Ph: 9585554590, 9585554599
Email: support@salemsmartreach.com
URL: www.salemsmartreach.com

A Supermodularity-Based Differential Privacy Preserving Algorithm for Data Anonymization

Abstract:

Maximizing data usage and minimizing privacy risk are two conflicting goals. Organizations always apply a set of transformations on their data before releasing it. While determining the best set of transformations has been the focus of extensive work in the database community, most of this work suffered from one or both of the following major problems: scalability and privacy guarantee. Differential Privacy provides a theoretical formulation for privacy that ensures that the system essentially behaves the same way regardless of whether any individual is included in the database. In this paper, we address both scalability and privacy risk of data anonymization. We propose a scalable algorithm that meets differential privacy when applying a specific random sampling. The contribution of the paper is two-fold: 1) we propose a personalized anonymization technique based on an aggregate formulation and prove that it can be implemented in polynomial time; and 2) we show that combining the proposed aggregate formulation with specific sampling gives an anonymization algorithm that satisfies differential privacy. Our results rely heavily on exploring the supermodularity properties of the risk function, which allow us to employ techniques from convex optimization. Through experimental studies we compare our proposed algorithm with other anonymization schemes in terms of both time and privacy risk.