

Wireless Networks

Series Editor

Xuemin Sherman Shen
University of Waterloo
Waterloo, ON, Canada

The purpose of Springer's new Wireless Networks book series is to establish the state of the art and set the course for future research and development in wireless communication networks. The scope of this series includes not only all aspects of wireless networks (including cellular networks, WiFi, sensor networks, and vehicular networks), but related areas such as cloud computing and big data. The series serves as a central source of references for wireless networks research and development. It aims to publish thorough and cohesive overviews on specific topics in wireless networks, as well as works that are larger in scope than survey articles and that contain more detailed background information. The series also provides coverage of advanced and timely topics worthy of monographs, contributed volumes, textbooks and handbooks.

More information about this series at <http://www.springer.com/series/14180>

Yuan Zhang • Chunxiang Xu •
Xuemin Sherman Shen


Data Security in Cloud Storage



Springer

Yuan Zhang 
School of Computer Science & Engineering
University of Electronic Science
and Technology of China
Chengdu, Sichuan, China

Chunxiang Xu 
School of Computer Science & Engineering
University of Electronic Science
and Technology of China
Chengdu, Sichuan, China

Xuemin Sherman Shen 
Department of Electrical and Computer
Engineering
University of Waterloo
Waterloo, ON, Canada

ISSN 2366-1186

Wireless Networks

ISBN 978-981-15-4373-9

<https://doi.org/10.1007/978-981-15-4374-6>

ISSN 2366-1445 (electronic)

ISBN 978-981-15-4374-6 (eBook)

© Springer Nature Singapore Pte Ltd. 2020

This work is subject to copyright. All rights are reserved 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

Preface

Cloud storage is a service that lets users store data by transferring it over the Internet or another network to an offsite storage system maintained by a third party. It is increasingly demanded along the users' data exploded and it provides users an efficient and convenient way to manage their data. Despite the appealing advantages of data outsourcing, cloud storage services are confronted with various threats from many aspects. Compared with traditional data storage systems (where users store their data locally), cloud storage provides users with a completely different paradigm to manage their data, which also introduces new and challenging threats towards data security. Specifically, in the cloud storage service, users do not physically own their data once outsourcing the data to a cloud server (which is subject to a cloud service provider), and these data are fully controlled by the cloud server. As such, once the cloud server (including insiders working at the cloud service provider) misbehaves, the outsourced data would suffer from threats, such as corruption, modification, removal, and privacy violation. In addition, since the data are transmitted over public and insecure networks, external adversaries (e.g., hackers) might eavesdrop on the communication channel between the user and the cloud server, tamper with the interaction messages between them, and extract the data contents from the cloud server for financial or political reasons.

This monograph gives a comprehensive overview of data security in cloud storage, which includes cloud storage reliability, cloud storage confidentiality, and data investigations in cloud storage. With these security issues, five research topics are introduced and studied, i.e., secure verification of data integrity, secure deduplication, secure keyword search, secure data provenance, and secure data time-stamping. This monograph not only presents basic paradigms and principles of the aforementioned research topics and the corresponding techniques that secure cloud storage but also provides a comprehensive survey on each of the research topics. In addition, this monograph also analyzes the relationship among these research topics.

As emerging techniques, such as indistinguishability obfuscation, blockchains, and trusted execution environments (TEEs), have been developed in the last decade, it has shown great potentials in enhancing data security. This monograph also introduces the latest advances in enhancing cloud storage reliability, confidentiality,

and investigations and analyzes their pros and cons. Finally, open research issues and future work on the related topics are also discussed.

We would like to thank Prof. Nan Cheng (Xidian University), Prof. Hongwei Li (University of Electronic Science and Technology of China), Prof. Xiaohui Liang (University of Massachusetts at Boston), Prof. Xiaodong Lin (University of Guelph), Prof. Jianbing Ni (Queen's University), Prof. Haomiao Yang (University of Electronic Science and Technology of China), Prof. Kan Yang (The University of Memphis), Prof. Shui Yu (University of Technology Sydney), Prof. Xiaojun Zhang (Southwest Petroleum University), and Prof. Jianying Zhou (Singapore University of Technology and Design) for their contributions in the presented research works. We would also like to thank Shanshan Li and Dongxiao Liu for reviewing parts of this monograph and all the members of BBCR group for the valuable discussions and their insightful suggestions, ideas, and comments. Special thanks also go to the staff at Springer Science+Business Media: Celine Chang, Susan Lagerstrom-Fife, Jane Li, and Suraj Kumar, for their help throughout the publication process.

Chengdu, China
Chengdu, China
Waterloo, Canada

Yuan Zhang
Chunxiang Xu
Xuemin Sherman Shen

Contents

- 1 Introduction** 1
 - 1.1 An Overview of Cloud Storage..... 2
 - 1.1.1 Cloud Storage Architecture 2
 - 1.1.2 Cloud Storage Applications 3
 - 1.2 Data Security in Cloud Storage..... 5
 - 1.3 Organization of the Monograph 7
 - References 9
- 2 Basic Techniques for Data Security** 11
 - 2.1 Data Authentication 11
 - 2.1.1 Message Authentication Code 12
 - 2.1.2 Hash Function..... 13
 - 2.1.3 Digital Signature 14
 - 2.2 Data Confidentiality..... 16
 - 2.2.1 Symmetric-Key Encryption 16
 - 2.2.2 Public-Key Encryption 17
 - 2.3 Threshold Cryptography 18
 - 2.4 Public-Key Cryptosystems 18
 - 2.4.1 PKI-Based Public-Key Cryptosystems 19
 - 2.4.2 Identity-Based Public-Key Cryptosystems 19
 - 2.4.3 Certificateless Public-Key Cryptosystems 20
 - 2.5 Blockchain..... 20
 - 2.6 Trusted Execution Environments 24
 - 2.7 Summary and Further Reading 25
 - References 25
- 3 Cloud Storage Reliability** 29
 - 3.1 Data Integrity 29
 - 3.2 Proofs of Storage: Definition and Criteria 30
 - 3.2.1 Threat Models..... 31
 - 3.2.2 Security Criteria 33
 - 3.3 Proofs of Storage for Cloud Storage Systems 34

3.3.1	Proofs of Storage for Dynamic Data	38
3.3.2	Enhancement of Security	40
3.3.3	Constructing Public Verification on Different Cryptosystems	42
3.3.4	Other Works	43
3.4	Latest Advances in Proofs of Storage	44
3.4.1	Proofs of Storage Based on Indistinguishability Obfuscation	44
3.4.2	Proofs of Storage Based on Blockchain	47
3.5	Summary and Further Reading	51
	References	52
4	Secure Deduplication	55
4.1	Deduplication Classification	55
4.2	Secure Deduplication: Threats and Countermeasures	57
4.2.1	Proofs of Ownership	58
4.2.2	Randomized Deduplication	60
4.3	Message-Locked Encryption	60
4.3.1	Overview	61
4.3.2	Threat Models of Encrypted Deduplication Storage Systems	64
4.3.3	Security Definition	65
4.4	Encrypted Deduplication Systems	65
4.4.1	Enhancement of Security	66
4.4.2	Practical Concern	73
4.4.3	Other Works	76
4.5	When Secure Deduplication Meets eHealth: A Case Study	76
4.5.1	Cloud-Based eHealth Systems	77
4.5.2	Adversary Model and Security Goals	78
4.5.3	Analysis of EMRs in Actual eHealth Systems	79
4.5.4	Study of HealthDep	81
4.6	Summary and Further Reading	84
	References	84
5	Secure Keyword Search	87
5.1	Keyword Search Over Encrypted Data	87
5.2	Symmetric-Key Searchable Encryption	89
5.2.1	System and Threat Models	89
5.2.2	Survey on Symmetric-Key Searchable Encryption	89
5.3	Public-Key Searchable Encryption	98
5.3.1	System model	99
5.3.2	Threat Model and Security Definition	100
5.3.3	Survey on Public-Key Searchable Encryption	100
5.4	Latest Advances in Public-Key Searchable Encryption	104
5.4.1	Public-Key Searchable Encryption Against Keyword Guessing Attacks	104

5.4.2	Remark and Further Discussion	112
5.5	Summary and Further Reading	113
	References	114
6	Secure Data Provenance	119
6.1	Introduction to Secure Data Provenance	119
6.1.1	Data Provenance vs. Secure Data Provenance	120
6.1.2	System and Threat Models	123
6.2	Survey on Secure Data Provenance	125
6.3	Blockchain: A Panacea for Secure Data Provenance	127
6.3.1	Blockchain-Based Secure Data Provenance	128
6.3.2	Implementation Based on Ethereum	135
6.3.3	Data Provenance and Beyond: Further Discussion	137
6.4	Summary and Further Reading	139
	References	140
7	Secure Data Time-Stamping	143
7.1	Introduction to Secure Data Time-Stamping	143
7.1.1	What Kinds of Data Would Benefit from Secure Time-Stamping?	144
7.1.2	System and Threat Models	145
7.2	Survey on Secure Time-Stamping	146
7.3	Secure Time-Stamping and Blockchain	149
7.3.1	Distributed Cryptocurrencies from Secure Time-Stamping	150
7.3.2	Secure Time-Stamping from Blockchain	151
7.4	Summary and Further Reading	164
	References	164
8	Summary and Future Research Directions	167
8.1	Summary	167
8.2	Future Work	169
8.2.1	Secure Data Integrity Verification from Smart Contract	169
8.2.2	Combination of Encrypted Deduplication and Symmetric-key Searchable Encryption	170
8.2.3	Secure Provenance Under Complex Models	171
8.2.4	Securely Time-stamping Operations in the Digital World	171

Acronyms

CA	Certificate authority
CDN	Content distribution network
CE	Convergent encryption
eHealth	Electronic healthcare
EHRs	Electronic health records
EMRs	Electronic medical records
EPD	Essential provenance data
FE	Functional encryption
FHE	Fully homomorphic encryption
HIPAA	Health Insurance Portability and Accountability Act
HVTs	Homomorphic verifiable tags
IdP	Identity provider
IMEI	International Mobile Equipment Identity
<i>iO</i>	Indistinguishability obfuscation
IoT	Internet of things
IRS	Index Repository Service
KGA	Keyword guessing attack
KGC	Key generation center
MHT	Merkle hash tree
MLE	Message-locked encryption
NPD	Nonessential provenance data
OPRF	Oblivious pseudorandom function
ORAM	Oblivious random access machine
PDP	Provable data possession
PIR	Private information retrieval
PKI	Public-key infrastructure
PKG	Private key generator
PoR	Proofs of retrievability
PoS	Proof of stake
PoW	Proof of work
POW	Proof of ownership

PSE	Public-key searchable encryption
SE	Searchable encryption
SSE	Symmetric-key searchable encryption
SIM	Subscriber identity module
TEEs	Trusted execution environments
TPA	Third-party auditor
TSP	Time-stamping service provider
WoL	Window of latching
WoT	Window of time-stamping