Chapter 6
Summary and Outlook

In this book, we summarize the latest work on personalized privacy protection in terms of information technology. Personalized privacy protection is still in its infancy. The theories, algorithms, and other conceptual designs surveyed in this book could be a starting point for forthcoming researchers and readers to probe this under-explored domain. We aim to offer a systematic summary of existing research and application outputs on personalized privacy protection, which also testifies the theoretical and practical applicability in diverse big data scenarios. We also subsequently present a couple of potentially promising directions, with which we expect to assist in avoiding superfluous efforts from subsequent interested explorers.

In general, an abundant volume of literature has been reviewed and analysed to show the current research and application status of personalized privacy protection solutions. We describe and compare the primary privacy concerns and attacks, some of which remain a bottleneck to personalized privacy protection. In particular, we have discussed the personalized privacy protection in cyber physical systems, social networks, smart homes, and location-based services. However, the proposed models are generalized models and able to be applied in more extensive scenarios.

We also include several mainstream theories for personalized privacy protection, including differential privacy, machine learning, game theory, and anonymity and clustering-based methods, and correspondingly explained and articulated while their feasibility has been demonstrated when fitting into various real-world practices.

Based on the existing research and results, we further discuss several future research directions, which are personalized privacy-preserving attribute-based encryption, personalized privacy-preserving federated learning using generative adversarial network, personalized privacy-preserving blockchain-enabled federated learning, collusion attack resistance in personalized privacy protection, and trade-off optimization between personalized privacy protection and data utility.

As aforementioned, privacy protection in the digital space is a new research domain. We have far more questions than answers, we definitely will counter many

Y. Qu et al., *Personalized Privacy Protection in Big Data*, Data Analytics,
https://doi.org/10.1007/978-981-16-3750-6_6

unprecedented problems, and many unknown of unknowns. Based on our study, we would like to share some big pictures with energetic young researcher as below.

First of all, we believe privacy protection needs the effort from multiple disciplines, for example, psychology, social science, law, information technology, and so on. It is certain that computer science herself cannot solve the privacy problem alone. According to the logic of science, we need firstly measure privacy, then represent privacy using mathematical models, and then confirm the ideas and conclusions by theoretical or experimental proof. Up to date, we do not have an effective way to get the first step done, namely measuring privacy. Similar to other soft concepts like happiness, madness, privacy is hard to measure. Issac Newton complained that "I can calculate the movement of stars, but not the madness of men", after 300 years, we still not face the similar difficulty. At the other hand, we see the dramatic development of all disciplines in the science family, we believe the light is on the horizon for us, but we do need to master multiple necessary skill sets to complete the mission.

Cross discipline research looks beautiful, but not easy to carry out. The Science magazine had a statistics several years ago, the result showed that among the research papers published on Science in the last 100 years, nearly 70% papers are cross disciplinary work. This result demonstrates that cross disciplinary study is powerful in research. However, our experience and the literature also show that it is hard to execute it. In general, a few coffees at the campus may generate some idea among colleagues from different discipline, but when we execute it, it is extremely hard as we speak different "languages". Strong leadership and financial support maybe the key for these kinds of collaboration. A common suggestion is we need to learn the skills of the other disciplines rather than bringing problems to the other party and waiting for solutions.

Secondly, theoretical tools for privacy is desperately needed, and it is a promising target for related communities. So far, differential privacy is the only new tool invented for privacy (we treat cryptography as the tool for secrecy sharing, not for privacy). However, differential privacy was invented for privacy protection in statistical information retrieval, which is only a very small part of the landscape of digital privacy. There are two possibilities on this issue.

- We extend the existing tools to deal with the new problems of privacy, e.g., upgrade cryptographic tools to fulfill the tasks of privacy preserving in big data publishing. The traditional symmetric and asymmetric encryption tools were designed to share a secret between two pairs (one-to-one communication), and the attribute-based encryption was developed to share a secret among a small group (one-to-many communication, we note the many here is a small number). However, in big data publishing, we release the data for anyone who wants to access under the condition of protecting the privacy of the data owners (one-to-any communication). So far cryptography cannot offer a suitable solution for it.
- Invent new tools. The invention of new tools is the result of application demands. We believe the research community will develop new tools for digital privacy as the demands are in place. It is a tough job, and also an exciting goal for hard working and talented people.

This small book is the short summary of our work in the recent years, and also the first step of our research group. We hope our shallow effort can attract interested readers to explore the promising land with us in both academia and industry domains.

We hope you enjoy reading the book, and sincerely looking forward to have your feedback, comments, suggestions and your work in the field.