Differential Privacy

CS 297

Pragya Rana

Outline

- Introduction
- Privacy Data Analysis: The Setting
- Impossibility of Absolute Disclosure Prevention
- Achieving Differential Privacy

Introduction

- Statistic: quantity computed from a sample
- Can we reveal useful information from the statistical database, while protecting the privacy of the individuals in the sample?
- A rigorous treatment of privacy requires definitions:
  - 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 even without access to the database?

- notion of semantic security in 1977 paper of Dalenius:
  - access to a statistical database should not enable one to learn anything about an individual that could not be learned without access.
- But this type of privacy cannot be achieved.
- Obstacle -> auxiliary information
- New approach-> the risk to one's privacy should not substantially increase as a result of participating in a statistical database -> **Differential Privacy**

Privacy Data Analysis: The Setting

- Two models for privacy mechanisms:

1. Non-Interactive setting: data collector (a trusted entity) publishes a "sanitized" version of the collected data (sanitization employs techniques such as data perturbation and sub-sampling, removing well-known identifiers such as names, birthdates, ssn)
2. Interactive setting: data collector provides an interface through which users may pose queries about the data and get answers.

Non-Interactive vs Interactive approach:

- Powerful results for the interactive approach
- Non-interactive approach more difficult due to difficulty of supplying utility that has not yet been specified at the time the sanitization is carried out.

Impossibility of Absolute Disclosure Prevention

- Requires some notion of utility:
  - for the mechanism to be useful its output should not be predictable by the user
- Let utility vector be denoted by w. This is a binary vector of some 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 purpoted privacy breach and outputs a single bit

  require that C always halt. Adversary wins, with respect to C and for a given (D, DB) pair, if it produces a string s such that C(D, DB, s) accepts.

- Auxiliary information generator:
  - a Turing machine that takes as input a description of the distribution D from which the database is drawn as well as the database DB itself, and outputs a string, z, of auxiliary information.
  - This string is given both to the adversary and to a simulator. Simulator has no access of any kind to the database; adversary has access to the database via the privacy mechanism.

- The theorem 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 privacy mechanism permits the adversary to win the probability arbitrarily close to 1.

- **Theorem 1.**
  - For 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 satisfying Assumption 3 and for all adversary simulators A*,
  - Pr[A(D, San(D, DB), X (D, DB)) wins] − Pr[A∗(D, X (D, DB)) wins] ≥ Δ
- The distribution D completely captures any information that the adversary (and the simulator) has about the database, prior to seeing the output of the auxiliary information generator.

- Assumption 2
  1. $\forall\, 0 < \gamma < 1\ \exists\, n\gamma\ Pr_{DB \in_R D}[|DB| > n\gamma] < \gamma$; moreover $n\gamma$ is computable by a machine given D as input.
  2. There exists an l such that both the following conditions hold:
     (a) Conditioned on any privacy breach of length l, the min-entropy of the utility vector is at least l.
     (b) Every DB $\in$ D has a privacy breach of length l.
  3. Pr[B(D,San(DB)) wins] $\leq \mu$ for all interactive Turing machines B, where μ is a suitably small constant. The probability is taken over the coin flips of B and the privacy mechanism San(), as well as the choice of DB $\in_R$ D

- Definition 1. An $(M, m, l, t, \varepsilon)$ fuzzy extractor is given by procedures (Gen,Rec).
  1. Gen is a randomized generation procedure. On input $w \in M$ outputs an "extracted" string $r \in \{0, 1\}l$ and a public string p. For any distribution W on M of min-entropy m, if $(R, P) \leftarrow Gen(W)$ then the distributions (R, P) and (Ul,P) are within statistical distance ε.
  2. Rec is a deterministic reconstruction procedure allowing recovery of $r = R(w)$ from the corresponding public string $p = P(w)$ together with any vector w' of distance at most t from w. That is, if $(r,p) \leftarrow Gen(w)$ and $||w-w'||1 \leq t$ then $Rec(w', p) = r$.

- Assumption 3:
  - For some l satisfying Assumption 2(2b), for any privacy breach $y \in \{0,1\}l$, the min-entropy of (San(W)|y) is at least k+l, where k is the length of the public strings p produced by the fuzzy extractor.
- 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 $S \subseteq$ Range(K),

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

Achieving Differential Privacy

- a concrete interactive privacy mechanism achieving ε-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
- **Exponential Noise and the L1-Sensitivity**
  achieve ε-differential privacy by the addition of random noise whose magnitude is chosen as a function of the largest change a single participant could have on the output to the query function; refer to this quantity as the sensitivity of the function.

- Definition 3.
  - For $f : D \rightarrow Rd$, the L1-sensitivity of f is
    $\Delta f = \max D1,D2 \; // f(D1) - f(D2) //$

  for all D1,D2 differing in at most one element.
- The privacy mechanism, denoted Kf for a query function f, computes f(X) and adds noise with a scaled symmetric exponential distribution with variance σ2 in each component, described by the density function
  - $Pr[Kf (X) = a] \propto \exp(- // f(X) - a // 1/\sigma)$

- Theorem 4.
  - For $f : D \rightarrow Rd$, the mechanism Kf gives $(\Delta f/\sigma)$-differential privacy.
- Theorem 5.
  - For query strategy $F = \{f\rho : D \rightarrow Rd\}$, the mechanism KF gives $(\Delta F/\sigma)$-differential privacy.

References

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