

Differential Privacy

In a Nutshell

M2 SIF - SED

Tristan Allard

Univ. Rennes / Irisa lab.

`tristan.allard@irisa.fr`

Autumn 2023

Progress of the Talk

Introduction

Reminder : Partition-Based Models

Differential Privacy

Conclusion

References

Differential Privacy and Privacy-Preserving Data Publishing

Privacy-Preserving Data Publishing (PPDP) :

- ▶ Publish *personal data* for analysis purposes (accurate aggregate queries) . . .
- ▶ . . . while preserving individuals' *privacy* (uncertain point queries)
- ▶ Also called *sanitization*

Differential privacy is one way to perform sanitization.

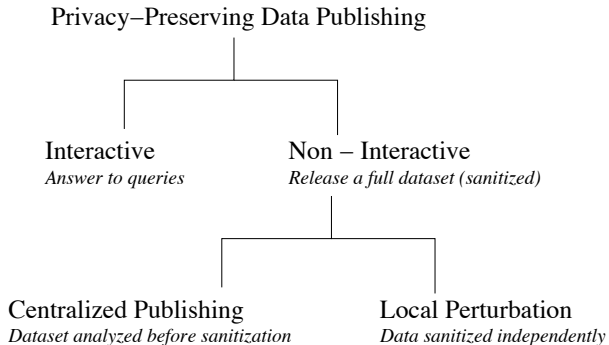
Components of a Privacy-preserving Data Publishing Solution (reminder)

Three components:

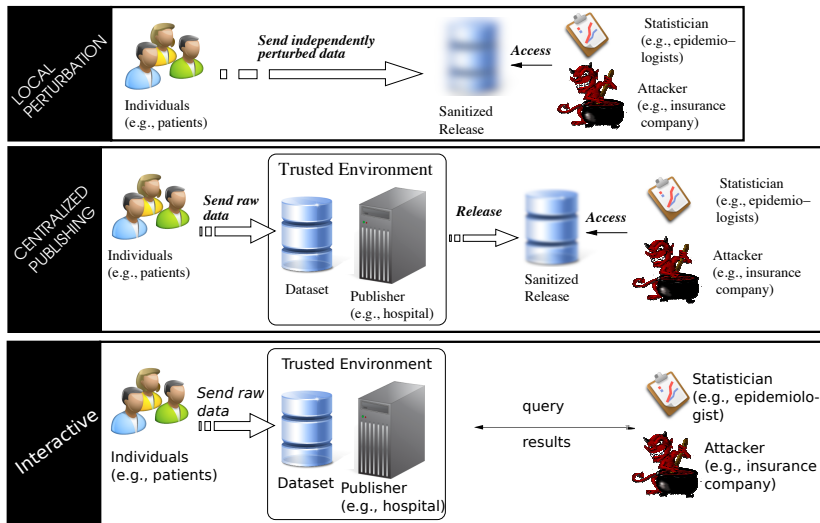
1. **Privacy model:** What does it mean for the data released to be privacy-preserving?
2. **Privacy mechanism:** How to produce the privacy-preserving data to be released?
3. **Utility metric:** How much useful is the released data?

Pseudonymity does not work. . . Which component(s) does it miss ?

Variations on a Theme (reminder) I



Variations on a Theme (reminder) II



This talk

This talk : introduction to the *differential privacy model*, a de facto standard today, and to its major mechanisms (no utility measures).

Warning : research about sanitization is still very active !

Please, go deeper !

- ▶ A vast litterature : big-bang of the PPDP works in computer science : early/mid 2000's
- ▶ For a good survey : Chen et al in *Foundations and Trends in Databases* 2009, see [1] ;
Note that : (1) other surveys exist and (2) 2009 is far.
- ▶ For more questions : email me

Progress of the Talk

Introduction

Reminder : Partition-Based Models

Differential Privacy

Conclusion

References

The k -ANONYMITY Model

A release is k -anonymous [7] if:

- ▶ it does not contain any identifying attribute ;
- ▶ any QI is indistinguishable from at least $(k - 1)$ others ;

A group of records indistinguishable wrt their QI is called an *equivalence class*.

Beyond k -ANONYMITY: the BAYES-OPTIMAL PRIVACY Attempt

Founding intuition

Background knowledge about SD should be **expressed** and **taken into account** by the privacy model.

The BAYES-OPTIMAL PRIVACY model [6] is an early attempt to this end (2006):

- ▶ **Background knowledge:** joint distribution between QI and SD
- ▶ **Prior belief:** given a targeted QI q and a SD s , probability of s given q
- ▶ **Posterior belief:** given a targeted QI q , a SD s , **and the sanitized release** \mathcal{V} , probability of s given q and \mathcal{V}
- ▶ **Privacy breach:** if $distance(posterior\ belief, prior\ belief) > \theta$ (too much gain in knowledge)

BAYES-OPTIMAL PRIVACY : Impractical

If BAYES-OPTIMAL PRIVACY were practical, it could permit to check that releases do not allow significant knowledge gains. . .

But :

- ▶ Obtaining the joint distribution f that represents the adversarial background knowledge ?
- ▶ What if there are several adversaries ?
- ▶ What about other kinds of knowledge ?
- ▶ Cost of checking all the possible (q, s) pair !

l -DIVERSITY (reminder)

l -DIVERSITY: a simple and easy-to-check condition for protecting against **SD homogeneity** and **adversarial negation statements**.

l -DIVERSITY [6]

An l -diverse equivalence class contains at least l *well-represented* sensitive values.

(See the previous CM for precise definitions of “well-represented”.)

Paradigm#1 : the Uninformative Principle

At the heart of I -DIVERSITY and followers, there is a vision of privacy : the uninformative principle.

Paradigm #1 : Uninformative Principle [6]

A privacy breach occurs when the *prior belief* of the adversary differs *significantly* from his *posterior belief* (i.e., below a user-defined threshold).

*“If the **release of the statistics S** make it possible to determine the value D_k **more accurately** than is possible **without access to S**, disclosure has taken place (...)”*

Dalenius 1977 [3]

Progress of the Talk

Introduction

Reminder : Partition-Based Models

Differential Privacy

Conclusion

References

Paradigm#2 : Differential Privacy

- ▶ Global trends are not private and must be learnt : there must be a knowledge gain !
- ▶ Privacy is about each individual value, i.e., **each individual contribution** to the global trend

Paradigm #2 : Differential Privacy

A function f satisfies differential privacy iif: the possible impact of any individual on its result (its possible outputs) is limited.



Paradigm#2 : Differential Privacy

- ▶ Global trends are not private and must be learnt : there must be a knowledge gain !
- ▶ Privacy is about each individual value, i.e., **each individual contribution** to the global trend

Paradigm #2 : Differential Privacy

A function f satisfies differential privacy iif: the possible impact of any individual on its result (its possible outputs) is limited.



Intuitions - Mechanism

- ▶ Differential privacy originally considers **aggregate queries** (counts, sums)...
- ▶ For ex : $q = \text{SELECT COUNT(*) FROM PATIENTS WHERE DIAGNOSIS LIKE 'FLU'}$
- ▶ How to hide the impact of any single individual participation to the aggregate result ?
 - ▶ Add random noise to the true result ! Answer $q(\mathcal{D}) + \text{noise}$
 - ▶ Such that the noise is **proportional to the participation of one individual**.
 - ▶ For ex : noise above should be proportionnal to the impact of one individual on q , *i.e.*, proportionnal to 1 !
 - ▶ What if q had been a sum of salaries ?

Intuitions - Aggregation is not enough I

A simple example :

- ▶ Private table : PATIENTS (AGE, ZIP, GENDER, DIAG)
- ▶ Background knowledge : Bill is a man, 26 years old, zipcode 12345, present in PATIENTS.
- ▶ Consider the following queries :
 - ▶ $q_1 = \text{SELECT COUNT(*) FROM PATIENTS WHERE (AGE = 26 AND ZIP=12345 AND GENDER='m')}$
 $\Rightarrow q_1(\mathcal{D}) = 1$
 - ▶ $q_2 = \text{SELECT COUNT(*) FROM PATIENTS WHERE DIAG LIKE 'FLU'}$
 $\Rightarrow q_2(\mathcal{D}) = 100$
 - ▶ $q_3 = \text{SELECT COUNT(*) FROM PATIENTS WHERE NOT (AGE = 26 AND ZIP=12345 AND GENDER='m')} \text{ AND DIAG LIKE 'FLU'}$
 $\Rightarrow q_3(\mathcal{D}) = 99$
- ▶ Does Bill have flu ?

Intuitions - Aggregation is not enough II

In general, with a sufficient number of COUNT queries, **and even with naive noise addition**, it is possible to build and solve a system of linear equations for obtaining the target values.

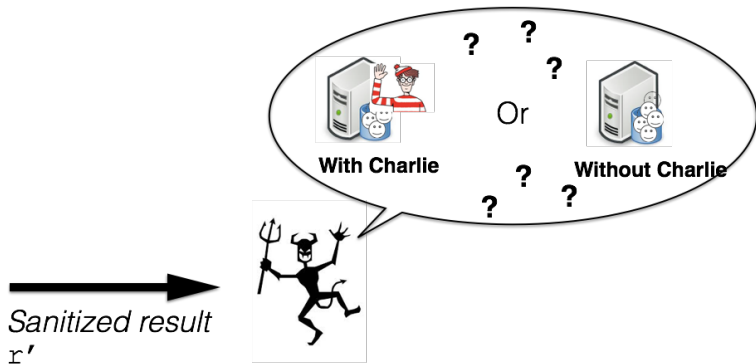
Real-life illustration¹

Assume a loan DB with a `clientId` column, a secret `loanStatus`, and an *exact* `count(·)` interface to the DB (perturbed answers might be vulnerable as well). Perform a sufficient number of queries as follows and solve the system.

```
SELECT count(clientId) FROM loans
WHERE clientId BETWEEN RANDOM1 and RANDOM2
AND loanStatus = 'C'
```

¹A real-life system <https://aircloak.com/> has been recently broken (2018) based on this kind of attack [2].

Intuitions



Intuitions

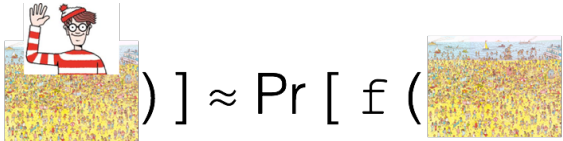
$$\Pr [\text{£} (\text{Charlie})] \approx \Pr [\text{£} (\text{No Charlie})]$$


Figure: Limited impact of any possible Charlie

Intuitions

$$\Pr [f (\text{[Image: Charlie in a crowd]})] \approx \Pr [f (\text{[Image: Crowd]})]$$

Close to an e^ϵ factor
(ϵ is the privacy parameter, set by DBA)

Figure: Limited impact of any possible Charlie

Initial Model

ϵ -differential privacy (from [4])

A **random function** f satisfies ϵ -differential privacy iff: **For all** \mathcal{D} and \mathcal{D}' **differing in at most one record**, and for any possible output \mathcal{S} of f , then it is true that:

$$\Pr[f(\mathcal{D}) = \mathcal{S}] \leq e^\epsilon \times \Pr[f(\mathcal{D}') = \mathcal{S}]$$

Initial Model

ϵ -differential privacy (from [4])

A **random function** f satisfies ϵ -differential privacy iff: **For all** \mathcal{D} and \mathcal{D}' **differing in at most one record**, and for any possible output \mathcal{S} of f , then it is true that:

$$\Pr[f(\mathcal{D}) = \mathcal{S}] \leq e^\epsilon \times \Pr[f(\mathcal{D}') = \mathcal{S}]$$

- ▶ f : here, an aggregate query perturbed by adding random noise to its output
- ▶ “For all \mathcal{D} and \mathcal{D}' ”: all possible datasets
- ▶ “ \mathcal{D} and \mathcal{D}' differing in at most one record”: here, \mathcal{D} is \mathcal{D}' with one tuple more or one tuple less (variant: one tuple with different values). Called *neighboring datasets*
- ▶ ϵ : the privacy parameter, public, common values: 0.01, 0.1, $\ln 2$, $\ln 3$
- ▶ $e^\epsilon \times \Pr[\dots]$: if one side is zero, the other must be zero too

Query Sensitivity

Different individuals, different impacts. . .



Query Sensitivity

Different individuals, different impacts. . .

- ▶ Presence/absence of an individual on the result of a COUNT: at worst +/- 1
- ▶ Presence/absence of an individual on the result of a SUM:
 $\max(|domain_{min}|, |domain_{max}|)$

Quantification of the worst-case impact of any possible individual on the output of a query g : called *query sensitivity*, and denoted S_g .

Query Sensitivity

Different individuals, different impacts. . .

- ▶ Presence/absence of an individual on the result of a COUNT: at worst +/- 1
- ▶ Presence/absence of an individual on the result of a SUM:
 $\max(|domain_{min}|, |domain_{max}|)$

Quantification of the worst-case impact of any possible individual on the output of a query g : called *query sensitivity*, and denoted S_g .

In general: $S_g = \max_{\mathcal{D}, \mathcal{D}'} \|g(\mathcal{D}) - g(\mathcal{D}')\|_1$ where \mathcal{D} and \mathcal{D}' are two neighboring datasets.

Laplace Mechanism for Real-Valued Interactive Queries

A - “Excellent, but how to achieve differential privacy ?”

B - “Just add random noise to each query output, he said !”

A - “But from which distribution ? Uniform ? Gaussian ? Gamma ? Poisson ? ... ? Any ?”

Laplace Mechanism for Real-Valued Interactive Queries

Given g and ϵ , adding a random variable sampled from a Laplace distribution with mean 0 and scale factor S_g/ϵ satisfies ϵ -differential privacy [5].

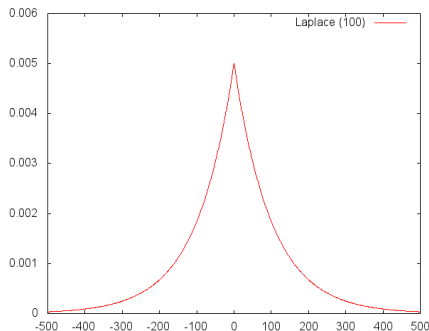


Figure: Laplace (0, 1/0.01)

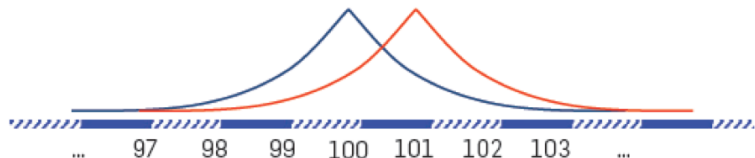
Laplace probability distribution function : $\Pr_{\text{Lap}(0,b)}(x) = \frac{e^{-|x|/b}}{2b}$

Laplace Mechanism for Real-Valued Interactive Queries

Given g and ϵ , adding a random variable sampled from a Laplace distribution with mean 0 and scale factor S_g/ϵ satisfies ϵ -differential privacy [5].

Assume that the COUNT when Bob participates to the dataset is $r = 101$:

- ▶ In red, distribution of perturbed outputs ($r' = r + n$) when Bob is in
- ▶ In blue, *idem* when Bob is out

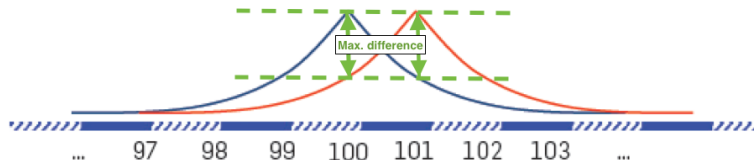


Laplace Mechanism for Real-Valued Interactive Queries

Given g and ϵ , adding a random variable sampled from a Laplace distribution with mean 0 and scale factor S_g/ϵ satisfies ϵ -differential privacy [5].

Assume that the COUNT when Bob participates to the dataset is $r = 101$:

- ▶ In red, distribution of perturbed outputs ($r' = r + n$) when Bob is in
- ▶ In blue, *idem* when Bob is out



Why does Adding Laplace Noise Work ?

Let n be the noise added to the exact result r of the function g computed over \mathcal{D} to satisfy DP : $r' = r + n$. The distribution from which n is sampled must hide the possible impact of the presence/absence of any individual.

Why is the Laplace distribution appropriate for satisfying DP ?

Why does Adding Laplace Noise Work ?

- ▶ With Bob in, the output is $r' = r + n$.
With Bob out, the output is $r' = (r - S_g) + n$

²**Homework** : Demonstrate that ! (play with Laplace probability distribution function)

Why does Adding Laplace Noise Work ?

- ▶ With Bob in, the output is $r' = r + n$.
With Bob out, the output is $r' = (r - S_g) + n$
- ▶ To satisfy DP, the probabilities of outputting the former or the latter must be *similar* (i.e., constraints of the DP model) :
Eq. 1 : $\Pr[r' = r + n] \leq e^\epsilon \cdot \Pr[r' = (r - S_g) + n]$
Eq. 2 : $\Pr[r' = (r - S_g) + n] \leq e^\epsilon \cdot \Pr[r' = r + n]$

²**Homework** : Demonstrate that ! (play with Laplace probability distribution function)

Why does Adding Laplace Noise Work ?

- ▶ With Bob in, the output is $r' = r + n$.
With Bob out, the output is $r' = (r - S_g) + n$
- ▶ To satisfy DP, the probabilities of outputting the former or the latter must be *similar* (i.e., constraints of the DP model) :
Eq. 1 : $\Pr[r' = r + n] \leq e^\epsilon \cdot \Pr[r' = (r - S_g) + n]$
Eq. 2 : $\Pr[r' = (r - S_g) + n] \leq e^\epsilon \cdot \Pr[r' = r + n]$
- ▶ **Eq. 1** : $\Pr[n = r' - r] \leq e^\epsilon \cdot \Pr[n - S_g = r' - r]$
Eq. 2 : ...

²**Homework** : Demonstrate that ! (play with Laplace probability distribution function)

Why does Adding Laplace Noise Work ?

- ▶ With Bob in, the output is $r' = r + n$.
With Bob out, the output is $r' = (r - S_g) + n$
- ▶ To satisfy DP, the probabilities of outputting the former or the latter must be *similar* (i.e., constraints of the DP model) :
 - Eq. 1** : $\Pr[r' = r + n] \leq e^\epsilon \cdot \Pr[r' = (r - S_g) + n]$
 - Eq. 2** : $\Pr[r' = (r - S_g) + n] \leq e^\epsilon \cdot \Pr[r' = r + n]$
- ▶ **Eq. 1** : $\Pr[n = r' - r] \leq e^\epsilon \cdot \Pr[n - S_g = r' - r]$
Eq. 2 : ...
- ▶ **Eq. 1** : $\Pr[n] \leq e^\epsilon \cdot \Pr[n - S_g]$
Eq. 2 : $\Pr[n - S_g] \leq e^\epsilon \cdot \Pr[n]$

²**Homework** : Demonstrate that ! (play with Laplace probability distribution function)

Why does Adding Laplace Noise Work ?

- ▶ With Bob in, the output is $r' = r + n$.
With Bob out, the output is $r' = (r - S_g) + n$
- ▶ To satisfy DP, the probabilities of outputting the former or the latter must be *similar* (i.e., constraints of the DP model) :
Eq. 1 : $\Pr[r' = r + n] \leq e^\epsilon \cdot \Pr[r' = (r - S_g) + n]$
Eq. 2 : $\Pr[r' = (r - S_g) + n] \leq e^\epsilon \cdot \Pr[r' = r + n]$
- ▶ **Eq. 1** : $\Pr[n = r' - r] \leq e^\epsilon \cdot \Pr[n - S_g = r' - r]$
Eq. 2 : ...
- ▶ **Eq. 1** : $\Pr[n] \leq e^\epsilon \cdot \Pr[n - S_g]$
Eq. 2 : $\Pr[n - S_g] \leq e^\epsilon \cdot \Pr[n]$
- ▶ Sampling n in $\text{Lap}(0, S_g/\epsilon)$ satisfies the two equations ! ²

²**Homework** : Demonstrate that ! (play with Laplace probability distribution function)

Differential Privacy Properties

- ▶ **Self-composability** : composing the outputs of two independent releases sanitized by differentially-private function(s) satisfies differential privacy :
 - ▶ Where $\epsilon_{final} = \sum \epsilon_i$ If input datasets are **not** disjoint
 - ▶ Or $\epsilon_{final} = \max \epsilon_i$ otherwise
- ▶ **No breach from post-processing** :
 - ▶ (*Laplace mechanism is independent from data*)
 - ▶ Any function applied to a differentially-private input produces a differentially-private output

A non exact statement hides in this slide, can you find it ?

Inherent Limits

- ▶ Noise distribution centered on 0 ...
 - ⇒ Sum of noises converges to 0 ...
 - ⇒ No unlimited number of queries !
- ▶ Composability properties ⇒ the privacy parameter ϵ can be seen as a **budget** that must be distributed over the queries to execute ($\epsilon_{final} = \sum \epsilon_i$)

Discussion

According to you:

- ▶ Against which adversary may differential privacy protect ?
- ▶ How could you set the value of ϵ ?
- ▶ Take your favorite analytical algorithm, how would you make it satisfy differential privacy ?

Progress of the Talk

Introduction

Reminder : Partition-Based Models

Differential Privacy

Conclusion

References

Conclusion

- ▶ Privacy-preserving data publishing : a bushy literature
- ▶ Differential privacy is the current *de facto* standard thanks to its sound privacy guarantees and composability properties.
 - ▶ Strong support in academia (e.g., the inventors of differential privacy were awarded the prestigious Gödel Prize in 2017)
 - ▶ Major data-centric organisations have switched to differential privacy (e.g., Google and its COVID19 statistics, the Census Bureau and its decennial 2020 census, LinkedIn and its analytics about the interests of its users)
- ▶ Mechanisms for satisfying differential privacy have been proposed in various PPDP settings (e.g., Laplace for real-valued interactive queries, randomized response for local perturbation, synthetic data generation for centralized publishing)
- ▶ Research still very active !

Progress of the Talk

Introduction

Reminder : Partition-Based Models

Differential Privacy

Conclusion

References

- [1] B.-C. Chen, D. Kifer, K. LeFevre, and A. Machanavajjhala.
Privacy-preserving data publishing.
Found. Trends in Databases, 2(1-2):1–167, January 2009.
- [2] A. Cohen and K. Nissim.
Linear program reconstruction in practice.
CoRR, abs/1810.05692, 2018.
- [3] T. Dalenius.
Towards a methodology for statistical disclosure control.
Statistik Tidskrift, 15(5):429–444, 1977.
- [4] C. Dwork.
Differential privacy.
In *Proceedings of the 33rd International Conference on Automata, Languages and Programming - Volume Part II*, ICALP'06, pages 1–12, Berlin, Heidelberg, 2006.
Springer-Verlag.
- [5] C. Dwork, F. McSherry, K. Nissim, and A. Smith.
Calibrating noise to sensitivity in private data analysis.

In *Proceedings of the Third Conference on Theory of Cryptography*, TCC'06, pages 265–284, Berlin, Heidelberg, 2006. Springer-Verlag.

- [6] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkatasubramanian.

ℓ -diversity: Privacy beyond κ -anonymity.

In *Proceedings of the 22nd IEEE International Conference on Data Engineering*, ICDE '06, pages 24–, Washington, DC, USA, 2006. IEEE Computer Society.

- [7] L. Sweeney.

k-anonymity: a model for protecting privacy.

Int. J. Uncertain. Fuzziness Knowl.-Based Syst., 10(5):557–570, 2002.