

SPRINGER BRIEFS ON
CYBER SECURITY SYSTEMS AND NETWORKS

Kwangjo Kim

Harry Chandra Tanuwidjaja

Privacy-Preserving Deep Learning A Comprehensive Survey

SpringerBriefs on Cyber Security Systems and Networks


Editor-in-Chief


Yang Xiang, Digital Research & Innovation Capability Platform, Swinburne University of Technology, Hawthorn, VIC, Australia

Series Editors

Liqun Chen , Department of Computer Science, University of Surrey, Guildford, Surrey, UK

Kim-Kwang Raymond Choo , Department of Information Systems, University of Texas at San Antonio, San Antonio, TX, USA

Sherman S. M. Chow , Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, Hong Kong

Robert H. Deng , School of Information Systems, Singapore Management University, Singapore, Singapore

Dieter Gollmann, E-15, TU Hamburg-Harburg, Hamburg, Hamburg, Germany

Kuan-Ching Li, Department of Computer Science & Information Engineering, Providence University, Taichung, Taiwan

Javier Lopez, Computer Science Department, University of Malaga, Malaga, Spain

Kui Ren, University at Buffalo, Buffalo, NY, USA

Jianying Zhou , Infocomm Security Department, Institute for Infocomm Research, Singapore, Singapore

The series aims to develop and disseminate an understanding of innovations, paradigms, techniques, and technologies in the contexts of cyber security systems and networks related research and studies.

It publishes thorough and cohesive overviews of state-of-the-art topics in cyber security, as well as sophisticated techniques, original research presentations and in-depth case studies in cyber systems and networks. The series also provides a single point of coverage of advanced and timely emerging topics as well as a forum for core concepts that may not have reached a level of maturity to warrant a comprehensive textbook.

It addresses security, privacy, availability, and dependability issues for cyber systems and networks, and welcomes emerging technologies, such as artificial intelligence, cloud computing, cyber physical systems, and big data analytics related to cyber security research. The main focus is on the following research topics:

Fundamentals and theories

- Cryptography for cyber security
- Theories of cyber security
- Provable security

Cyber Systems and Networks

- Cyber systems security
- Network security
- Security services
- Social networks security and privacy
- Cyber attacks and defense
- Data-driven cyber security
- Trusted computing and systems

Applications and others

- Hardware and device security
- Cyber application security
- Human and social aspects of cyber security


More information about this series at <http://www.springer.com/series/15797>

Kwangjo Kim · Harry Chandra Tanuwidjaja

Privacy-Preserving Deep Learning

A Comprehensive Survey

Kwangjo Kim 
School of Computing
Korea Advanced Institute of Science
and Technology (KAIST)
Daejeon, Korea (Republic of)

Harry Chandra Tanuwidjaja 
School of Computing
Korea Advanced Institute of Science
and Technology (KAIST)
Daejeon, Korea (Republic of)

ISSN 2522-5561 ISSN 2522-557X (electronic)
SpringerBriefs on Cyber Security Systems and Networks
ISBN 978-981-16-3763-6 ISBN 978-981-16-3764-3 (eBook)
<https://doi.org/10.1007/978-981-16-3764-3>

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

*To our friends and families for their lovely
support.*

Preface

This monograph aims to give a survey on the state of the art of Privacy-Preserving Deep Learning (PPDL), which is considered to be one of the emerging technologies by combining classical privacy-preserving and cryptographic protocols with deep learning in a systematic way.

Google and Microsoft announced a big investment in PPDL in the early 2019, followed by the announcement of “Private Join and Compute”, an open-source PPDL tools that is based on Secure Multi-Party Computation (Secure MPC) and Homomorphic Encryption (HE) on June 2019 by Google. One of the main issues in PPDL is about its applicability, e.g., to understand the gap between the theory and practice exists. In order to solve this, there are many advances relying on the classical privacy-preserving method (HE, secure MPC, differential privacy, secure enclaves, and its hybrid) and together with deep learning. The basic architecture of PPDL is to build a cloud framework that enables collaborative learning while keeping the training data on the client device. After the model is fully trained, the privacy during the sensitive data exchange or storage must be strictly preserved and the overall framework must be feasible for the real applications.

This monograph plans to provide the fundamental understandings for privacy preserving and deep learning, followed by comprehensive overview of the state of the art of PPDL methods, suggesting the pros and cons of each method, and introducing the recent advances of the federated learning and split learning-based PPDL called as Privacy-Preserving Federated Learning (PPFL). In addition, this monograph gives a guideline to general people and students, and practitioners who are interested to know about PPDL and also helping early-stage researcher who wants to explore PPDL area. We hope that the early-stage researchers can grasp the basic theory of PPDL, understand the pros and cons of current PPDL and PPFL methods, addressing the gap between theory and practice in the most recent approach, so that they can propose their own method later.

Daejeon, Korea (Republic of)
March 2021

Kwangjo Kim
Harry Chandra Tanuwidjaja

Acknowledgements

This monograph was partially supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea Government Ministry of Science and ICT (MSIT) (No. 2017-0-00555, Towards Provable-secure Multi-party Authenticated Key Exchange Protocol based on Lattices in a Quantum World and No. 2019-0-01343, Regional Strategic Industry Convergence Security Core Talent Training Business.)

The authors sincerely appreciate the contribution of the lovely alumni and members of Cryptology and Information Security Lab (CAISLAB), Graduate School of Information Security, School of Computing, KAIST, including but not limiting to Rakyong Choi, Jeeun Lee, Seungeun Baek, Muhamad Erza Aminanto, and Edwin Ayisi Opare for their volunteering help and inspiring discussions.

In particular, we would like to mention our sincere gratitude to Indonesia Endowment Fund for Education, Lembaga Pengelola Dana Pendidikan (LPDP) for supporting Harry Chandra Tanuwidjaja during his Ph.D. study at KAIST.

We gratefully acknowledge the editors of this monograph series on security for their valuable comments and the Springer for giving us the opportunity to write and publish this monograph.

Finally, we also are very grateful for our families for their strong support and endless love *at the foot of the spiritual Kyerong Mt. which is located at the center of the Republic of Korea.*

Contents

- 1 Introduction** 1
 - 1.1 Background 1
 - 1.2 Motivation 2
 - 1.3 Outline 3
 - References 4
- 2 Preliminaries** 7
 - 2.1 Classical Privacy-Preserving Technologies 7
 - 2.1.1 Group-Based Anonymity 7
 - 2.1.2 Cryptographic Method 8
 - 2.1.3 Differential Privacy 10
 - 2.1.4 Secure Enclaves 11
 - 2.2 Deep Learning 11
 - 2.2.1 Outline of Deep Learning 11
 - 2.2.2 Deep Learning Layers 12
 - 2.2.3 Convolutional Neural Network (CNN) 14
 - 2.2.4 Generative Adversarial Network (GAN) 14
 - 2.2.5 Support Vector Machine 14
 - 2.2.6 Recurrent Neural Network 15
 - 2.2.7 *K*-Means Clustering 16
 - 2.2.8 Reinforcement Learning 17
 - References 19
- 3 X-Based PPDL** 23
 - 3.1 HE-Based PPDL 23
 - 3.2 Secure MPC-Based PPDL 29
 - 3.3 Differential Privacy-Based PPDL 34
 - 3.4 Secure Enclaves-Based PPDL 38
 - 3.5 Hybrid-Based PPDL 40
 - References 42