Advances in Information Security

Volume 69

Series editor
Sushil Jajodia, George Mason University, Fairfax, VA, USA

More information about this series at http://www.springer.com/series/5576

Tianqing Zhu • Gang Li • Wanlei Zhou
Philip S. Yu

Differential Privacy and Applications

Springer

Tianqing Zhu
Deakin University
Melbourne, Australia

Wanlei Zhou
Deakin University
Melbourne, Australia

Gang Li
Deakin University
Melbourne, Australia

Philip S. Yu
University of Illinois at Chicago
Chicago, IL, USA

ISSN 1568-2633
Advances in Information Security
ISBN 978-3-319-62002-2 ISBN 978-3-319-62004-6 (eBook)
DOI 10.1007/978-3-319-62004-6

Library of Congress Control Number: 2017946488

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Corporations, organizations, and governments have collected, digitized, and stored information in digital forms since the invention of computers, and the speed of such data collection and volumes of stored data have increased dramatically over the past few years, thanks to the pervasiveness of computing devices and the associated applications that are closely linked to our daily lives. For example, hospitals collect records of patients, search engines collect online behaviors of users, social network sites collect connected friends of people, e-commerce sites collect shopping habits of customers, and toll road authorities collect travel details of vehicles. Such huge amounts of datasets provide excellent opportunities for businesses and governments to improve their services and to bring economic and social benefits, especially through the use of technologies dealing with big data, including data mining, machine learning, artificial intelligence, data visualization, and data analytics. For example, by releasing some statistics of hospital records may help medical research to combat diseases. However, as most of the collected datasets are personally related and contain private or sensitive information, data releasing may also provide a fertile ground for adversaries to obtain certain private and sensitive information, even though simple anonymization techniques are used to hide such information. Privacy preserving has, therefore, become an urgent issue that needs to be addressed in the digital age.

Differential privacy is a new and promising privacy framework and has become a popular research topic in both academia and industry. It is one of the most prevalent privacy models as it provides a rigorous and provable privacy notion that can be applied in various research areas, and can be potentially implemented in various application scenarios. The goal of this book is to summarize and analyze the state-of-the-art research and investigations in the field of differential privacy and its applications in privacy-preserving data publishing and releasing, so as to provide an approachable strategy for researchers and engineers to implement this new framework in real world applications.

This is the first book with a balanced view on differential privacy theory and its applications, as most existing books related to privacy preserving either do not

touch the topic of differential privacy or only focus on the theoretical analysis of differential privacy. Instead of using abstract and complex notions to describe the concepts, methods, algorithms, and analysis on differential privacy, this book presents these difficult topics in a combination of applications, in order to help students, researchers, and engineers with less mathematical background understand the new concepts and framework, enabling a wider adoption and implementation of differential privacy in the real world. The striking features of the book, differs from others, can be illustrated from three basic aspects:

- A detailed coverage on differential privacy in the perspective of engineering, rather than computing theory. The most difficult part in comprehending differential privacy is the complexity and the level of abstract of the theory. This book presents the theory of differential privacy in a more natural and easy to understand way.
- A rich set of state-of-the-art examples on various application areas helps readers to understand how to implement differential privacy in real world scenarios. Each application example includes a brief introduction to the problem and its challenges, a detailed implementation of differential privacy to solve the problem, and an analysis on the privacy and utility.
- A comprehensive collection of contemporary research results and issues in differential privacy. Apart from the basic theory, most of the contents of the book are from the recent publications in the last 5 years, reflecting the state-of-the-art of research and development in the area of differential privacy.

This book intends to enable readers, especially postgraduate and senior undergraduate students, to study up-to-date concepts, methods, algorithms, and analytic skills for building modern privacy-preserving applications through differential privacy. It enables students not only to master the concepts and theories in relation to differential privacy but also to readily use the material introduced into implementation practices. Therefore, the book is divided into two parts: theory and applications. In the theory part, after an introduction of the differential privacy preliminaries, the book presents detailed descriptions from an engineering viewpoint on areas of differentially private data publishing and differentially private data analysis where research on differential privacy has been conducted. In the applications part, after a summary on the steps to follow when solving the privacy-preserving problem in a particular application, the book then presents a number of state-of-the-art application areas of differential privacy, including differentially private social network data publishing, differentially private recommender system, differential location privacy, spatial crowdsourcing with differential privacy preservation, and correlated differential privacy for non-IID datasets. The book also includes a final chapter on the future direction of differential privacy and its applications.

Acknowledgments

We are grateful to many research students and colleagues at Deakin University in Melbourne and University of Illinois at Chicago, who have made a lot of comments to our presentations as their comments inspire us to write this book. We would like to acknowledge some support from research grants we have received, in particular, the Australian Research Council Grant no. DP1095498, LP120200266, and DP140103649, NSF through grants IIS-1526499, and CNS-1626432, and NSFC (Nos. 61672313, 61502362). Some interesting research results presented in the book are taken from our research papers that indeed (partially) supported through these grants. We also would like to express our appreciations to the editors in Springer, especially Susan Lagerstrom-Fife, for the excellent professional support. Finally we are grateful to the family of each of us for their consistent and persistent supports. Without their support, the book may just become some unpublished discussions.

Melbourne, Australia — Tianqing Zhu
Melbourne, Australia — Wanlei Zhou
Melbourne, Australia — Gang Li
Chicago, IL, USA — Philip S. Yu
May 2017

Contents