Privacy-preserving Computing

Privacy-preserving computing aims to protect the personal information of users while capitalizing on the possibilities unlocked by big data. This practical introduction for students, researchers, and industry practitioners is the first cohesive and systematic presentation of the field's advances over four decades. The book shows how to use privacy-preserving computing in real-world problems in data analytics and AI, and includes applications in statistics, database queries, and machine learning. The book begins by introducing cryptographic techniques such as secret sharing, homomorphic encryption, and oblivious transfer, and then broadens its focus to more widely applicable techniques such as differential privacy, trusted execution environment, and federated learning. The book ends with privacy-preserving computing in practice in areas like finance, online advertising, and healthcare, and finally offers a vision for the future of the field.

KAI CHEN is Professor at the Department of Computer Science and Engineering of the Hong Kong University of Science and Technology, where he leads the Intelligent Networking and Systems (iSING) Lab and the WeChat-HKUST Joint Lab on Artificial Intelligence Technology. His research interests include data center networking, high-performance networking, machine learning systems, and hardware acceleration.

QIANG YANG is Chief AI Officer at Webank and Professor Emeritus at the Department of Computer Science and Engineering of the Hong Kong University of Science and Technology. He is an AAAI, ACM, and IEEE Fellow and Fellow of the Canadian Royal Society. He has authored books such as *Intelligent Planning*, *Crafting Your Research Future*, *Transfer Learning*, and *Federated Learning*. His research interests include artificial intelligence, machine learning and data mining, automated planning, transfer learning, and federated learning.

Privacy-preserving Computing for Big Data Analytics and AI

KAI CHEN
Hong Kong University of Science and Technology

QIANG YANG
WeBank and Hong Kong University of Science and Technology

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781009299510

DOI: 10.1017/9781009299534

Originally published in Chinese as "Privacy-Preserving Computing" by Publishing House
of Electronics Industry in 2022

First published in English by Cambridge University Press & Assessment 2024

A catalogue record for this publication is available from the British Library

ISBN 978-1-009-29951-0 Hardback

Contents

978-1-009-29951-0 — Privacy-preserving Computing
Kai Chen , Qiang Yang
Frontmatter
More Information

Cambridge University Press & Assessment
978-1-009-29951-0 — Privacy-preserving Computing
Kai Chen , Qiang Yang
Frontmatter
More Information

Preface

We are in an era of big data where daily user activities generate huge amounts of data that fuel the advances of data-driven technologies, such as artificial intelligence (AI). However, these data inevitably contain private information of users, the disclosure of which would result in severe consequences. Therefore, how to exploit the knowledge contained within large-scale data without compromising user privacy becomes an important but challenging goal. The term *privacy-preserving computing* thus emerges as a summary of the theoretical and technical advances in pursuit of this goal.

Privacy-preserving computing is a field of rich history and fruitful achievements. Over 40 years ago, the theory of secure multiparty computation, which aims to jointly execute computing tasks while concealing partial inputs, marked the advent of privacy-preserving computing. In recent years, privacy-preserving computing remains an active research topic as we witness the technology of federated learning, enabling joint training of machine learning models without disclosing private data. Over the decades, privacy-preserving computing has grown into an inclusive and fruitful field, comprising secret sharing (SS), garbled circuits (GC), oblivious transfer (OT), differential privacy (DP), homomorphic encryption (HE), trusted execution environment (TEE), and federated learning (FL). In addition, with its applications in real-world tasks (such as database queries, data analytics, and machine learning) and scenarios (such as finance and health care), privacy-preserving computing is also a versatile subject that contributes to social well-being.

Despite the success and advances of privacy-preserving computing, we note that a comprehensive book that systematically describes the field is still absent. In fact, existing advances in privacy-preserving computing are still scattered in journal papers, technical talks, blogs, tutorials, and other publications without a unified and comprehensive taxonomy to summarize them. Consequently, the authors believe that the lack of a unified and systematic introduction hampers the development and application of privacy-preserving computing, as illustrated by the following examples:

- We gave a presentation entitled "Privacy-Preserving Computing: Theory and Efficiency" during a seminar organized by the China Computer Federation (CCF), where the audience mainly consisted of interested professors and students from universities in China. The presentation was a great success, and from the many questions received from the audience, we observed that despite their interests in privacy-preserving computing, their understanding of the topic was still vague and fragmented. Specifically, they were rather unclear about the scope, categorization, and detailed techniques in privacy-preserving computing. Thus, a comprehensive introduction that covers a wide range of privacy-preserving computing techniques would be helpful to students and researchers.
- We often met with organizations who were passionate about privacy-preserving computing but were not equipped with sufficient knowledge. A typical example would be the Hong Kong Science and Technology Park (HKSTP). As hundreds of sci-tech companies are located in HKSTP, it has the motivation to create a better environment for innovative startups. However, corporate data generally contains sensitive information about the companies and is thus not easily accessible. Therefore, we extensively discussed with HKSTP the concepts, techniques, and practical issues of federated learning. We believe that the interests in federated learning and other privacy-preserving computing techniques are general, and that a book that covers practical aspects and case studies of privacy-preserving computing would be helpful to industrial practitioners.

Motivated by our observations, we wrote this book on privacy-preserving computing in an attempt to build a unified taxonomy on privacy-preserving computing and also to guide its practical real-world applications. The whole process of writing the book lasted for over a year and involved the efforts of many students from the HKUST Intelligent Systems and Networking (iSING) Lab. We read and summarized many research papers, including some of our own, trying to introduce the fundamental techniques, case studies, and large-scale platforms of privacy-preserving computing in plain and comprehensible language. We finally envisioned the future directions and challenges of privacy-preserving computing.

To summarize, we hope that with this book on privacy-preserving computing we can build a unified and comprehensive taxonomy and overview of the field. Meanwhile, we are also aware that this book is still far from being an encyclopedia, in that it cannot cover every aspect of privacy-preserving computing. Nonetheless, we still hope that our efforts can mark the first step toward this goal and motivate future researchers to make new contributions.

Summary of Contents

The contents of this book can be divided into three parts:

(i) Encrypted computation (Chapters 2–5). This part of the book aims to introduce cryptographic techniques to achieve privacy-preserving computing, including secret sharing (SS), homomorphic encryption (HE), oblivious transfer (OT), and garbled circuits (GC). These cryptographic techniques serve as foundations of many privacy-preserving computing protocols and applications. In each chapter, we cover basic knowledge about the cryptographic technique and some practical examples of applications.

(ii) Privacy-preserving computation (Chapters 6–8). This part of the book aims to introduce noncryptographic techniques to achieve privacy-preserving computing, including differential privacy (DP), trusted execution environment (TEE), and federated learning (FL). These techniques focus on protecting data privacy in a more diverse range of application scenarios.

(iii) Privacy-preserving computing platforms and case studies (Chapters 9–10). This part of the book aims to show how the introduced techniques are successfully applied in practice and on a large scale. Chapter 9 introduces the federated learning platform, FATE, as well as some platforms for encrypted databases. It also covers the efficiency problem in real-world privacy-preserving computing platforms and potential solutions. Chapter 10 introduces some case studies where privacy-preserving computing techniques are applied, including finance, risk management, online advertising, database queries, health care, and public services.

Acknowledgments

First, we would like to express our gratitude toward a group of outstanding Ph.D. students, researchers, and engineers who have dedicated huge amounts of effort to this book, including (in alphabetical order)

- Di Chai, who contributed to the writing of Chapters 2 and 10.
- Tianjian Chen, who contributed to the writing of Chapter 10.
- Xiaodian Cheng, who contributed to the writing of Chapter 9.
- Kun Guo, who contributed to the writing of Chapter 10.
- Shuihai Hu, who contributed to the writing of Chapter 9.
- Yilun Jin, who contributed to the writing of Chapters 3, 4, 5, 6, and 10.
- Zhenghang Ren, who contributed to the writing of Chapters 4, 5, 7, 8, and 10.
- Han Tian, who contributed to the writing of Chapters 3, 6, and 10.
- Liu Yang, who contributed to the writing of Chapters 2 and 10.
- Junxue Zhang, who contributed to the writing of Chapter 7.

During the preparation of this book, we consulted over 200 related books, articles, and research papers. We would also like to thank the authors of these works for their contributions to the field of privacy-preserving computing.

Finally, we would like to thank our families for their understanding and continued support. Without them, the book would not have been possible.