CSE660
Differential Privacy

October 17, 2017

Marco Gaboardi
Room: 338-B
gaboardi@buffalo.edu
http://www.buffalo.edu/~gaboardi

Outline of the class

Week 1

Introduction, motivation and privacy limitations. Definition of Differential Privacy and the curator model.

Week 2

Basic mechanisms: Randomized Response, Laplace Mechanism,

Week 3

Basic properties following from the definition, Exponential Mechanism and comparison with the other basic mechanisms.

Week 4

The Report Noisy max algorithm.

Week 5

The Sparse Vector technique. Releasing Many Counting Queries with Correlated Noise. The smallDB algorithm.

Week 6

The MWEM algorithm.

Outline of the class

Week 7
Revisiting MWEM, The DualQuery algorithm.
Week 8
Advanced Composition and variations on differential privacy: Renyi DP, zero-concentrated DP.
Week 9
Studying the experimental accuracy.
The local model for differential privacy.
Week 10
More algorithms for the local model.
Week 11
PAC learning and private PAC learning
Week 12
Differentially Private Hypothesis Testing
Week 13
Differential Privacy and Generalization in Adaptive Data Analysis
Week 14
Project presentations

Differential privacy

Definition

Given $\varepsilon, \delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε, δ)-differentially private iff
for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon) \Pr[Q(b_2) \in S] + \delta$$

Blatantly non-privacy

The privacy mechanism $M:X^n \rightarrow R$ is blatantly non-private if an adversary can build a candidate database $D' \in X^n$, that agrees with the real database D in all but $o(n)$ entries:

$$d_H(D,D') \in o(n)$$

Differential privacy prevents blatantly non-privacy

Consider a uniformly random dataset $D \in X^n$.
Suppose $Q: X^n \rightarrow R$ is (ε,δ)-differentially private.
Then the the expected fraction of rows that any adversary can reconstruct is at most:

$$\frac{e^{\epsilon}}{|X|} + \delta$$