

DOCTORAL THESIS

Authenticated query processing in the cloud

Xu, Cheng

Date of Award:
2019

[Link to publication](#)

General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

Abstract

With recent advances in data-as-a-service (DaaS) and cloud computing, outsourcing data to the cloud has become a common practice. In a typical scenario, the data owner (DO) outsources the data and delegates the query processing service to a service provider (SP). However, as the SP is often an untrusted third party, the integrity of the query results cannot be guaranteed and is thus imperative to be authenticated. To tackle this issue, a typical approach is letting the SP provide a cryptographic proof, which can be used to verify the soundness and completeness of the query results by the clients. Despite extensive research on authenticated query processing for outsourced databases, existing techniques have only considered limited query types. They fail to address a variety of needs demanded by enterprise customers such as supporting aggregate queries over set-valued data, enforcing fine-grained access control, and using distributed computing paradigms. In this dissertation, we take the first step to comprehensively investigate the authenticated query processing in the cloud that fulfills the aforementioned requirements. Security analysis and performance evaluation show that the proposed solutions and techniques are robust and efficient under a wide range of system settings.

Keywords: Query Authentication, Query Processing, Aggregate Queries, Access Control, Distributed Systems, Data Integrity

Table of Contents

Declaration	i
Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures	viii
List of Tables	x
List of Algorithms	xi
Chapter 1 Introduction	1
1.1 Background	1
1.2 Authenticating Aggregate Queries over Set-Valued Data	2
1.3 Authenticating Relational Queries with Fine-Grained Access Control	5
1.4 Authenticating kNN Queries in Distributed Settings	7
1.5 Dissertation Organization	8
Chapter 2 Literature Review	9
2.1 Authenticated Query Processing	9
2.2 Aggregate Queries over Set-Valued Data	12
2.3 Access Control	12
2.4 Distributed Spatial Queries	13
Chapter 3 Authenticating Aggregate Queries over Set-Valued Data	15
3.1 Problem Formulation	15

3.2	Preliminaries	18
3.3	PA ² : Privacy-Preserving Authentication Framework for Aggregate Queries	21
3.3.1	Privacy-Preserving Authentication Protocols on Multiset Operations	21
3.3.2	Privacy-Preserving Authentication Algorithms on Aggregate Queries	25
3.3.3	Privacy-Preserving Authentication on Candidate Object Selection	29
3.3.4	Cost Analysis	32
3.4	Security Analysis	35
3.4.1	Security of PA ² Algorithms	35
3.4.2	Privacy Guarantee on Sensitive Features	36
3.4.3	Discussion on Privacy Guarantee	37
3.5	Optimizations	38
3.5.1	Optimized MG-Tree	38
3.5.2	Accumulating $sum(\cdot)$ by Linear Ordering	41
3.5.3	Acceleration by Parallelism	42
3.6	Performance Evaluation	42
3.6.1	Performance of Privacy-Preserving Multiset Authentication Protocols	44
3.6.2	Query Authentication Performance	45
3.6.3	Impact of Optimizations	47
3.6.4	Scalability Test on Dimensionality and Object Cardinality	48
3.7	Chapter Summary	50

Chapter 4 Authenticating Relational Queries with Fine-Grained Access

Control	51
4.1 Problem Formulation	51
4.2 Preliminaries	54
4.3 Equality Query Authentication	56
4.3.1 ADS Generation and Query Processing	57
4.3.2 ABS with Predicate Relaxation	60
4.4 Range & Join Query Authentication	67
4.4.1 Range Query Authentication	67
4.4.2 Join Query Authentication	71
4.5 Security Analysis	73
4.5.1 Security Analysis on ABS	73
4.5.2 Security Analysis on Query Authentication	75
4.6 Optimizations	77
4.6.1 Hierarchical Role Assignment	77
4.6.2 Acceleration by Parallelism	77
4.7 Relaxing Zero-Knowledge Confidentiality Requirement	78
4.7.1 Using k -d Tree Index Structure	78
4.7.2 Supporting Continuous Query Attributes	80
4.8 Handling Duplicate Records	81
4.9 Handling Updates	82
4.9.1 Client Access Role Updates	82
4.9.2 Data Record Updates	83
4.10 Performance Evaluation	84
4.10.1 Equality Query Performance	86
4.10.2 Range and Join Query Performance	86
4.10.3 Impact of the Optimizations	88

4.10.4	Performance with Duplicate Records and Updates	90
4.11	Chapter Summary	92
Chapter 5	Authenticating kNN Queries in Distributed Settings	93
5.1	Problem Formulation	93
5.2	Preliminaries	94
5.2.1	Cryptographic Primitives	94
5.2.2	Spatial Authenticated Data Structure	95
5.2.3	Local Authenticated kNN Processing	96
5.3	Distributed kNN Authentication	97
5.3.1	Distributed MR-Tree	97
5.3.2	Authenticating Distributed kNN Query Processing	98
5.3.3	Client Verification	103
5.3.4	Robustness Analysis	104
5.4	Optimization of VO Size	105
5.5	Performance Evaluation	108
5.5.1	Cost of Index Construction	109
5.5.2	Distributed Authenticated kNN Query Cost	110
5.6	Chapter Summary	113
Chapter 6	Conclusions	114
6.1	Contributions	114
6.2	Future Directions	115
Appendix A	Proof of Theorem 3.1	118
Appendix B	Proof of Theorem 4.1	122
Appendix C	Proof of Theorem 4.2	128
Bibliography		130
List of Publications		140
Curriculum Vitae		142