat{f}(v_i) = \frac{N_i - nq}{n(p-q)}$$

N_i = number of times the value v_i or bit i has been reported.

- Variance of the estimator*:

$f(v_i) = 0 \longrightarrow$ Approximate Var^*

$p + q = 1$ "symmetric"

$$\mathrm{Var}[\hat{f}(v_i)] = \frac{q(1-q)}{n(p-q)^2} + \frac{f(v_i)(1-p-q)}{n(p-q)}$$

* Wang, T., Blocki, J., Li, N. and Jha, S. Locally differentially private protocols for frequency estimation. In: USENIX Security Symposium (2017).

Outline

1. Introduction
2. Multiple Frequency Estimates Under Local Differential Privacy
3. Privacy-Utility Trade-off of Differentially Private Machine Learning Models
4. Further Contributions
5. Conclusion & Perspectives

Outline

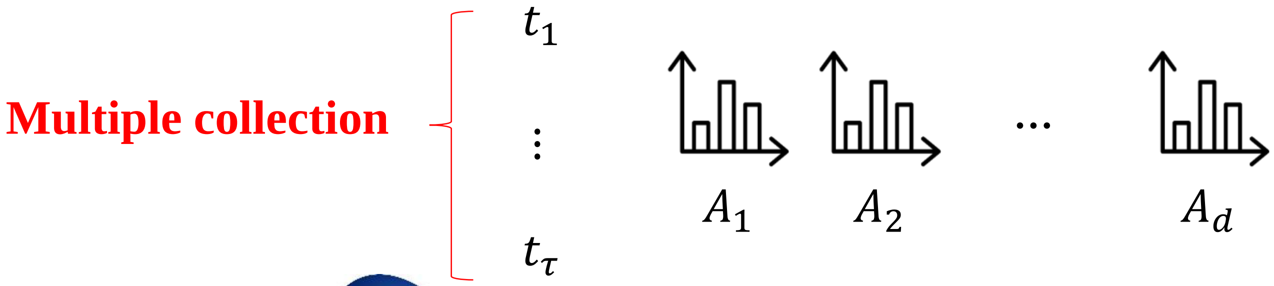
1. Introduction
2. Multiple Frequency Estimates Under Local Differential Privacy
   i. **Longitudinal and Multidimensional Data Collection**
   ii. Multidimensional Data Collection
3. Privacy-Utility Trade-off of Differentially Private Machine Learning Models
4. Further Contributions
5. Conclusion & Perspectives

Problem Statement: Statistical Learning

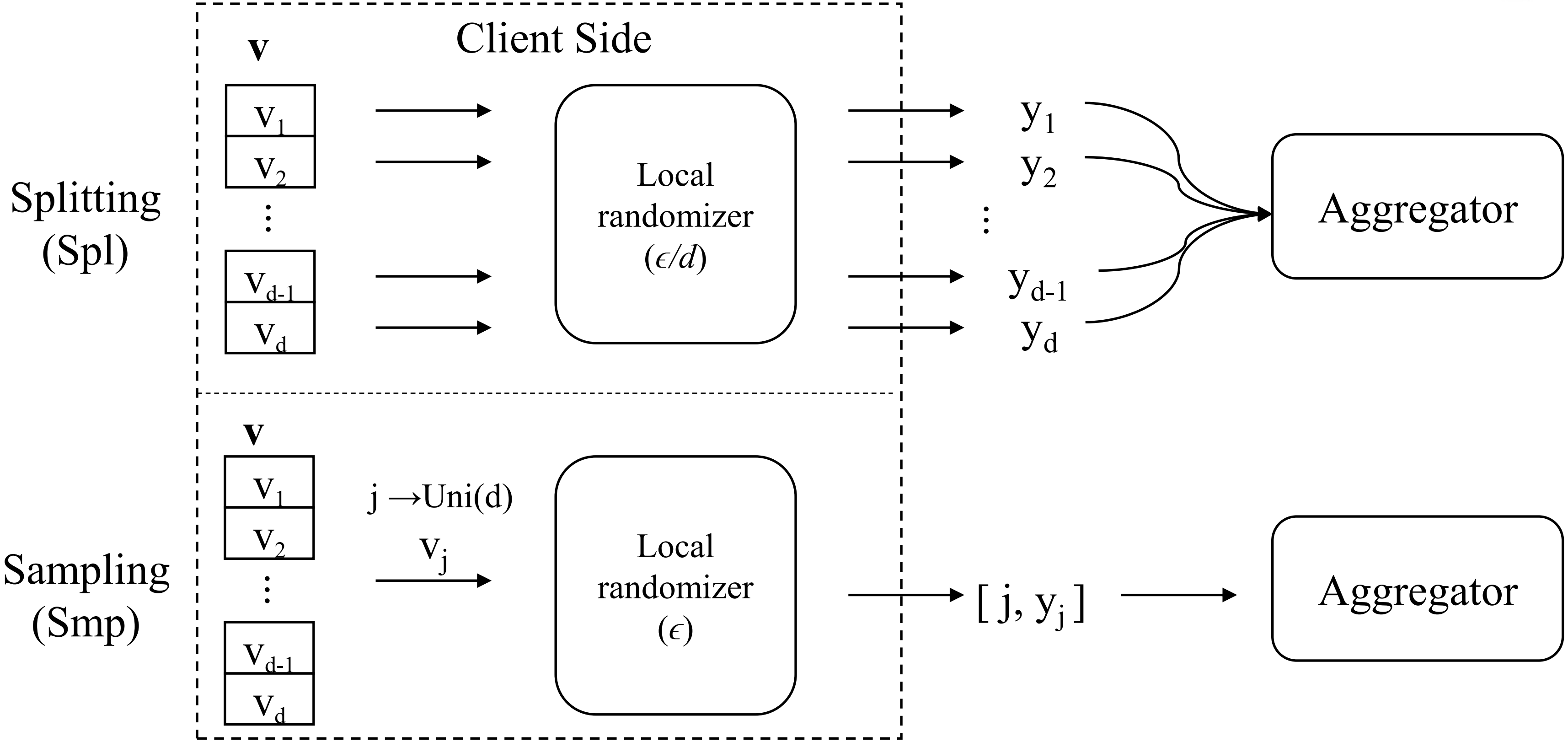
- **Tackled Issue:** Collecting *multidimensional* data under ϵ-*LDP* throughout time (i.e., *longitudinal study*) for *frequency estimation*.

- **More formally (notation):**

  - d attributes $A = \{A_1, A_2, \ldots, A_d\}$; **Multiple attributes**
  - Each attribute A_j has a discrete domain of size $|A_j| = k_j$;
  - Each user u_i for $1 \leq i \leq n$ has a tuple $\mathbf{v}^i = (v_1^i, v_2^i, \ldots, v_d^i)$;
  - **Analyzer:** estimate a k_j-bins histogram for each attribute $j \in [1, d]$.



State-of-the-Art for Multiple Attributes*, **



* Nguyên, T.T., Xiao, X., Yang, Y., Hui, S.C., Shin, H., Shin, J. Collecting and analyzing data from smart device users with local differential privacy. In: arXiv:1606.05053 (2016).
** Wang, N., Xiao, X., Yang, Y., Zhao, J., Hui, S.C., Shin, H., Shin, J., Yu, G. Collecting and analyzing multidimensional data with local differential privacy. In: ICDE (2019).

Multidimensional Frequency Estimates

- ϵ : privacy budget;
- d : total number of attributes;
- n : total number of users.

number of attributes each user will sample

Sampling-based solution[*]: Find r that minimizes the variance of each protocol[**].

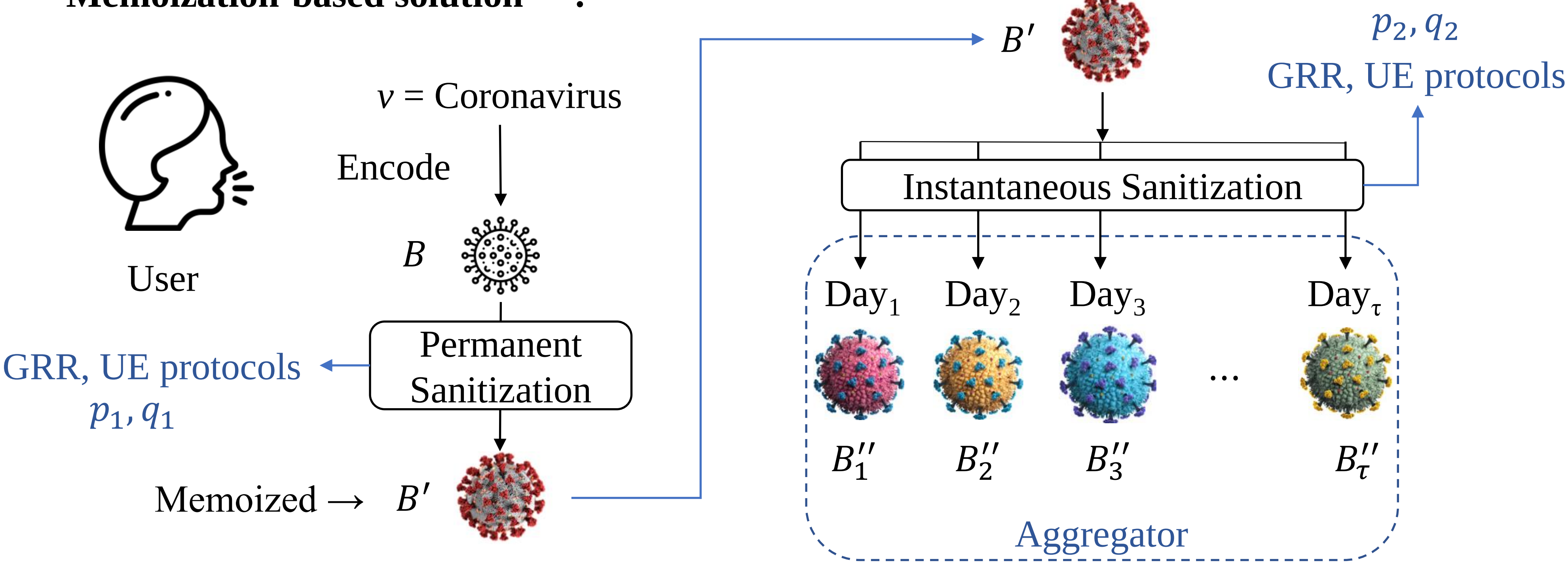
$$Var[\hat{f}_{GRR}] = \frac{d\left(e^{\epsilon/r} + k_j - 2\right)}{nr(e^{\epsilon/r} - 1)^2} \qquad Var[\hat{f}_{SUE}] = \frac{d\left(e^{\epsilon/2r}\right)}{nr(e^{\epsilon/2r} - 1)^2} \qquad Var[\hat{f}_{OUE}] = \frac{d\left(4e^{\epsilon/r}\right)}{nr(e^{\epsilon/r} - 1)^2}$$

- Variance is minimized for sampling (Smp, i.e., $r = 1$), as in[*, **].

[*] Nguyên, T.T., Xiao, X., Yang, Y., Hui, S.C., Shin, H., Shin, J. Collecting and analyzing data from smart device users with local differential privacy. In: arXiv:1606.05053 (2016).
** Wang, T., Blocki, J., Li, N. and Jha, S. Locally differentially private protocols for frequency estimation. In: USENIX Security Symposium (2017).

Longitudinal Frequency Estimates

Memoization-based solution*, **:



* Erlingsson, Ú., Pihur, V., Korolova, A. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. In: ACM SIGSAC (2014).

** Ding, B., Kulkarni, J., Yekhanin, S. Collecting telemetry data privately. In: NeurIPS (2017).

Memoization-based: Estimator and Variance

- Unbiased normalized longitudinal frequency estimation $f_L(v_i)$ for $v_i \in A_j$:

$$\hat{f}_L(v_i) = \frac{\frac{N_i - nq_2}{(p_2 - q_2)} - nq_1}{n(p_1 - q_1)} \rightarrow \frac{N_i - nq_1(p_2 - q_2) - nq_2}{n(p_1 - q_1)(p_2 - q_2)}$$

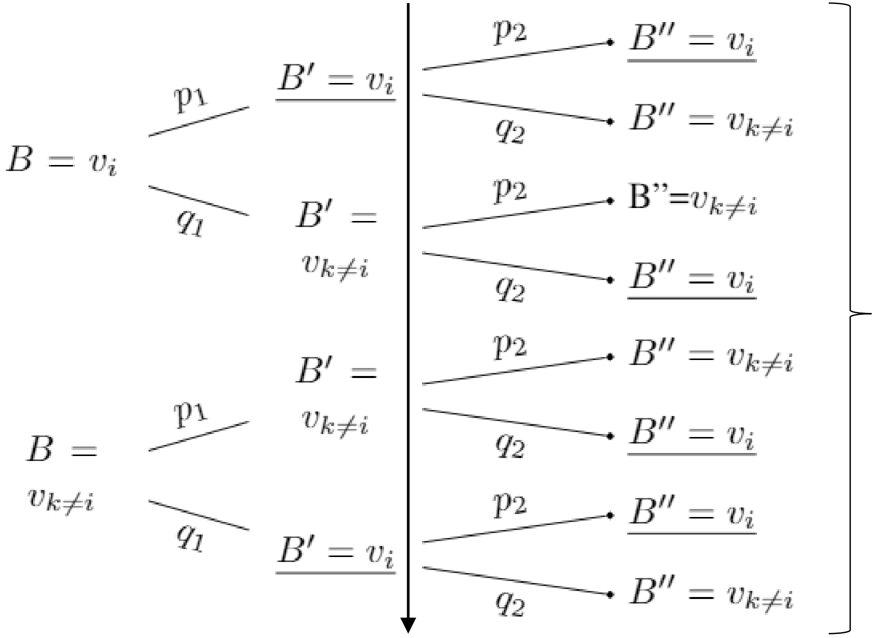
N_i = number of times the value v_i or bit i has been reported.

- Approximate variance of the estimator:

$$\mathrm{Var}^*[\hat{f}_L(v_i)] = \frac{\big(p_2 q_1 - q_2(q_1 - 1)\big)(-p_2 q_1 + q_2(q_1 - 1) + 1)}{n(p_1 - q_1)^2 (p_2 - q_2)^2}$$

Unbiased estimation and variance development in the manuscript

Longitudinal GRR: ϵ study



$$\Pr[B''|B] = \begin{cases} \Pr[B'' = v_i | B = v_i] = p_1 p_2 + q_1 q_2 \\ \Pr[B'' = v_{k\neq i} | B = v_i] = p_1 q_2 + q_1 p_2 \\ \Pr[B'' = v_i | B = v_{k\neq i}] = p_1 q_2 + q_1 p_2 \\ \Pr[B'' = v_{k\neq i} | B = v_{k\neq i}] = p_1 p_2 + q_1 q_2 \end{cases}$$

First report: $\epsilon_1 = \ln\left(\frac{p_1 p_2 + q_1 q_2}{p_1 q_2 + q_1 p_2}\right)$

Infinity reports:

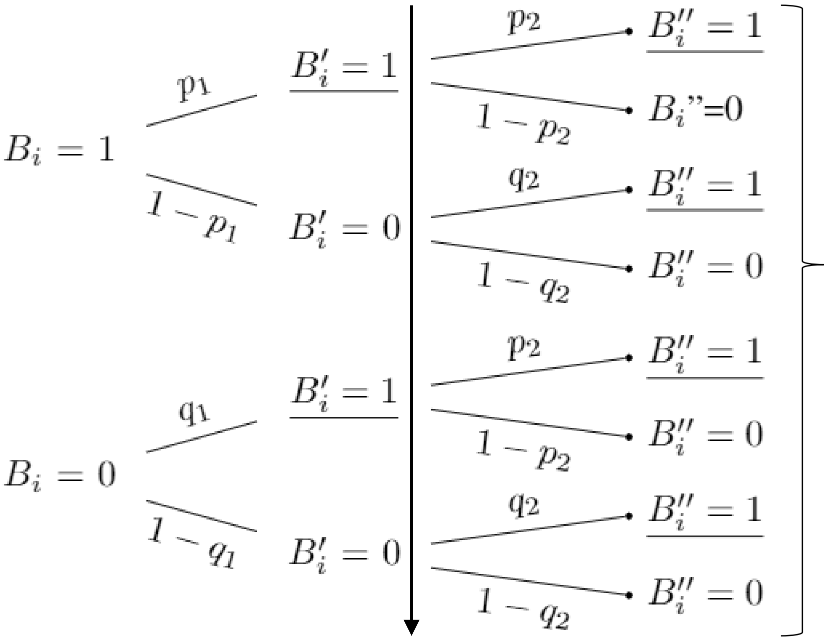
$$\epsilon_\infty = \ln\left(\frac{p_1}{q_1}\right)$$

Given ϵ_∞ and ϵ_1:

$$p_1 = \frac{e^{\epsilon_\infty}}{e^{\epsilon_\infty} + k_j - 1}, q_1 = \frac{1 - p_1}{k_j - 1}$$

$$p_2 = \frac{e^{\epsilon_1 + \epsilon_\infty} - 1}{-k_j e^{\epsilon_1} + (k_j - 1)e^{\epsilon_\infty} + e^{\epsilon_1} + e^{\epsilon_\infty + \epsilon_1} - 1}, q_2 = \frac{1 - p_2}{k_j - 1}$$

Longitudinal UE: ϵ study



$$\Pr[B''_i|B_i] = \begin{cases} \Pr[B''_i = 1|B_i = 1] = p_1p_2 + (1-p_1)q_2 \\ \Pr[B''_i = 0|B_i = 1] = p_1(1-p_2) + (1-p_1)(1-q_2) \\ \Pr[B''_i = 1|B_i = 0] = q_1p_2 + (1-q_1)q_2 \\ \Pr[B''_i = 0|B_i = 0] = q_1(1-p_2) + (1-q_1)(1-q_2) \end{cases}$$

First report:

$$\epsilon_1 = \ln\left(\frac{(p_1p_2 - q_2(p_1-1))(p_2q_1 - q_2(q_1-1) - 1)}{(p_2q_1 - q_2(q_1-1))(p_1p_2 - q_2(p_1-1) - 1)}\right)$$

Infinity reports:

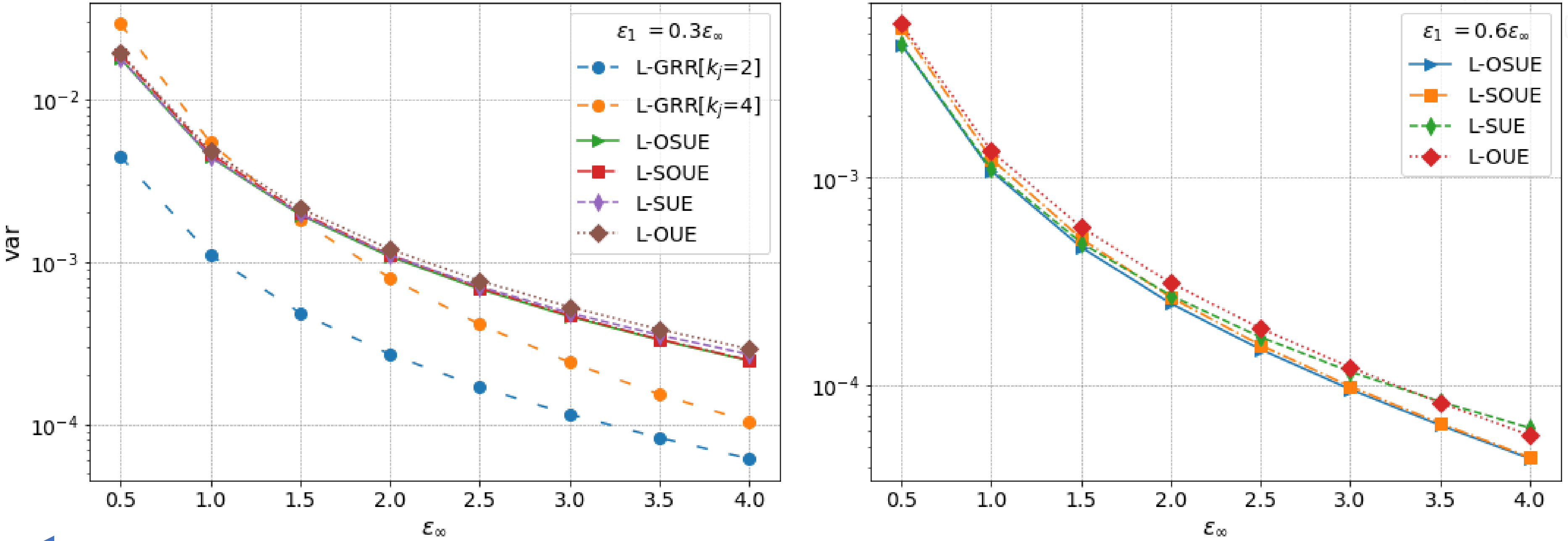
$$\epsilon_\infty = \ln\left(\frac{p_1(1-q_1)}{(1-p_1)q_1}\right)$$

Given SUE and OUE:

- Apply OUE twice (L-OUE);
- Apply SUE twice (L-SUE);
- OUE then SUE (L-OSUE);
- SUE then OUE (L-SOUE).

Num. Eval. of L-GRR and L-UE Variances

More accurate



More private

Adaptive LDP for LOngitudinal and Multidimensional FREquency Estimates (ALLOMFREE): $\min\left(Var^*\left[\hat{f}_{L_{(L-GRR)}}\right], Var^*\left[\hat{f}_{L_{(L-OSUE)}}\right]\right)$

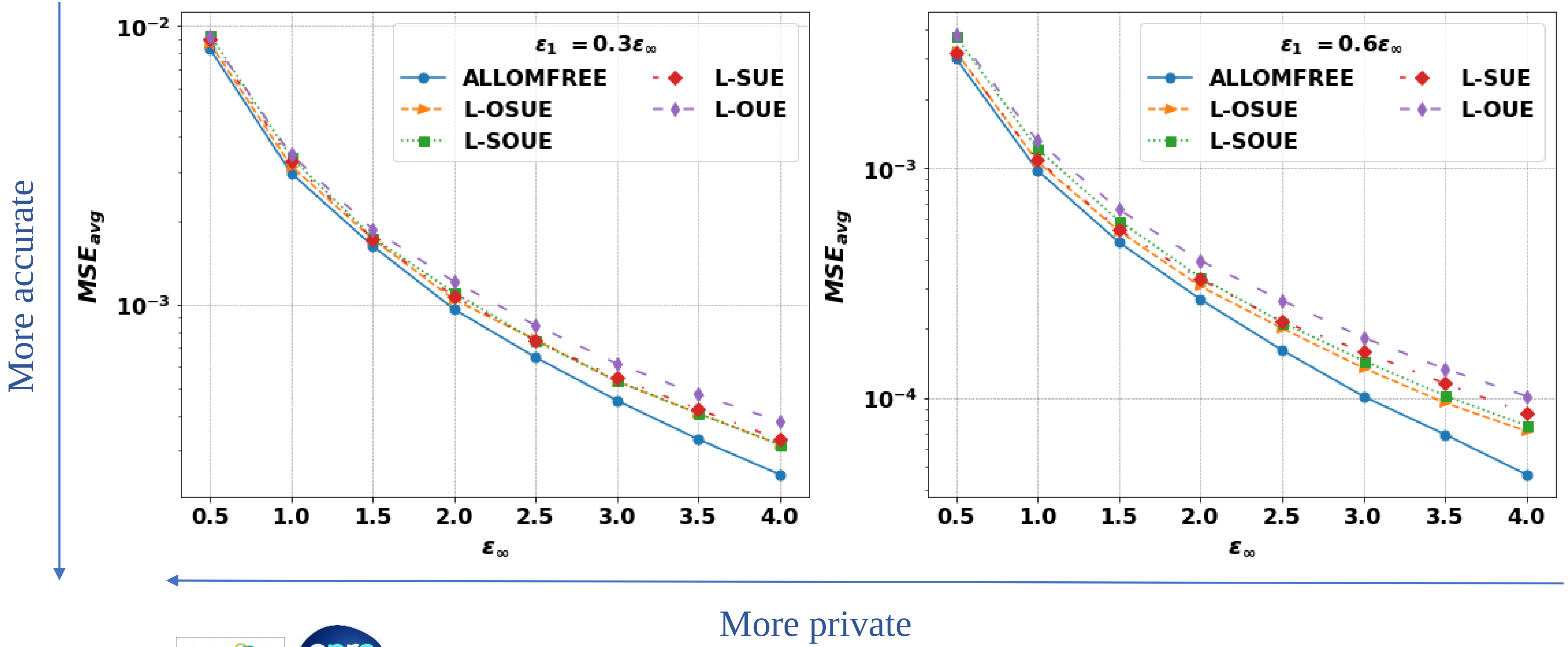
Experiments

- Dataset:
  - Census-Income*: $n = 299285$, $d = 33$, $\mathbf{k} = [9,52,47,17, \dots, 3,3,2]$
- Evaluation: $\epsilon_\infty = [0.5, 1, \dots, 3.5, 4]$ with $\epsilon_1 = \{0.3\epsilon_\infty,\ 0.6\epsilon_\infty\}$.
- Methods:
  - Smp: L-SUE, L-OUE, L-OSUE, L-SOUE;
  - ALLOMFREE (i.e., L-GRR or L-OSUE).
- Metric: Averaged MSE with $\tau = 1$ (a single collection),

$$\mathrm{MSE}_{avg} = \frac{1}{\tau}\sum_{t\in[1,\tau]}\frac{1}{d}\sum_{j\in[1,d]}\frac{1}{|A_j|}\sum_{v_i\in A_j}\left(f(v_i) - \hat{f}(v_i)\right)^2.$$

* Dheeru Dua and Casey Graff. 2017. UCI Machine Learning Repository: http://archive.ics.uci.edu/ml/index.php

Experimental Results on Census Dataset



Outline

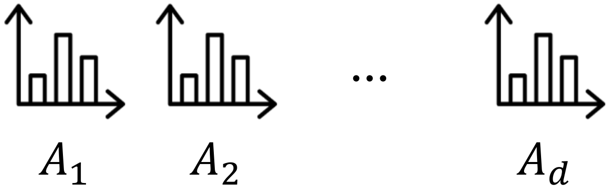
1. Introduction
2. Multiple Frequency Estimates Under Local Differential Privacy
   i. Longitudinal and Multidimensional Data Collection
   ii. **Multidimensional Data Collection**
3. Privacy-Utility Trade-off of Differentially Private Machine Learning Models
4. Further Contributions
5. Conclusion & Perspectives

Problem Statement: Statistical Learning

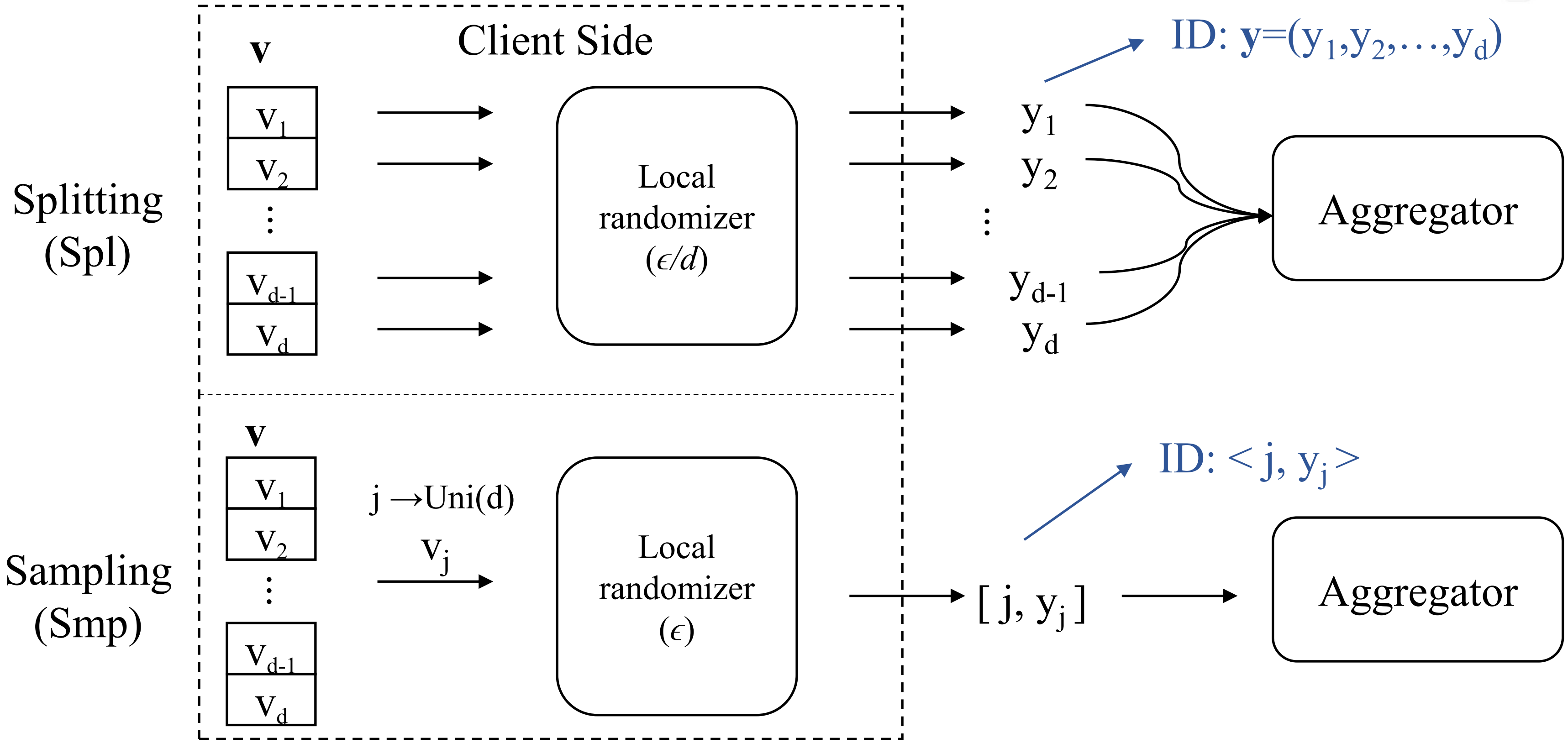
- **Tackled Issue:** Collecting *multidimensional* data under ϵ-*LDP* for *frequency estimation*.

- **More formally (notation):**

  - d attributes $A = \{A_1, A_2, \ldots, A_d\}$; **Multiple attributes**
  - Each attribute A_j has a discrete domain of size $|A_j| = k_j$;
  - Each user u_i for $1 \leq i \leq n$ has a tuple $\mathbf{v}^i = (v_1^i, v_2^i, \ldots, v_d^i)$;
  - **Analyzer:** estimate a k_j-bins histogram for each attribute $j \in [1, d]$.

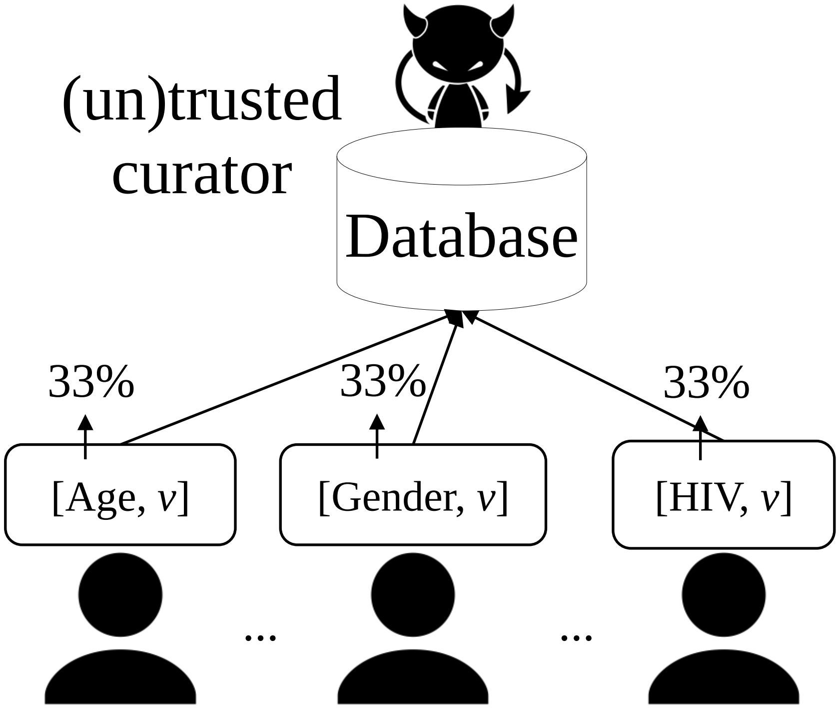


State-of-the-Art for Multiple Attributes*, **



[*] Nguyên, T.T., Xiao, X., Yang, Y., Hui, S.C., Shin, H., Shin, J. Collecting and analyzing data from smart device users with local differential privacy. In: arXiv:1606.05053 (2016).
[**] Wang, T., Blocki, J., Li, N. and Jha, S. Locally differentially private protocols for frequency estimation. In: USENIX Security Symposium (2017).

Why not *Smp*?



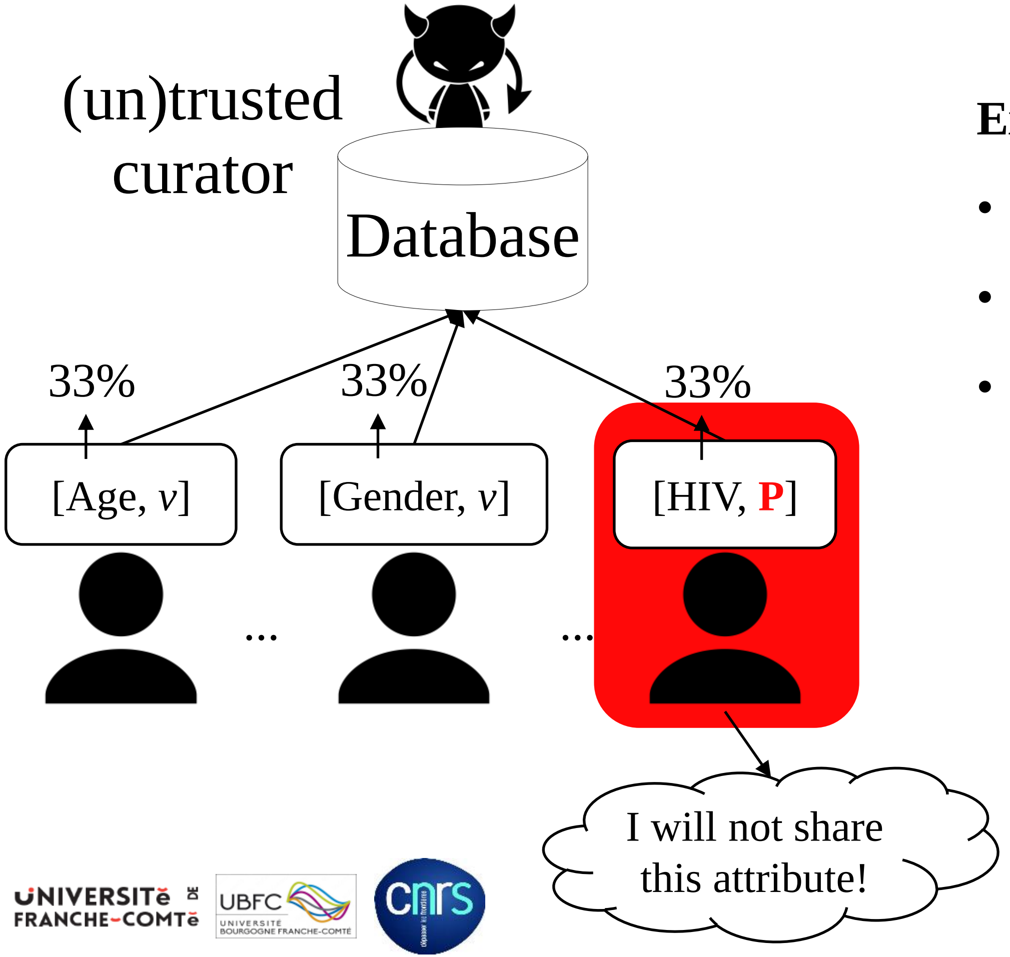
Example:

GRR for attributes with small domain
OUE otherwise

- *Smp*[ADP] → (attribute, ϵ-LDP value)
- Application scenario: health data
- $\epsilon = 2$, $d = 3$ attributes: age ($k_1 = [1, \ldots, 100]$), gender ($k_2 = [\mathrm{M}, \mathrm{F}]$), and HIV ($k_3 = [\mathrm{P}, \mathrm{N}]$).

Why not *Smp*?

All attributes have equal 'weight' in terms of privacy.



GRR for attributes with small domain
OUE otherwise

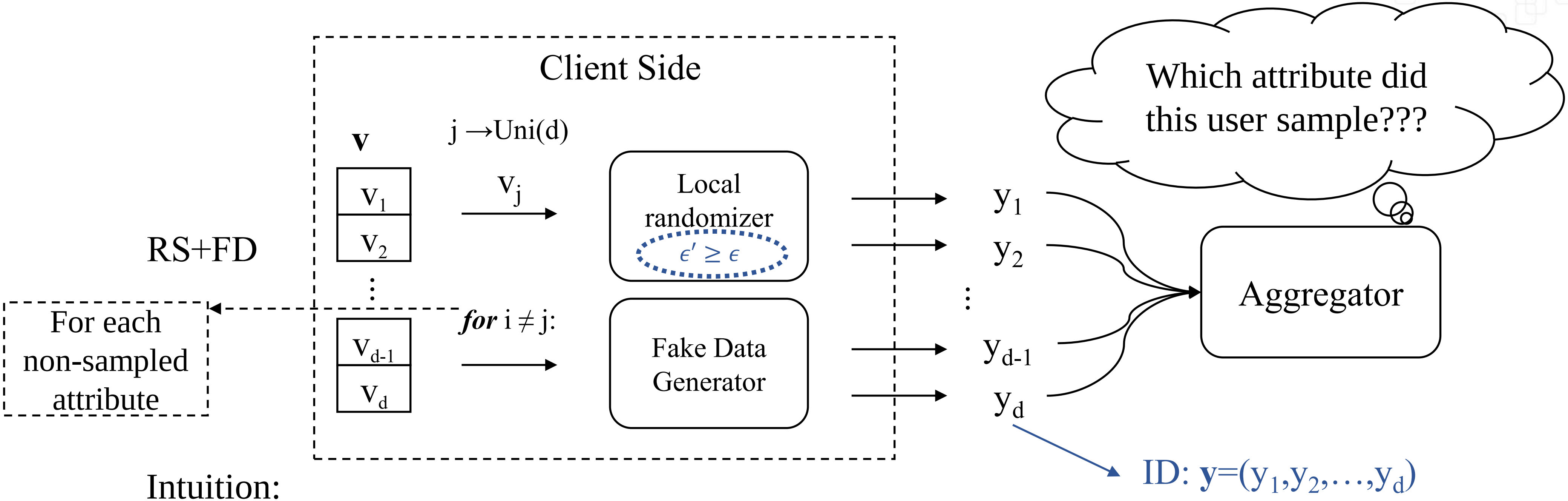
Example:

- *Smp*[ADP] → (attribute, ϵ-LDP value)
- Application scenario: health data
- $\epsilon = 2$, $d = 3$ attributes: age ($k_1 = [1, \ldots, 100]$), gender ($k_2 = [\mathrm{M}, \mathrm{F}]$), and HIV ($k_3 = [\mathrm{P}, \mathrm{N}]$).

$$p_{grr} = \frac{e^{\epsilon}}{e^{\epsilon} + k_j - 1} \approx 0.88 \text{ (probability of 'being honest')}$$

$$q_{grr} = \frac{1 - p_{grr}}{k_j - 1} \approx 0.12 \text{ (probability of 'lying')}$$

RS+FD: Random Sampling + Fake Data



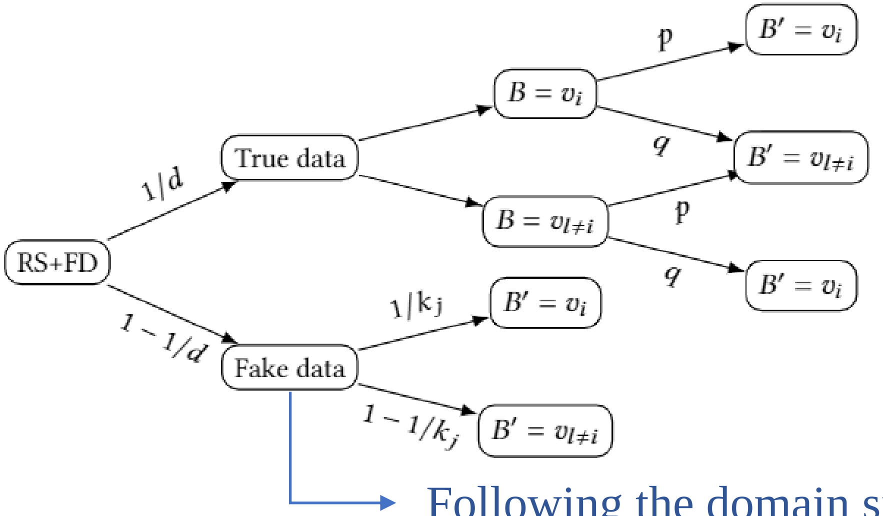
Intuition:

- RS+FD introduces **uncertainty** in the view of the aggregator.
- **Sampling result is not disclosed**, what is the impact in terms of privacy[*]**?**

[*] Li, N., Qardaji, W., Su, D. On sampling, anonymization, and differential privacy or, k-anonymization meets differential privacy. In: ASIACCS'12 (2012).

RS+FD with GRR

Client-Side of RS+FD[GRR]:



Following the domain size

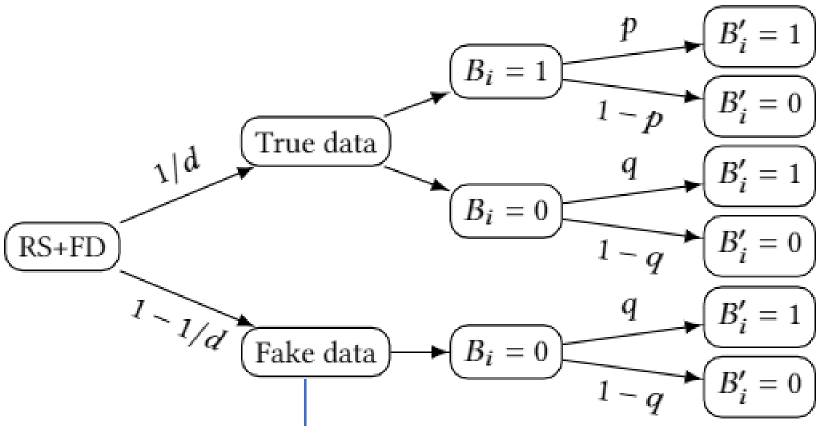
Aggregator → For each attribute $j \in [1, d]$, estimate:

$$\hat{f}(v_i) = \frac{N_i d k_j - n(d - 1 + q k_j)}{n k_j (p - q)}$$

Unbiased estimation and variance development in the manuscript

RS+FD with OUE

Client-Side of RS+FD[OUE-z]:



OUE applied to zero-vectors → $[0, 0, \dots, 0, 0]$

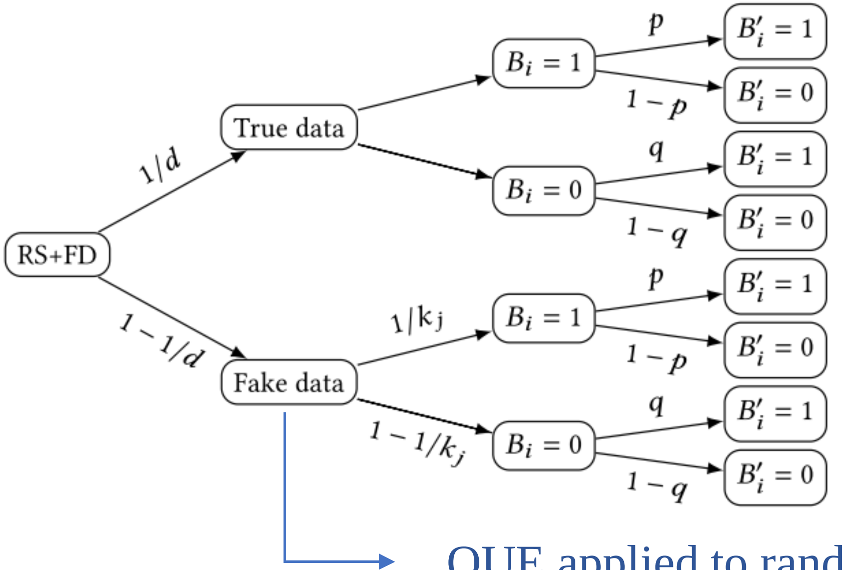
Aggregator → For each attribute $j \in [1, d]$, estimate:

$$\hat{f}(v_i) = \frac{d(N_i - nq)}{n\,(p - q)}$$

Unbiased estimation and variance development in the manuscript

RS+FD with OUE

Client-Side of RS+FD[OUE-r]:



OUE applied to random unary-encoded vectors

Aggregator → For each attribute $j \in [1, d]$, estimate:

$$\hat{f}(v_i) = \frac{N_i d k_j - n\left[q k_j + (p-q)(d-1) + q k_j (d-1)\right]}{n k_j (p-q)}$$

Unbiased estimation and variance development in the manuscript

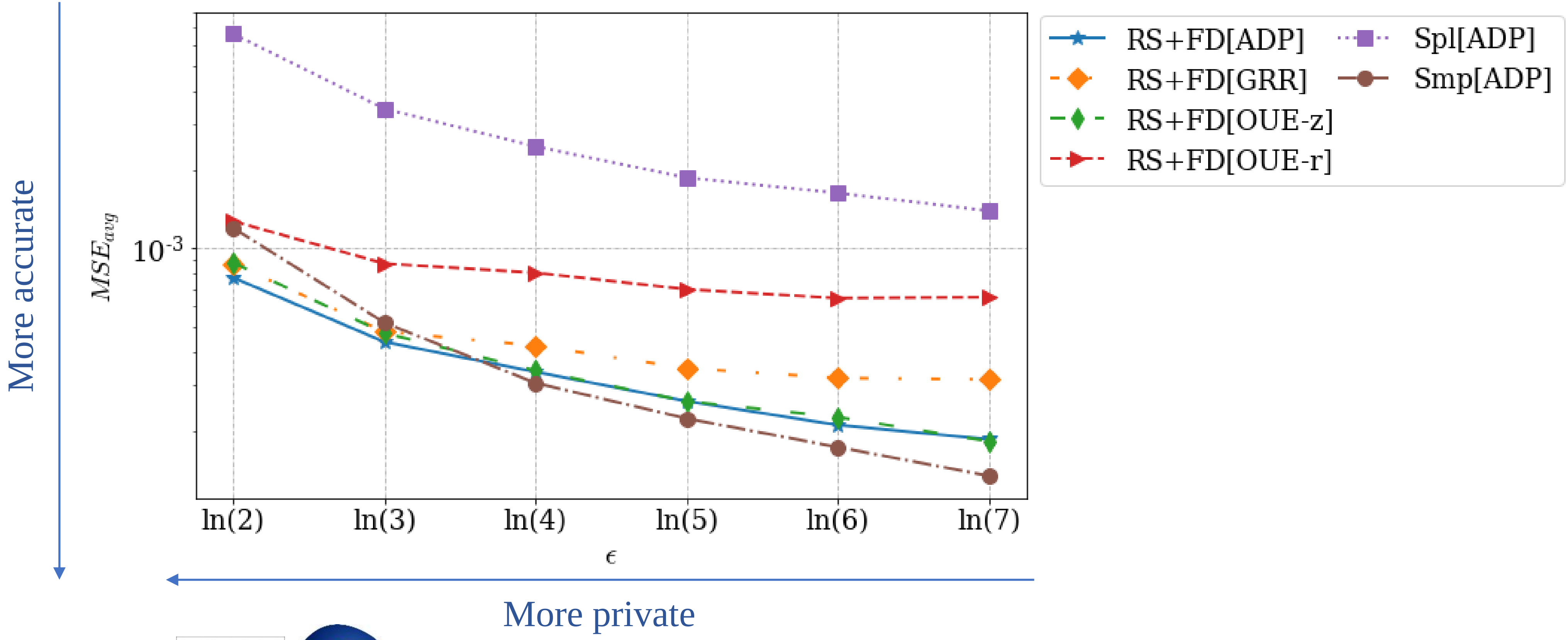
Experiments

- Dataset:
  - Census-Income*: $n = 299285$, $d = 33$, $\mathbf{k} = [9,52,47,17,\ldots,3,3,2]$
- Evaluation: ϵ = [ln(2), ln(3), ..., ln(7)].
- Methods:
  - Spl: ADP (i.e., either GRR or OUE);
  - Smp: ADP;
  - RS+FD: GRR, OUE-z, OUE-r, and ADP (i.e., either GRR or OUE-z).
- Metric: Averaged MSE,

$$\mathrm{MSE}_{avg} = \frac{1}{d}\sum_{j\in[1,d]}\frac{1}{|A_j|}\sum_{v_i\in A_j}\left(f(v_i)-\hat{f}(v_i)\right)^2.$$

* Dheeru Dua and Casey Graff. 2017. UCI Machine Learning Repository: http://archive.ics.uci.edu/ml/index.php

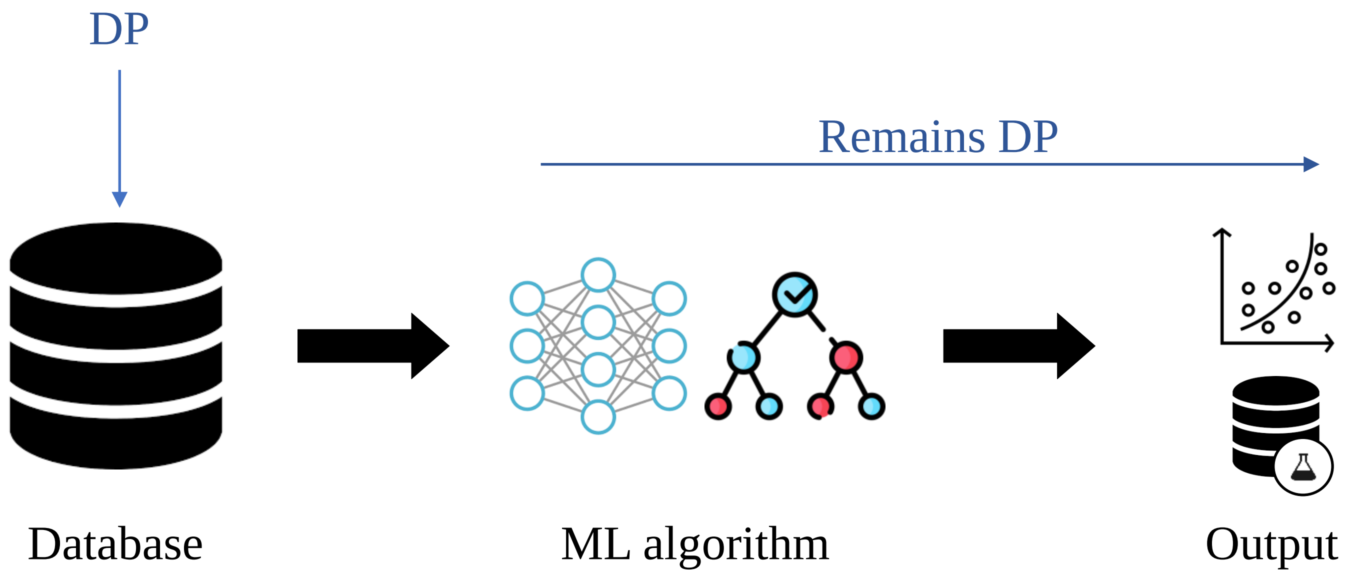
Experimental Results on Census Dataset



Outline

Problem Statement: Machine Learning

- **Tackled Issue:** Evaluation of the privacy-utility trade-off of training machine learning algorithms over differentially private data.
- **Motivation:** ML models are also succeptible to privacy attacks[*, **].



[*] Shokri, R., Stronati, M., Song, C., Shmatikov, V. Membership inference attacks against machine learning models. In: IEEE S&P (2017).
[**] Song, C., Ristenpart, T., Shmatikov, V. Machine learning models that remember too much. In: ACM SIGSAC (2017).

Outline

Aggregated Firemen Operation: Open Data*

YEAR	WEEK	CITY	REASON	NB_OPE
2018	10	AUVERS-SAINT-GEORGES	AID_TO_PEOPLE	4
2018	34	BROUY	AID_TO_PEOPLE	1
2018	35	BOUTIGNY-SUR-ESSONNE	AID_TO_PEOPLE	3
2018	32	ITTEVILLE	AID_TO_PEOPLE	1
2018	5	GUILLERVAL	AID_TO_PEOPLE	1

YEAR_MONTH	ZIP_CODE	CITY	AID_TO_PEOPLE
2008-4	71232	HAUTEFOND	1.0
2013-6	71450	ST MARTIN DE COMMUNE	0.0
2010-10	71469	ST PIERRE LE VIEUX	1.0
2009-5	71520	SEVREY	1.0
2013-7	71016	AZE	3.0

Brouy

Commune in France

Brouy is a commune in the Essonne department in Île-de-France in northern France. Inhabitants of Brouy are known as Brogaçois. Wikipedia

Area: 8.39 km²

Population: 144 (2015) INSEE

Generic Time ?

Generic Location ?

Generic Reason/Type

Hautefond

Commune in France

Hautefond is a commune in the Saône-et-Loire department in the region of Bourgogne-Franche-Comté in eastern France. Wikipedia

Area: 13.62 km²

Weather: 13°C, Wind S at 8 km/h, 72% Humidity weather.com

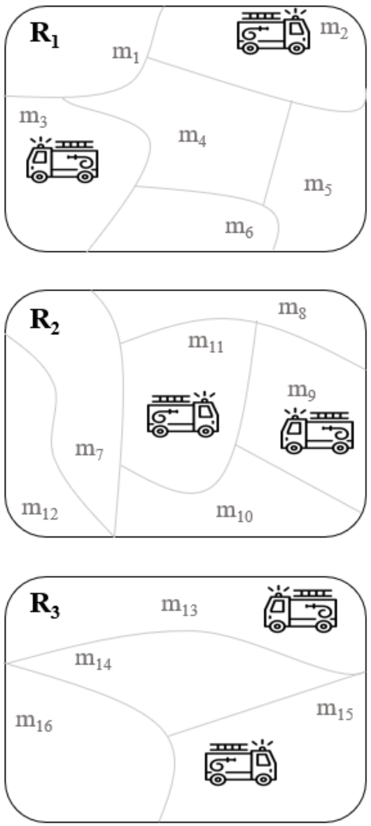
Population: 213 (2015) INSEE

Target: Multivariate Operational Demand Forecast

* Open platform for French public data: https://www.data.gouv.fr/en/

Our Solution: Generalization + DP

Agglomeration of small cities to larger regions



Sanitization

Interv. ID	***SDate*** (YYYY-MM-DD HH:MM:SS)	**Region**	**ϵ-LDP**
ID_1	2006-01-01 ~~00:05:23~~	R_1	[0, 1, 0]
ID_2	2006-01-01 ~~00:15:55~~	R_1	[1, 0, 1]
ID_3	2006-01-01 ~~01:24:12~~	R_2	[1, 1, 0]
...	...	...	...
...	...	...	...
ID_n-2	2018-12-31 ~~23:30:25~~	R_2	[1, 1, 1]
ID_n-1	2018-12-31 ~~23:45:23~~	R_3	[0, 1, 1]
ID_n	2018-12-31 ~~23:59:30~~	R_3	[0, 0, 0]

Daily Generalization

Sanitized Region

Sanitized Dataset

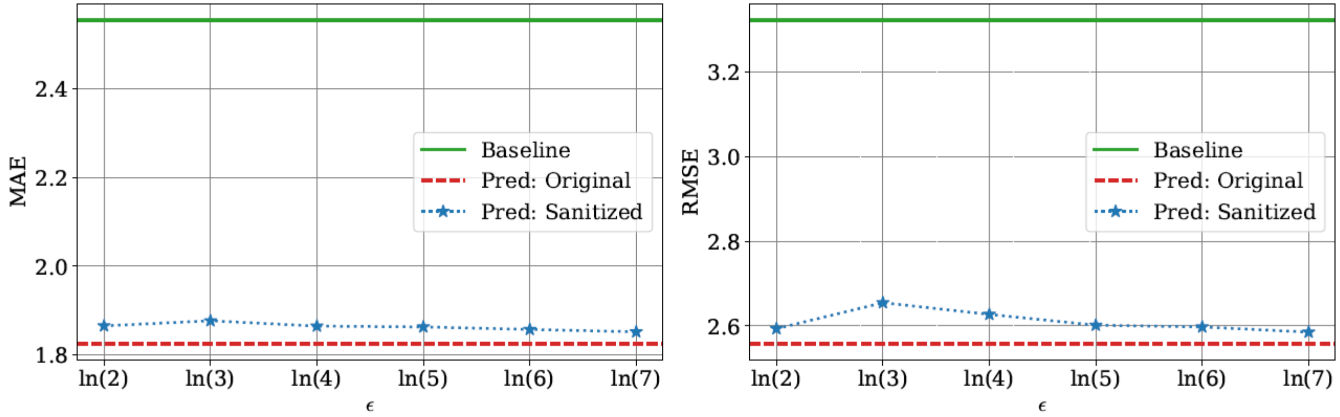
GRR, SUE, OUE, ...

The data analyst can aggregate by any period s/he wishes (e.g., 1-day, 3-days, 1-week, 1-month, ...

Dataset: Intervention's history of SDIS 25

Impact on Predictions of Daily Demand

- **Target:** Number of operations per day and per region.
- **Metrics:** Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE);
- **ML technique:** eXtreme Gradient Boosting (XGBoost).
- **Methods:** Baseline (average per day of the week), XGBoost trained over original and sanitized data.



Outline

Firemen Operation: Open Data*

Date/Time	Incident #	Level	Units	Location	Type
12/23/2021 4:28:37 AM	F210141750	1	M17	3900 7th Ave Ne	Medic Response
12/23/2021 4:27:22 AM	F210141748	1	A5	607 3rd Ave	Aid Response
12/23/2021 4:28:37 AM	F210141750	1	E17	3900 7th Ave Ne	Medic Response
12/23/2021 4:10:09 AM	F210141747	1	E31	2140 N Northgate Way	Aid Response
12/23/2021 3:50:06 AM	F210141743	1	M28	6900 37th Ave S	Medic Response

Precise Time
Precise Location
Generic Reason/Type

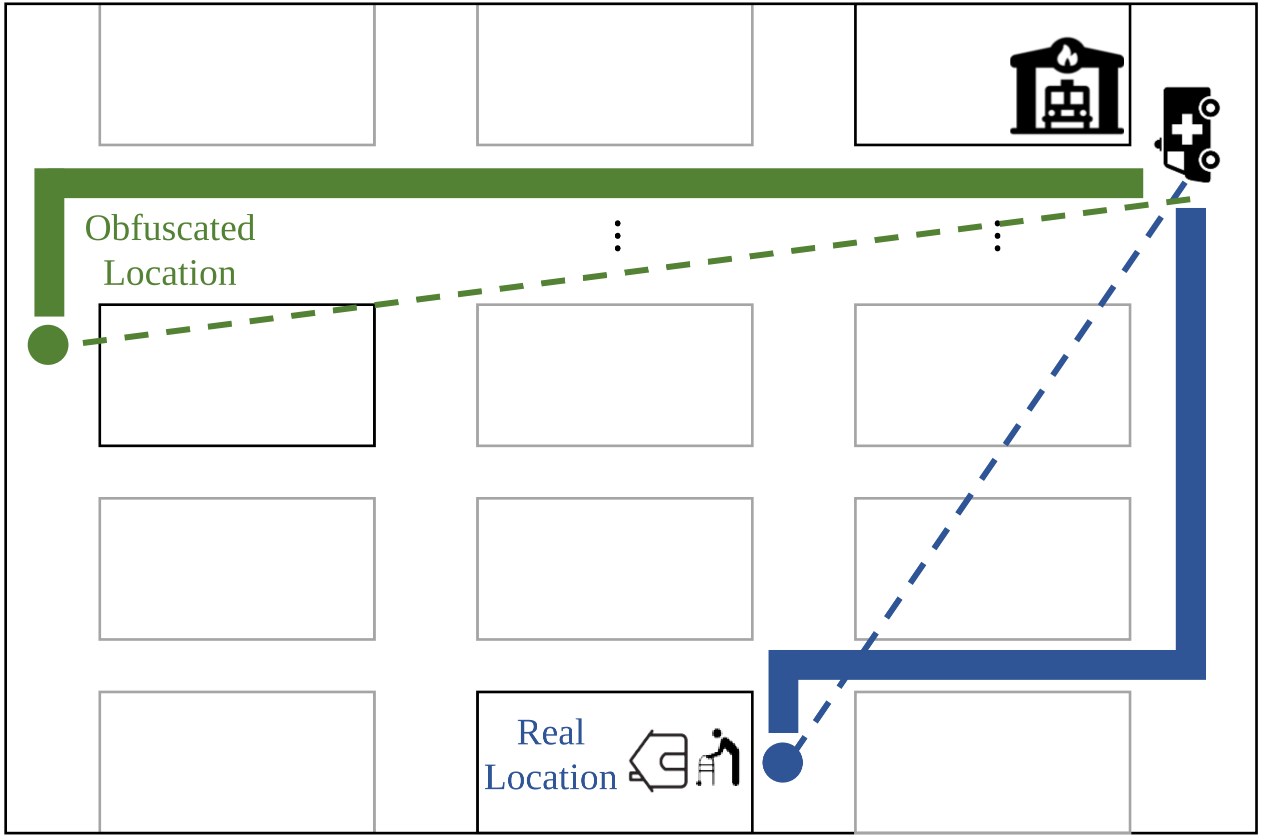
With both locations: Fire brigade and intervention

Target: Predict ambulance response time (ART)

Time measured from the call until an ambulance arrives at the emergency scene.

* Seattle Fire Department: http://www2.seattle.gov/fire/realtime911/

Need a Precise Location to Predict ART?



Obfuscation of emergency location data (i.e., latitude & longitude) using Planar Laplace Mechanism*;

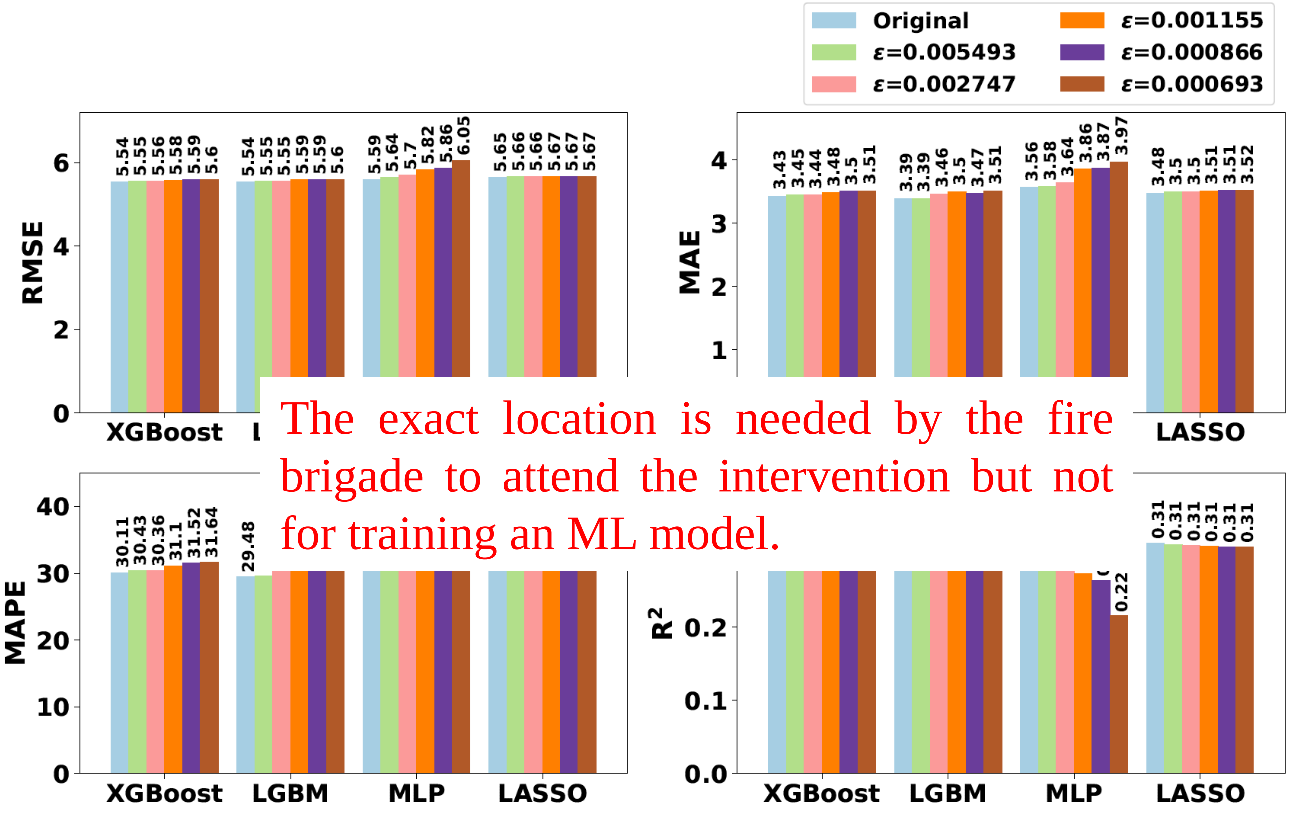
Additional perturbation:

- Estimated travel time;
- Estimated travel distance;
- Euclidean distance;
- Neighborhood, city, zone;
- …

Dataset: Departure's history of SDIS 25 ambulances

* Andrés, M.E., Bordenabe, N.E., Chatzikokolakis, K., Palamidessi, C. Geo-indistinguishability: Differential privacy for location-based systems. In: SIGSAC (2013).

Impact on Predictions of ART



The exact location is needed by the fire brigade to attend the intervention but not for training an ML model.

Metrics:

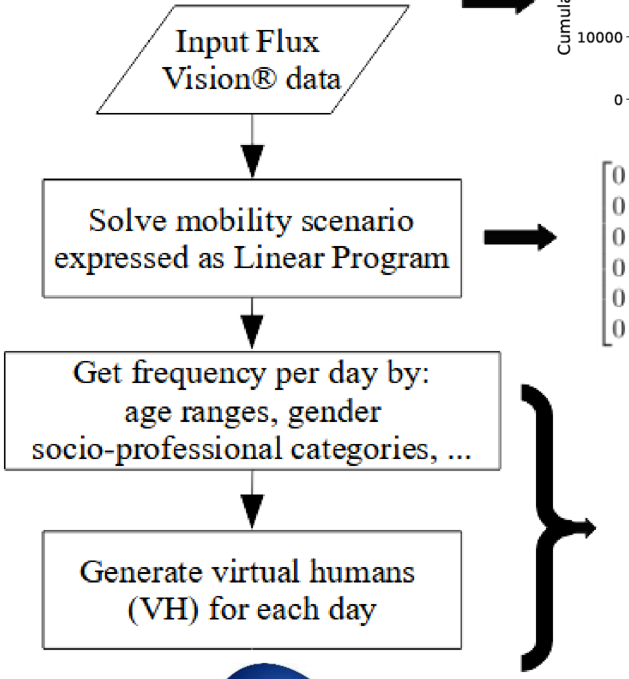
- Root Mean Squared Error (RMSE)
- Mean Absolute Error (MAE)
- Mean Absolute Percentage Error (MAPE)
- Coefficient of determination (R^2)

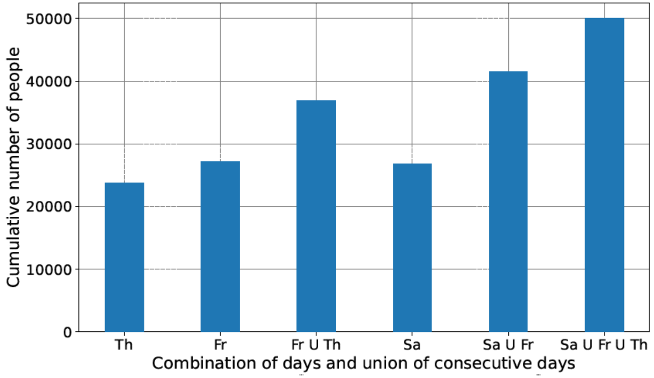
ML Techniques:

- eXtreme Gradient Boosting (XGBoost)
- Light Gradient Boosted Machine (LGBM)
- Multilayer Perceptron (MLP)
- Least Absolute Shrinkage and Selection Operator (LASSO)

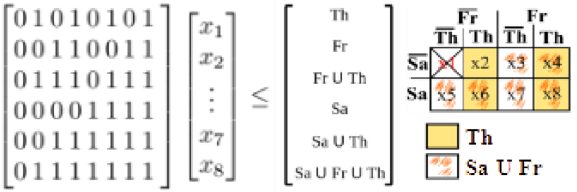
Outline

Providing Synthetic Data for Mobility

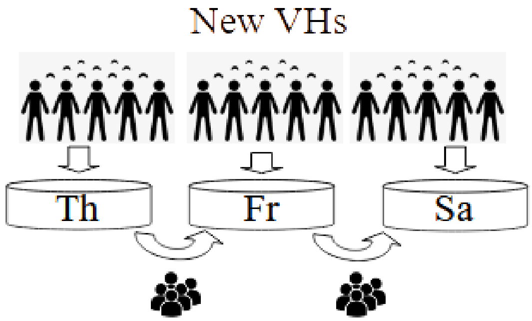




Multiple Attributes: Gender, Age-ranges, Sleeping Area, ...



Solves for *Nb* days: $2^{Nb} - 1$ combinations of day intersections.



Open Dataset: Mobility Scenario FIMU*

MS-FIMU→ Longitudinal and Multidimensional Dataset of Categorical Attributes:

- $d = 7$ attributes; $n = 88{,}935$ unique users; $Nb = 7$ days;
- Averaged Mean Relative Error $\approx 8\%$

Person ID	Name	Gender	Age	...	Visitor category	Region
91	Adrien Clement	M	45-54	...	French tourist	Alsace
32947	Grégoire Didier	M	25-34	...	French tourist	Franche-Comté
53990	Marie Le Lemaitre	F	25-34	...	Resident	Franche-Comté
58664	Michelle-Céline Marion	F	25-34	...	Resident	Franche-Comté

Date ID	Date
1	2017-05-31
2	2017-06-01
...	...
7	2017-06-06

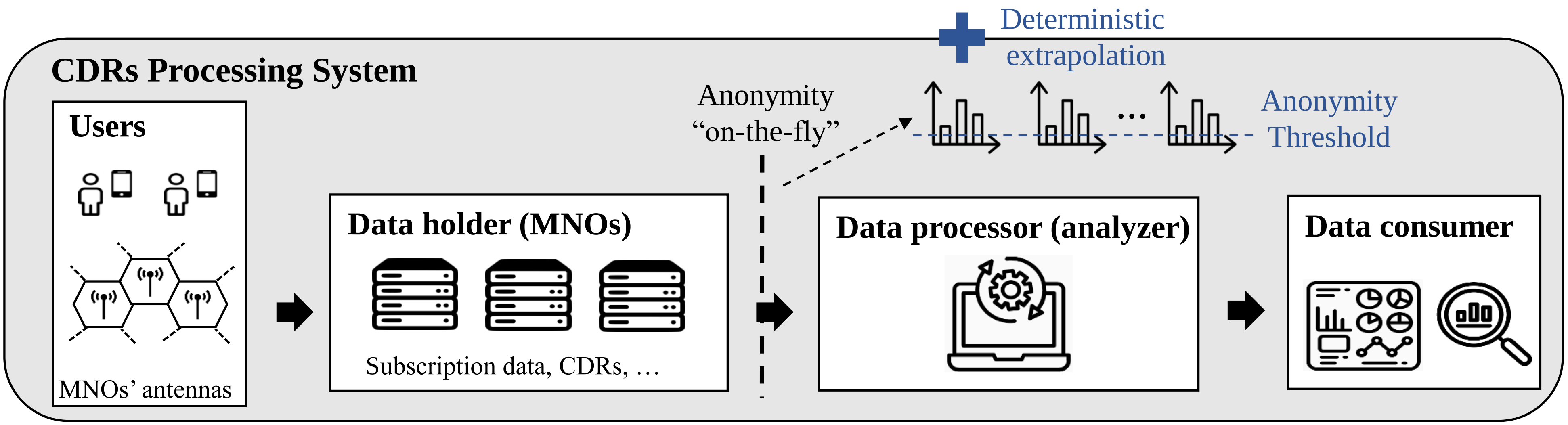
Index	Person ID	Date ID	Visit Duration
1	5385	2	6h
2	234	5	4h

* Open Dataset: MS-FIMU: https://github.com/hharcolezi/OpenMSFIMU

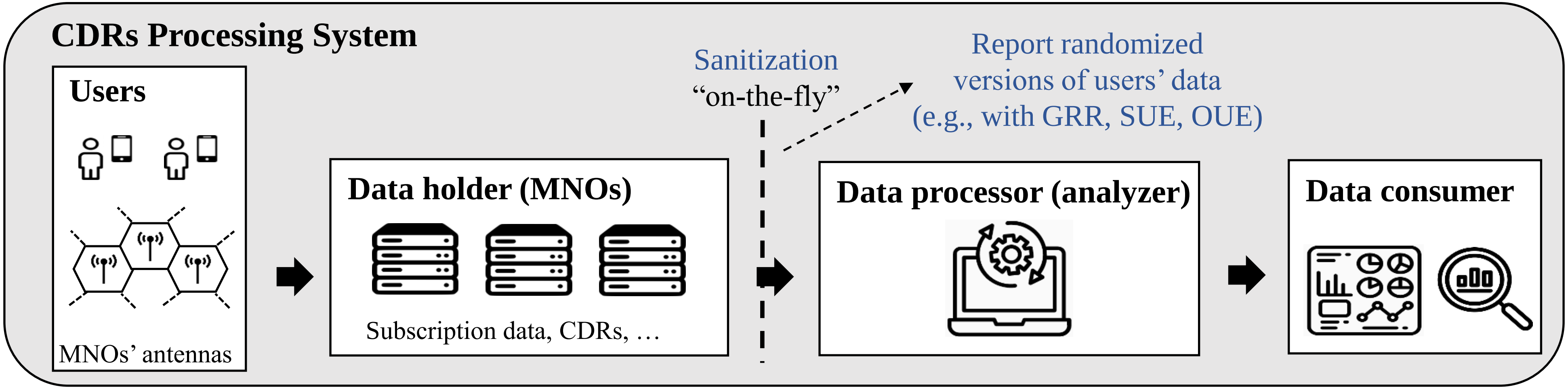
Outline

1. Introduction
2. Multiple Frequency Estimates Under Local Differential Privacy
3. Privacy-Utility Trade-off of Differentially Private Machine Learning Models
4. Further Contributions
    i. Generating Synthetic Data
    ii. **Privacy-Preserving Mobility Reports**
5. Conclusion & Perspectives

Current Anonymity-Based Mobility Reports



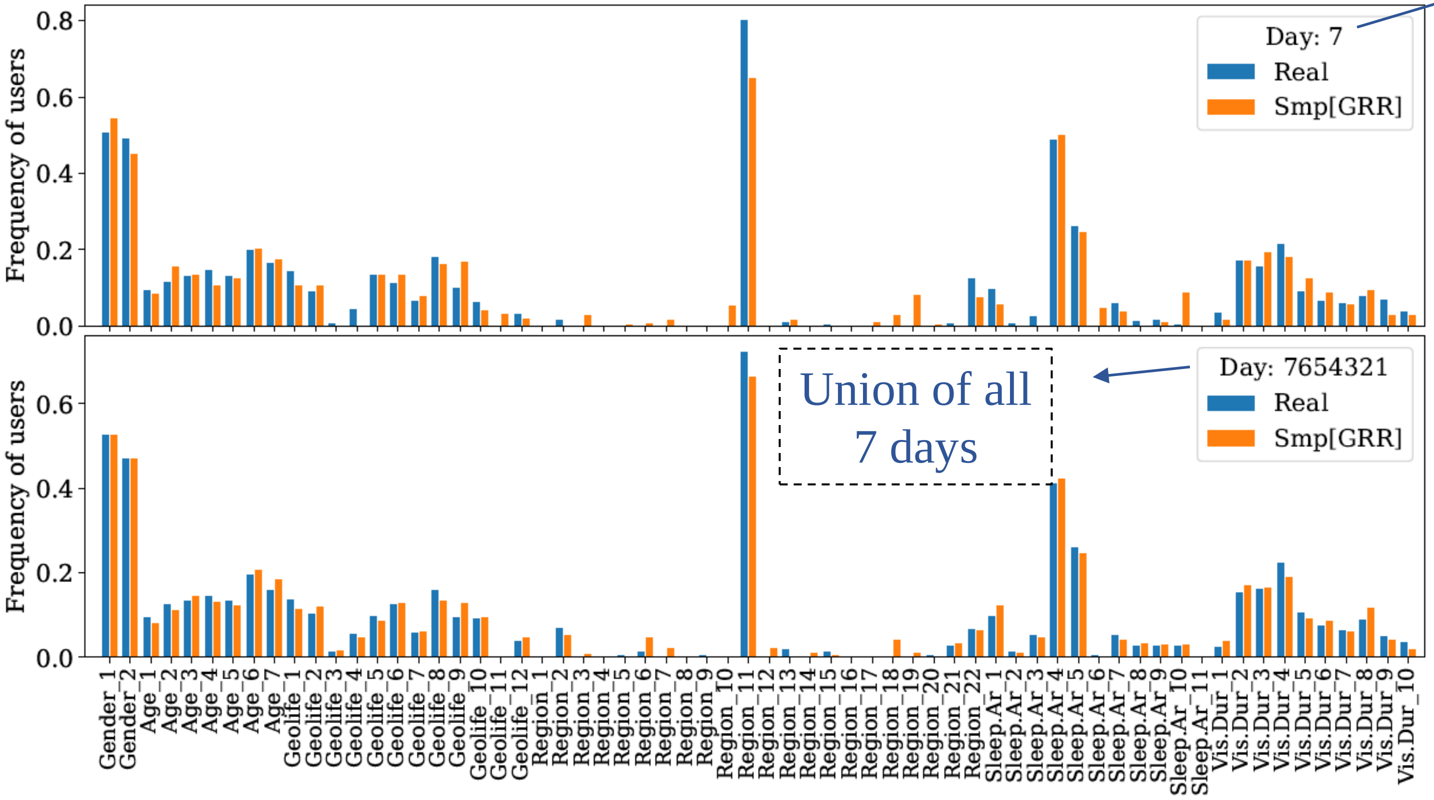
Proposed LDP-Based Mobility Reports



- Advantage: This scenario considers a *strong adversary* and *strong restrictions* for MNOs.
- Issue: The use of local randomizers can lead to great loss of utility.

LDP-Based Mobility Reports

A single day



Dataset:

- MS-FIMU

Method:

- Smp[GRR];

Privacy bugdet:

- $\epsilon = 1$

Outline

Conclusion & Perspectives

General Conclusion:

- We published an open dataset MS-FIMU of categorical attributes based on real-world mobility analytics (longitudinal and multidimensional);
- We proposed a CDRs processing system with DP guarantees at the user level for human mobility analytics;
- We optimized the utility of LDP protocols (i.e., L-GRR and L-OSUE) for longitudinal frequency estimates through memoization with theoretical proofs;
- We improved utility and privacy in multiple frequency estimates under LDP through generic frameworks (i.e., ALLOMFREE and RS+FD);
- We empirically evaluated the privacy-utility trade-off of differentially private machine learning models on real-world datasets/tasks.

Conclusion & Perspectives

Perspectives:

- Improve RS+FD with realistic fake data;
- Design more enhanced post-processing methods (e.g., Expectation-Maximization algorithm) for ALLOMFREE and RS+FD;
- Cast other LDP protocols into RS+FD, including longitudinal ones;
- Evaluate performance VS privacy protection of ALLOMFREE and RS+FD on generating synthetic data for ML classification/regression tasks;

- Attack RS+FD, i.e., try to correctly guess the sampled attribute of each user;
- Evaluate the privacy-utility trade-off of differentially private ML models against attacks (e.g., membership inference attacks).
- Build a python library for multiple frequency estimates under LDP.

Thank you for your attention!

Héber HWANG ARCOLEZI

heber.hwang_arcolezi@univ-fcomte.fr