Differential Privacy

A paper by Cynthia Dwork

Slides by Prajna Puranik

Introduction

- **Statistic:** Quantity computed from a sample
- When do we say a statistical database protects privacy?

  - When the database enables the user to learn properties of the population as a whole, while protecting the privacy of the individuals in the sample

- Before exploring the idea of privacy, we need to define the following:

➔ What constitutes a failure to preserve privacy?

➔ What is the power of the adversary whose goal it is to compromise privacy?

➔ What auxiliary information is available to the adversary (newspapers, medical studies, labor statistics) even without access to the database in question?

Notion of security defined by Dalenius in 1977 paper:

- access to a statistical database should not enable one to learn anything about an individual that could not be learned without access

This type of privacy cannot be achieved.

The obstacle = auxiliary information (information available to the adversary other than from access to the statistical database)

So, a new approach to privacy: the risk to one's privacy should not substantially increase as a result of participating in a statistical database.

This is captured by **differential privacy.**

Private Data Analysis: The Setting

There are two natural models for privacy mechanisms:

Non- interactive: The data collector, a trusted entity, publishes a “sanitized” version of the collected data -> remove well known identifiers like dob

Interactive: The data collector provides an interface through which users may pose queries about the data, and get (possibly noisy) answers.

Very powerful results for the interactive approach have been obtained while the non-interactive case has proven to be more difficult

Differential Privacy

The risk to one's privacy should not substantially increase as a result of participating in the statistical database

Definition 2: A randomized function K gives ε-differential privacy if for all data sets D1 and D2 differing on at most one element, and all $S \subseteq \text{Range}(K)$,

$$\Pr[K(D1) \in S] \leq \exp(\varepsilon) \times \Pr[K(D2) \in S]$$

where ε is a positive real number

S is a subset of image K -> set of all output values that it might produce

A mechanism K that satisfies this definition addresses concerns that any participant might have about the leakage of her personal information x: even if the participant removed her data from the data set, no outputs (and thus consequences of outputs) would become significantly more or less likely

Achieving Differential Privacy

The mechanism works by adding appropriately chosen random noise to the answer a = f(X), where f is the query function and X is the database

Magnitude of the random noise is chosen as a function of the largest change a single participant could have on the output to the query function => **sensitivity** of the function

Definition 3: For $f : D \rightarrow R^d$, the L1-sensitivity of f is

$\Delta f = \max || f(D1) - f(D2)||$

for all D1,D2 differing in at most one element and D is the collection of datasets

The privacy mechanism (Kf for a query function f) computes f(X) and adds noise -

Noise is added with a scaled symmetric exponential distribution with variance $\sigma2$ (to be determined in Theorem 4) in each component, described by the density function:

$Pr[Kf\ (X) = a] \propto exp(- // f(X) - a // /\sigma)$

Theorem 4: For $f : D \rightarrow Rd$, the mechanism Kf gives $(\Delta f/\sigma)$-differential privacy.

To achieve ε-differential privacy, one must choose $\sigma \geq \varepsilon/\Delta f$.

For many types of queries Δf will be quite small.

Our techniques work best – ie, introduce the least noise – when Δf is small

Theorem 5: For query strategy F = {fp : D → Rd}, the mechanism KF gives (ΔF/σ)-differential privacy.

Query strategy F is specified by a set of query functions fp where fp (X)i is the function describing the ith query given that the first i − 1 responses have been p1,p2,...,pi−1.

We require that fp(X)i = fp′(X)i if the first i − 1 responses in p and p′ are equal.

We define the sensitivity of a query strategy F = {fp : D → (R+)d} to be the largest sensitivity of any of its possible functions

Impossibility of Absolute Disclosure Prevention

Impossibility Result: Dalenius postulated that *access to a statistical database should not enable one to learn anything about an individual that could not be learned without access*

This is impossible if we want the database to be useful

To prove that this result is impossible, we need to define the following:

- **Utility:** Privacy mechanism is useful if its output is not predictable and the unpredictability should not just come from using random functions
- There should be a vector of questions whose answers should be learnable by the user

This is the utility vector w -> binary vector of fixed length k

A **privacy breach** for a database is described by a Turing machine C that takes as input:

- a description of a distribution D on databases
- a database DB drawn according to this distribution
- a string – the purported privacy breach

– and outputs a single bit

An **auxiliary information** generator is a Turing machine that takes as input:

- a description of the distribution D from which the database is drawn
- the database DB itself

- and outputs a string, z, of auxiliary information

Theorem 1: Fix any privacy mechanism San() and privacy breach decider C.

There is an auxiliary information generator X and an adversary A such that for all distributions D and for all adversary simulators A∗:

Pr[A(D, San(D, DB), X (D, DB)) wins] − Pr[A∗(D, X (D, DB)) wins] ≥ Δ where Δ is a suitably chosen (large) constant

The theorem below says that for any privacy mechanism San() and any distribution D satisfying certain technical conditions with respect to San(), *there is always some particular piece of auxiliary information, z, so that z alone is useless to someone trying to win, while z in combination with access to the data through the privacy mechanism permits the adversary to win* with probability arbitrarily close to 1.

Reference:

C. Dwork, "Differential Privacy, 33rd International Colloquium on Automata, Languages and Programming, part II", 2006