Lecturer: Kobbi Nissim
Department of Computer Science, Ben-Gurion University

Title: Smooth Sensitivity and Sampling in Private Data Analysis

Abstract:

The goal of private data analysis is to release aggregate statistics about a dataset while protecting the privacy of individuals whose information is contained in the dataset. A popular technique for privacy is output perturbation where a small amount of random noise is added to the released statistics, carefully crafted to protect individual privacy.

We introduce a new framework that expands the applicability of output perturbation. Departing from previous works, the noise magnitude in out framework is determined not only by the function F() we want to release, but also by the content of the dataset itself. One of the challenges is to ensure that the noise magnitude does not become a source of information leakage by itself. For that we introduce a quantity that we call smooth sensitivity. This is a measure of the variability of F() in the neighborhood of the particular instance dataset x. To our knowledge, this is the first formal analysis of the effect of instance-based noise in the context of data privacy.

Our framework raises many interesting algorithmic questions. Namely, to apply the framework one must compute or approximate the smooth sensitivity of F(x) on particular dataset instances. We show how to do this efficiently for several functions, including the median and the cost of the minimum spanning tree.

Finally, we also give a generic procedure based on sampling that allows one to release F(x) accurately on many `well behaved' datasets x. This procedure is applicable even when no efficient algorithm for approximating smooth sensitivity of F() is known or when F() is given as a black box. We illustrate the procedure by applying k-means clustering and learning mixtures of Gaussians.

The talk with be self contained.

Joint work with Sofya Raskhodnikova and Adam Smith. STOC 2007.