

Differentially Private Data Publishing and Analysis: a Survey

Abstract:

Digital information collected by corporations, organizations, and governments has resulted in a vast number of datasets, and the speed of such data collection has increased dramatically over the last few years. Typically, a data collector, also known as a *curator*, is in charge of publishing data for further analysis. However, most collected datasets contain private or sensitive information. Even though curators can apply several simple anonymization techniques, sensitive personal information still has a high probability of being disclosed. Privacy preservation has, therefore, become an urgent issue that needs to be addressed.

Existing System:

Differential privacy is an essential and prevalent privacy model that has been widely explored in recent decades. This survey provides a comprehensive and structured overview of two research directions: differentially private data publishing and differentially private data analysis. We compare the diverse release mechanisms of differentially private data publishing given a variety of input data in terms of query type, the maximum number of queries, efficiency, and accuracy. We identify two basic frameworks for differentially private data analysis and list the typical algorithms used within each framework.

Disadvantages:

- More Privacy budget consumed.

Proposed System:

Multi-disciplinary survey of work on differential privacy including an overview of the huge amount of literature in two major differential privacy research streams: data publishing and data analysis. We identified different publishing mechanisms for data publishing and compared various types of input and output data. We presented two basic dataset publishing methods: anonymized and learning-based. We discussed two basic frameworks for data analysis and illustrated their respective analysis scenarios. The basic techniques in differential privacy look

simple and intuitively appealing, and when combined with specific problems, differential privacy demonstrates itself as a powerful and useful tool for a diverse range of applications.

Advantages:

- Less Privacy budget consumed.

Modules:

- Differentially private data publishing.
- Differentially private data analysis.

SYSTEM REQUIREMENTS

H/W System Configuration:-

Processor	-	Pentium –III
RAM	-	256 MB (min)
Hard Disk	-	20 GB
Key Board	-	Standard Windows Keyboard
Mouse	-	Two or Three Button Mouse
Monitor	-	SVGA

S/W System Configuration:-

Operating System	:	Windows95/98/2000/XP
Application Server	:	Tomcat5.0/6.X
Front End	:	HTML, Jsp
Scripts	:	JavaScript.
Server side Script	:	Java Server Pages.
Database	:	MySQL 5.0

Database Connectivity : JDBC

www.takeoffprojects.com