

Guojun Wang · Mohammed Atiquzzaman  
Zheng Yan · Kim-Kwang Raymond Choo (Eds.)

LNCS 10656

# Security, Privacy, and Anonymity in Computation, Communication, and Storage

10th International Conference, SpaCCS 2017  
Guangzhou, China, December 12–15, 2017  
Proceedings



 Springer

  
GU CN  
SpaCCS 2017

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, Lancaster, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

Josef Kittler

*University of Surrey, Guildford, UK*

Jon M. Kleinberg

*Cornell University, Ithaca, NY, USA*

Friedemann Mattern

*ETH Zurich, Zurich, Switzerland*

John C. Mitchell

*Stanford University, Stanford, CA, USA*

Moni Naor

*Weizmann Institute of Science, Rehovot, Israel*

C. Pandu Rangan

*Indian Institute of Technology, Madras, India*

Bernhard Steffen

*TU Dortmund University, Dortmund, Germany*

Demetri Terzopoulos

*University of California, Los Angeles, CA, USA*

Doug Tygar

*University of California, Berkeley, CA, USA*

Gerhard Weikum

*Max Planck Institute for Informatics, Saarbrücken, Germany*

More information about this series at <http://www.springer.com/series/7409>

Guojun Wang · Mohammed Atiquzzaman  
Zheng Yan · Kim-Kwang Raymond Choo (Eds.)

# Security, Privacy, and Anonymity in Computation, Communication, and Storage

10th International Conference, SpaCCS 2017  
Guangzhou, China, December 12–15, 2017  
Proceedings

*Editors*

Guojun Wang  
Guangzhou University  
Guangzhou  
China

Mohammed Atiquzzaman  
Edith Kinney Gaylord Presidential Professor  
University of Oklahoma  
Norman, OK  
USA

Zheng Yan   
Aalto University  
Espoo  
Finland

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

ISSN 0302-9743                      ISSN 1611-3349 (electronic)  
Lecture Notes in Computer Science  
ISBN 978-3-319-72388-4              ISBN 978-3-319-72389-1 (eBook)  
<https://doi.org/10.1007/978-3-319-72389-1>

Library of Congress Control Number: 2017961795

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

© Springer International Publishing AG 2017

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, express 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.

Printed on acid-free paper

This Springer imprint is published by Springer Nature  
The registered company is Springer International Publishing AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

The 10th International Conference on Security, Privacy, and Anonymity in Computation, Communication and Storage (SpaCCS 2017) was held in Guangzhou, China, during December 12–15, 2017, and was jointly organized by Guangzhou University and Central South University.

The SpaCCS conference series provides a forum for researchers to gather and share their research findings, achievements, innovations, and perspectives in information security and related fields. Previous SpaCCS conferences were held in Zhangjiajie, China (2016), Helsinki, Finland (2015), Beijing, China (2014), Melbourne, Australia (2013), Liverpool, UK (2012), and Changsha, China (2011).

This year, the conference received 140 submissions. All submissions were reviewed by at least three reviewers in a high-quality review process. Based on the recommendations of the reviewers and discussions by the Program Committee members, 47 papers were selected for oral presentation at the conference and inclusion in this Springer volume (i.e., an acceptance rate of 33.6%). In addition to the technical presentations, the program included 12 keynote speeches by world-renowned researchers. We are very grateful to the keynote speakers for their time and willingness to share their expertise with the conference attendees. The keynote speakers were: Prof. Geoffrey Charles Fox, Prof. Lajos Hanzo, Prof. Azzedine Boukerche, Prof. Jie Wu, Prof. Robert Deng, Prof. Ljiljana Trajkovic, Prof. Kin K. Leung, Prof. Vijay Varadharajan, Prof. Hai Jin, Prof. Jianhua Ma, Prof. Jinjun Chen, and Xiaofeng Chen. A big thank you to all of them!

SpaCCS 2017 was made possible by the joint effort of a large number of individuals and organizations worldwide. There is a long list of people who volunteered their time and energy to put together the conference and who deserve special thanks. First and foremost, we would like to offer our gratitude to the Steering Committee chairs, Prof. Guojun Wang from Guangzhou University, China, and Prof. Gregorio Martinez from University of Murcia, Spain, for guiding the entire process of the conference. We are also deeply grateful to all the Program Committee members for their time and effort in reading, commenting, debating, and finally selecting the papers. We also thank all the external reviewers for assisting the Program Committee in their particular areas of expertise.

We would like to offer our gratitude to the general chairs, Prof. Robert Deng, Prof. Yang Xiang, and Prof. Jose M. Alcaraz Calero, for their tremendous support and advice in ensuring the success of the conference. Thanks also go to: the workshop chairs, Prof. Georgios Kambourakis, Prof. Ryan Ko, and Prof. Sancheng Peng; the publicity chairs, Prof. Carlos Becker Westphall, Prof. Scott Fowler, and Dr. Mir Sajjad Hussain Talpur; the publication chair, Prof. Fang Qi; the organization chairs, Prof. Dongqing Xie, Dr. Shuhong Chen, and Dr. Xiaofei Xing; the registration chair, Ms. Pin Liu; and the Web chair, Mr. Yang Shu.

Last but not the least, we would like to thank all the authors, participants, and session chairs for their valuable contributions. Many of them traveled long distances to attend this conference and contribute to the success of SpaCCS2017.

December 2017

Guojun Wang  
Mohammed Atiquzzaman  
Zheng Yan  
Kim-Kwang Raymond Choo

# Organization

## Executive General Chair

Guojun Wang                      Guangzhou University, China

## General Chairs

Robert Deng                      Singapore Management University, Singapore  
Yang Xiang                         Deakin University, Australia  
Jose M. Alcaraz Calero            University of the West of Scotland, UK

## Program Chairs

Mohammed Atiquzzaman        University of Oklahoma, USA  
Zheng Yan                          Aalto University, Finland  
Kim-Kwang                         University of Texas at San Antonio, USA  
    Raymond Choo

## Program Vice Chairs

Christian Esposito                University of Salerno, Italy  
Jin Li                                  Guangzhou University, China  
Marinella Petrocchi                Istituto di Informatica e Telematica, CNR, Italy  
Zeeshan Pervez                      University of the West of Scotland, UK  
Sabu M. Thampi                      Indian Institute of Information Technology  
    and Management, India  
Wenjun Jiang                         Hunan University, China

## Program Committee

### (1) Security Track

#### Program Vice Chairs

Christian Esposito                University of Salerno, Italy  
Jin Li                                  Guangzhou University, China

#### Technical Program Committee

Mohammad Aazam                 Carnegie Mellon University, Qatar  
Habtamu Abie                        Norwegian Computing Center, Norway  
Avishek Adhikari                    University of Calcutta, India



Hamid Alasadi	Basra University, Iraq
Abdul Halim Ali	Universiti Kuala Lumpur, Malaysia
Flora Amato	University of Naples Federico II, Italy
Moreno Ambrosin	University of Padua, Italy
Jose Andre Morales	Carnegie Mellon University, USA
Kamran Arshad	Ajman University and University of Greenwich, UK
Zubair Baig	Edith Cowan University, Australia
Muneer Bani Yassein	Jordan University of Science and Technology, Jordan
Cataldo Basile	Politecnico di Torino, Italy
Jorge Bernal Bernabe	University of Murcia, Spain
Simona Bernardi	Centro Universitario de la Defensa, Zaragoza, Spain
Sulabh Bhattarai	Dakota State University, USA
Carlo Blundo	University of Salerno, Italy
Nidhal Bouaynaya	Rowan University, USA
Salah Bourennane	Ecole Centrale Marseille, France
Abdelkrim Brahm	École de Technologie Supérieure, Canada
Andrea Bruno	University of Salerno, Italy
Christian Callegari	CNIT, Italy
Eddy Caron	ENS de Lyon, France
Anupam Chattopadhyay	National Technological University, Singapore
Ankit Chaudhary	Northwest Missouri State University, USA
Pin-Yu Chen	AI Foundations, IBM T. J. Watson Research Center, USA
Thomas Chen	City, University of London, UK
Chang Choi	Chosun University, South Korea
Pablo Corral Gonzalez	Universidad Miguel Hernandez, Spain
Stefano Cresci	Istituto di Informatica e Telematica CNR, Italy
Hai Dao	Hanoi University of Industry, Vietnam
Narottam Das	University of Southern Queensland, Australia
Ashok Kumar Das	International Institute of Information Technology, Hyderabad, India, India
Alessandra De Benedictis	University of Naples Federico II, Italy
Isabel de la Torre Díez	University of Valladolid, Spain
Tony de Souza-Daw	Melbourne Polytechnic, Australia
Roberto Di Pietro	University of Padova, Italy
Soufiene Djahel	Manchester Metropolitan University, UK
Oscar Esparza	Universitat Politècnica de Catalunya, Spain
Ed Fernandez	FAU Florida Atlantic University, USA
Massimo Ficco	University of Campania Luigi Vanvitelli, Italy
Dimitris Geneiatakis	European Commission, Joint Research Centre, Italy
Angelo Genovese	University of Milan, Italy
Dieter Gollmann	Hamburg University of Technology, Germany
Ying Guo	Central South University, China
Sheikh Mahbub Habib	TU Darmstadt, Germany
Hovhannes Harutyunyan	Concordia University, Canada
Ragib Hasan	The University of Alabama at Birmingham, USA

Keqiang He	University of Wisconsin-Madison and Google, USA
Yu Hua	Huazhong University of Science and Technology, China
Xinyi Huang	Fujian Normal University, China
Biju Issac	Teesside University, UK
Celestine Iwendi	Bangor College, WSN Consults Ltd., Sweden, China
Hai Jiang	Arkansas State University, USA
Rajgopal Kannan	University of Southern California, USA
George Karakostas	McMaster University, Canada
Kwangjo Kim	KAIST, South Korea
Ryan Ko	University of Waikato, New Zealand
Nikos Komninos	City, University of London, UK
Ram Krishnan	The University of Texas at San Antonio, USA
Xia Li	Qualcomm, USA
Ruidong Li	National Institute of Information and Communications Technology (NICT), Japan
Xin Liao	Hunan University, China
Feng Lin	University of Colorado Denver, USA
Jialin Liu	Lawrence Berkeley National Lab, USA
Pascal Lorenz	University of Haute Alsace, France
Pavel Loskot	Swansea University, UK
Edwin Lughofer	Johannes Kepler University Linz, Austria
Leandros Maglaras	De Montfort University, UK
Wissam Mallouli	Montimage, France
Ahmed Mehaoua	University of Paris Descartes, France
Christoph Meinel	Hasso Plattner Institute, Germany
Aleksandra Mileva	University Goce Delcev, Republic of Macedonia
Nader Mir	San Jose State University, California, USA
Moeiz Miraoui	University of Gafsa, Tunisia
Chadli Mohammed	University of Picardie Jules Verne, France
Robert Morelos-Zaragoza	San Jose State University, USA
Paolo Mori	Istituto di Informatica e Telematica CNR, Italy
Vincenzo Moscato	University of Naples Federico II, Italy
Abhishek Murthy	Philips Lighting Research North America, USA
Roberto Nardone	University of Naples Federico II, Italy
Pouya Ostovari	Temple University, USA
Pouya Ostovari	Temple University, Iran
Keyurkumar Patel	Australian Defence Force, Australia
Sancheng Peng	Guangdong University of Foreign Studies, China
Antonio Pescapè	University of Naples Federico II, Italy
Pedro R. M. Inácio	Universidade da Beira Interior, Portugal
Sherif Rashad	Florida Polytechnic University, USA
Vaibhav Rastogi	Northwestern University, USA
Mubashir Husain Rehmani	COMSATS Institute of Information Technology, Pakistan
Roberto Rojas-Cessa	New Jersey Institute of Technology, USA

Imed Romdhani	Edinburgh Napier University, UK
Bimal Roy	Indian Statistical Institute, India
Walid Saad	Virginia Tech, USA
Fahad Saeed	WMU, USA
Altair Santin	Pontifical Catholic University of Parana, Brazil
Andrea Saracino	Istituto di Informatica e Telematica CNR, Italy
Damien Sauveron	University of Limoges, France
Frank Schulz	SAP SE, Germany
Ricky Sethi	Fitchburg State University, USA
Meral Shirazipour	Ericsson Research, USA
Patrick Siarry	Université Paris-Est Creteil, France
Nicolas Sklavos	University of Patras, Greece
Houbing Song	Embry-Riddle Aeronautical University, USA
Martin Strohmeier	University of Oxford, UK
Chunhua Su	University of Aizu, Japan
Chang-ai Sun	University of Science and Technology Beijing, China
Ljiljana Trajkovic	Simon Fraser University, Canada
Chaitanya S. K. Vadrevu	Amazon, USA
Quoc-Tuan Vien	Middlesex University, UK
Artemios Voyiatzis	SBA Research, Austria
Yong Wang	Dakota State University, USA
Jie Wang	University of Massachusetts Lowell, USA
Hejun Wu	Sun Yat-sen University, China
Yongdong Wu	Institute for Infocomm Research, Singapore
Yang Xiao	The University of Alabama, USA
Ping Yang	State University of New York at Binghamton, USA
Ping Yang	SUNY at Binghamton, USA
Muneer Masadeh	Jordan University of Science and Technology, Jordan
Bani Yassein	
Nicolas Younan	Mississippi State University, USA
Chau Yuen	Singapore University of Technology and Design, Singapore
Go Yun II	HWUM, Malaysia
Nicola Zannone	Eindhoven University of Technology, The Netherlands
Xiao Zhang	Syracuse University, USA
Qingchen Zhang	St. Francis Xavier University, Canada
Xinliang Zheng	Frostburg State University, USA
Congxu Zhu	Central South University, China
Natasa Zivic	University of Siegen, Germany

## (2) Privacy Track

### Program Vice Chairs

Marinella Petrocchi                      Istituto di Informatica e Telematica, CNR, Italy  
 Zeeshan Pervez                              University of the West of Scotland, UK

### Technical Program Committee

Wajahat Ali Khan                              Kyung Hee University, South Korea  
 Panagiotis Andriotis                          University of the West of England, UK  
 Bruhadeshwar Bezawada                      Mahindra Ecole Centrale, India  
 Arcangelo Castiglione                          University of Salerno, Italy  
 Sudip Chakraborty                              Valdosta State University, USA  
 Cheng Kang Chu                                Huawei, Singapore  
 Gianpiero Costantino                          Istituto di Informatica e Telematica CNR, Italy  
 Vittoria Cozza                                  University of Padua, Italy  
 Sabrina De Capitani di  
     Vimercati                                      University of Milan, Italy  
 Josep Domingo-Ferrer                          Universitat Rovira i Virgili, Spain  
 Subrata Dutta                                    Haldia Institute of Technology, India  
 Muhammad Fahim                                Istanbul Sabahhatin Zaim University, Turkey  
 Ugo Fiore                                        University of Naples Federico II, Italy  
 Sara Foresti                                      University of Milan, Italy  
 Felix J. Garcia Clemente                      University of Murcia, Spain  
 Saurabh Garg                                    University of Tasmania, India  
 Abdul Ghafoor                                  Acreo Swedish ICT, Sweden  
 Michele Girolami                                CNR- ISTI, Italy  
 Donghai Guan                                  Nanjing University of Aeronautics and Astronautics,  
     China  
 Yao Guo                                        Peking University, China  
 Selena He                                        Kennesaw State University, USA  
 Xiaojun Hei                                      Huazhong University of Science and Technology,  
     China  
 Patrick Hung                                    University of Ontario Institute of Technology, Canada  
 Abdessamad Imine                              Lorraine University, France  
 Farkhund Iqbal                                 Zayed University, United Arab Emirates  
 Murtuza Jadliwala                              Wichita State University, USA  
 Vana Kalogeraki                                Athens University of Economics and Business, Greece  
 Zaheer Khan                                    University of the West of England, UK  
 Marek Klonowski                                TU Wroclaw, Poland  
 Xin Li    Nanjing University of Aeronautics and Astronautics,  
     China  
 Chi Lin                                         Dalian University of Technology, China  
 Giovanni Livraga                                University of Milan, Italy  
 Songtao Lu                                      Iowa State University, USA  
 Pietro Manzoni                                 Polytechnic University of Valencia, Spain

Guazzone Marco	University of Piemonte Orientale, Italy
Asad Masood Khattak	Zayed University, United Arab Emirates
Iliaria Matteucci	Istituto di Informatica e Telematica CNR, Italy
Guerroumi Mohamed	USTHB University, Algeria
Vincenzo Piuri	University of Milan, Italy
Raffaele Pizzolante	University of Salerno, Italy
Abdul Razzaque	University of the West of Scotland, UK
Vincent Roca	Inria, France
Madhu Kumar S. D	NIT Calicut, India
Jun Shen	University of Wollongong, Australia
Angelo Spognardi	University of Rome 1, Italy
Traian Marius Truta	Northern Kentucky University, USA
Damla Turgut	University of Central Florida, USA
Omaid Uthmani	Glasgow Caledonian University, UK
Shuang Wang	University of California San Diego, USA
Mingzhong Wang	University of the Sunshine Coast, Australia
Yuanzhang Xiao	Northwestern University, USA
Xiaolong Xu	Nanjing University of Posts and Telecommunications, China
Xuanxia Yao	University of Science and Technology Beijing, China
Baoliu Ye	Nanjing University, China
Yu Yong	Shaanxi Normal University, China
Sherali Zeadally	University of Kentucky, USA
Mingwu Zhang	Hubei University of Technology, China
Youwen Zhu	Nanjing University of Aeronautics and Astronautics, China

### **(3) Applications Track**

#### **Program Vice Chairs**

Sabu M. Thampi	Indian Institute of Information Technology and Management, India
Wenjun Jiang	Hunan University, China

#### **Technical Program Committee**

Alfredo Cuzzocrea	University of Trieste, Italy
Kalman Graffi	Heinrich Heine University of Düsseldorf, Germany
Dimitrios Karrs	Stereia Hellas Institute of Technology, Greece
Mirco Marchetti	University of Modena and Reggio Emilia, Italy
Juan Pedro Munoz-Gea	Universidad Politécnica de Cartagena, Spain
Thinagaran Perumal	Universiti Putra Malaysia, Malaysia
Antonio Ruiz-Martínez	University of Murcia, Spain
Jorge S. A. Silva	University of Coimbra, Portugal
Saratha Sathasivam	Universiti Sains Malaysia, Malaysia
Junggab Son	Kennesaw State University, USA

Chao Song	University of Electronic Science and Technology of China, China
Yunwei Zhao	Nanyang Technological University, China
Ping Zhou	Qualcomm Inc., USA

### Steering Committee

Guojun Wang	Guangzhou University, China (Chair)
Gregorio Martinez	University of Murcia, Spain (Chair)
Jemal H. Abawajy	Deakin University, Australia
Jose M. Alcaraz Calero	University of the West of Scotland, UK
Yiannong Cao	Hong Kong Polytechnic University, Hong Kong, SAR China
Hsiao-Hwa Chen	National Cheng Kung University, Taiwan
Jinjun Chen	University of Technology, Sydney, Australia
Kim-Kwang Raymond Choo	University of Texas at San Antonio, USA
Robert Deng	Singapore Management University, Singapore
Mario Freire	The University of Beira Interior, Portugal
Minyi Guo	Shanghai Jiao Tong University, China
Weijia Jia	Shanghai Jiao Tong University, China
Wei Jie	University of West London, UK
Georgios Kambourakis	University of the Aegean, Greece
Ryan Ko	University of Waikato, New Zealand
Constantinos Koliass	George Mason University, USA
Jianbin Li	Central South University, China
Jie Li	University of Tsukuba, Japan
Jianhua Ma	Hosei University, Japan
Felix Gomez Marmol	University of Murcia, Spain
Geyong Min	University of Exeter, UK
Peter Mueller	IBM Zurich Research Laboratory, Switzerland
Indrakshi Ray	Colorado State University, USA
Kouichi Sakurai	Kyushu University, Japan
Juan E. Tapiador	Carlos III University of Madrid, Spain
Sabu M. Thampi	Indian Institute of Information Technology and Management, India
Jie Wu	Temple University, USA
Yang Xiao	The University of Alabama, USA
Yang Xiang	Deakin University, Australia
Zheng Yan	Aalto University, Finland
Laurence T. Yang	St. Francis Xavier University, Canada
Wanlei Zhou	Deakin University, Australia

## Workshop Chairs

Georgios Kambourakis	University of the Aegean, Samos, Greece
Ryan Ko	University of Waikato, New Zealand
Sancheng Peng	Guangdong University of Foreign Studies, China

## Publicity Chairs

Carlos Becker Westphall	Federal University of Santa Catarina, Brazil
Scott Fowler	Linköping University, Sweden
Mir Sajjad Hussain Talpur	Sindh Agriculture University, Pakistan

## Publication Chair

Fang Qi	Central South University, China
---------	---------------------------------

## Organizing Chairs

Dongqing Xie	Guangzhou University, China
Shuhong Chen	Guangzhou University, China
Xiaofei Xing	Guangzhou University, China

## Registration Chair

Pin Liu	Central South University, China
---------	---------------------------------

## Web Chair

Yang Shu	Central South University, China
----------	---------------------------------

## Sponsors



# Contents

Optimized Analysis Based on Improved Mutation and Crossover Operator for Differential Evolution Algorithm. . . . .	1
<i>Zhenlan Liu, Jian-bin Li, and Qiang Song</i>	
Revisiting Localization Attacks in Mobile App People-Nearby Services. . . . .	17
<i>Jialin Wang, Hanni Cheng, Minhui Xue, and Xiaojun Hei</i>	
Lengthening Unidimensional Continuous-Variable Quantum Key Distribution with Noiseless Linear Amplifier . . . . .	31
<i>Yu Cao, Jianwu Liang, and Ying Guo</i>	
Research on Internet of Vehicles' Privacy Protection Based on Tamper-Proof with Ciphertext . . . . .	42
<i>Qifan Wang, Guihua Duan, Entao Luo, and Guojun Wang</i>	
An Attack to an Anonymous Certificateless Group Key Agreement Scheme and Its Improvement . . . . .	56
<i>Xuefei Cao, Lanjun Dang, Kai Fan, and Yulong Fu</i>	
A Space Efficient Algorithm for LCIS Problem . . . . .	70
<i>Daxin Zhu and Xiaodong Wang</i>	
Improving the Efficiency of Dynamic Programming in Big Data Computing . . . . .	78
<i>Xiaodong Wang and Daxin Zhu</i>	
Traceable and Complete Fine-Grained Revocable Multi-authority Attribute-Based Encryption Scheme in Social Network . . . . .	87
<i>Yanmei Li, Fang Qi, and Zhe Tang</i>	
The All Seeing Eye: Web to App Intercommunication for Session Fingerprinting in Android. . . . .	93
<i>Efthimios Alepis and Constantinos Patsakis</i>	
An Efficient Hierarchical Identity-Based Encryption Scheme for the Key Escrow . . . . .	108
<i>Yuanlong Li, Fang Qi, and Zhe Tang</i>	
An Improved Pre-copy Transmission Algorithm in Mobile Cloud Computing . . . . .	121
<i>Xianfei Huang, Nao Wang, and Gaocai Wang</i>	



Motivation of DDOS Attack-Aware Link Assignment between Switches to SDN Controllers . . . . . 131  
*Sameer Ali, Saw Chin Tan, Lee Ching Kwang, Zulfadzli Yusoff, Reazul Haque, Ir. Rizaludin Kaspin, and Salvatore Renato Ziri*

TIM: A Trust Insurance Mechanism for Network Function Virtualization Based on Trusted Computing . . . . . 139  
*Guangwu Xu, Yankun Tang, Zheng Yan, and Peng Zhang*

Personalized Semantic Location Privacy Preservation Algorithm Based on Query Processing Cost Optimization . . . . . 153  
*Mengzhen Xu, Hongyun Xu, and Cheng Xu*

Smartphone Bloatware: An Overlooked Privacy Problem . . . . . 169  
*Haroon Elahi, Guojun Wang, and Xu Li*

An ECC-Based Off-line Anonymous Grouping-Proof Protocol . . . . . 186  
*Zhibin Zhou, Pin Liu, Qin Liu, and Guojun Wang*

PCSD: A Tool for Android Malware Detection. . . . . 201  
*Bo Leng, Jianbin Li, Yang Xu, Liang She, Wuqiang Gao, and Quanrun Zeng*

Authorship Analysis of Social Media Contents Using Tone and Personality Features. . . . . 212  
*Athira Usha and Sabu M. Thampi*

Privacy-Preserving Handover Authentication Protocol from Lightweight Identity-Based Signature for Wireless Networks . . . . . 229  
*Changji Wang, Shengyi Jiang, and Yuan Yuan*

Spatial Outlier Information Hiding Algorithm Based on Complex Transformation . . . . . 241  
*Zhaoyu Shou, Akang Liu, Simin Li, and Xiawei Cheng*

A Reputation Model Considering Repurchase Behavior and Mechanism Design to Promote Repurchase . . . . . 256  
*Yuan Liu, Jin Bai, Guibing Guo, Xingwei Wang, and Zhenhua Tan*

Chinese Named Entity Recognition Based on B-LSTM Neural Network with Additional Features . . . . . 269  
*Liubo Ouyang, Yuan Tian, Hui Tang, and Boyun Zhang*

Amplified Locality-Sensitive Hashing for Privacy-Preserving Distributed Service Recommendation. . . . . 280  
*Lianyong Qi, Wanchun Dou, Xuyun Zhang, and Shui Yu*

Learn to Accelerate Identifying New Test Cases in Fuzzing . . . . . 298  
*Weiwei Gong, Gen Zhang, and Xu Zhou*

Service Selection Based on User Privacy Risk Evaluation . . . . . 308  
*Mingdong Tang, Sumeng Zeng, Jianxun Liu, and Buqing Cao*

An Efficient Lattice-Based Proxy Signature with Message Recovery . . . . . 321  
*Faguo Wu, Wang Yao, Xiao Zhang, and Zhiming Zheng*

FABAC: A Flexible Fuzzy Attribute-Based Access Control Mechanism. . . . . 332  
*Yang Xu, Wuqiang Gao, Quanrun Zeng, Guojun Wang, Ju Ren, and Yaoxue Zhang*

Security Analysis and Improvement of An Anonymous Attribute-Based Proxy Re-encryption . . . . . 344  
*Hongjian Yin and Leyou Zhang*

Relacha: Using Associative Meaning for Image  
 Captcha Understandability . . . . . 353  
*Songjie Wei, Qianqian Wu, and Milin Ren*

Identification of Natural Images and Computer Generated Graphics  
 Based on Multiple LBPs in Multicolor Spaces . . . . . 368  
*Fei Peng, Xiao-hua Hu, and Min Long*

A Formal Android Permission Model Based on the B Method . . . . . 381  
*Lu Ren, Rui Chang, Qing Yin, and Yujia Man*

S-SurF: An Enhanced Secure Bulk Data Dissemination  
 in Wireless Sensor Networks . . . . . 395  
*Jian Shen, Tiantian Miao, Qi Liu, Sai Ji, Chen Wang, and Dengzhi Liu*

MCloud: Efficient Monitoring Framework for Cloud  
 Computing Platforms. . . . . 409  
*Jijun Zeng, Zhenyue Long, Guiquan Shen, Lihao Wei, and Yunkui Song*

Secure Vibration Control of Flexible Arms Based on Operators’ Behaviors. . . 420  
*Jiantao Li, Hua Deng, and Wenjun Jiang*

A New Color Image Encryption Scheme Based on Chaotic Hénon  
 Map and Lü System . . . . . 432  
*Chong Fu, Gao-yuan Zhang, Bei-li Gao, Jing Sun, and Xue Wang*

A Distributed Authentication Protocol Using Identity-Based Encryption  
 and Blockchain for LEO Network. . . . . 446  
*Shuai Li, Meilin Liu, and Songjie Wei*

A Detection System for Distributed DoS Attacks Based on Automatic Extraction of Normal Mode and Its Performance Evaluation . . . . .	461
<i>Yaokai Feng, Yoshiaki Hori, and Kouichi Sakurai</i>	
A Unified Model for Detecting Privacy Leakage on Android . . . . .	474
<i>Xueqi Ren, Xinming Wang, Hua Tang, Zhaohui Ma, Jiechao Wu, and Gansen Zhao</i>	
Multi-party Security Computation with Differential Privacy over Outsourced Data . . . . .	487
<i>Ping Li, Heng Ye, and Jin Li</i>	
REW-SMT: A New Approach for Rewriting XACML Request with Dynamic Big Data Security Policies . . . . .	501
<i>Ha Xuan Son, Tran Khanh Dang, and Fabio Massacci</i>	
Decoupling Security Services from IaaS Cloud Through Remote Virtual Machine Introspection . . . . .	516
<i>Huaizhe Zhou, Haihe Ba, Jiangchun Ren, Yongjun Wang, Zhiying Wang, and Yunshi Li</i>	
Privacy Preserving Hierarchical Clustering over Multi-party Data Distribution . . . . .	530
<i>Mina Sheikhalishahi and Fabio Martinelli</i>	
Improving MQTT by Inclusion of Usage Control . . . . .	545
<i>Antonio La Marra, Fabio Martinelli, Paolo Mori, Athanasios Rizos, and Andrea Saracino</i>	
Using JSON to Specify Privacy Preserving-Enabled Attribute-Based Access Control Policies . . . . .	561
<i>Que Nguyet Tran Thi, Tran Khanh Dang, Huy Luong Van, and Ha Xuan Son</i>	
Comprehensive Diversity in Recommender Systems . . . . .	571
<i>Tranos Zuva and Raoul Kwuimi</i>	
Towards Intelligent System Wide Information Management for Air Traffic Management . . . . .	584
<i>Li Weigang, Alessandro F. Leite, Vitor F. Ribeiro, Jose A. Fregnani, and Italo R. de Oliveira</i>	
A Security Risk Management Model for Cloud Computing Systems: Infrastructure as a Service . . . . .	594
<i>Mouna Jouini and Latifa Ben Arfa Rabai</i>	
<b>Author Index . . . . .</b>	<b>609</b>

# Optimized Analysis Based on Improved Mutation and Crossover Operator for Differential Evolution Algorithm

Zhenlan Liu<sup>1</sup>, Jian-bin Li<sup>2(✉)</sup>, and Qiang Song<sup>1</sup>

<sup>1</sup> School of Information Science and Engineering,  
Central South University, Changsha 410083, China

<sup>2</sup> Institute of Information Security and Big Data,  
Central South University, Changsha 410083, China  
lijianbin@csu.edu.cn

**Abstract.** Differential evolution algorithm is a better algorithm for global numeric optimization in the evolution algorithm. The excellent individual in the evolution algorithm contains more abundant information. We propose a mutation method in order of cosine function in order to ensure that the information of such excellent individual will be better inherited and avoid prematurity of the algorithm. Firstly, vector selection probability is calculated according to the cosine computational model in the paper. Secondly, the individual is selected in the sorting order of probability value. Higher sorting position will lead to higher probability of selection. In addition, existing differential evolution algorithms rarely deal with population distribution information. Instead, they only allow for the distribution information of a generation in population evolution. The result is that the performance of the differential evolution algorithm will be affected due to the insufficiency of population information. For this problem, we allow for multi-generation cumulative distribution information of the population in the algorithm and perform eigen decomposition for the covariance matrix. The eigenvector so generated is used to establish the characteristic coordinate system. Crossover operation of the algorithm is performed in the new coordinate system. For such improvement operation, we use IEEECEC2014 as the test function. The experimental results show that this improved algorithm has more improved performance than existing improved DE algorithms.

**Keywords:** Differential evolution · Sorting algorithm · Eigen decomposition  
Population multi-generation cumulative distribution information  
Covariance matrix

## 1 Introduction

The differential evolution algorithm (DE) is a simple and effective method of searching uncertainty proposed by Storn and Price [1, 2] in 1995. Simple principles and fewer controlled parameters of DE make it applicable to random and parallel global search. It has also yielded good results in artificial intelligence and pattern recognition fields.

DE has defects including prone to prematurity, low degree of convergence as other evolution algorithms though it is advantageous in solving function optimization problems. To improve DE performance, many scholars have optimized and improved DE algorithm, such as the recent JADE [3], jDE [4] algorithms. Fan and Lampinen [5] proposed a new type of triangular mutation operation for DE mutation operation. Kaelo and Ali [6] improved mutation operation with attraction-repulsion concept of electromagnetism-like mechanism (EM). However, both improvement ways lead to increased algorithm difficulty and also offer slow convergence. The mutation method based on sorting algorithm enhances the search capability of DE while avoiding prematurity of DE.

DE is a population-based optimization algorithm. Despite of the fact, population distribution information has not been widely applied to DE, which results in much poor performance of DE in solving certain highly complex optimization problems [7]. For this, literature [8, 9] applies population distribution information to DE. However, it only allows for the population distribution information of a single generation and omits the multi-generation cumulative distribution information of the population in the process of evolution. For convenience of computing the multi-generation cumulative distribution information of the population in the process of evolution, we use CMA-ES model of Hansen and Ostermeier [10]. In the model, the covariance matrix will self-adaptively to update the covariance matrix according to the distribution information of prior and present population. The covariance matrix will then be subject to eigen decomposition. The eigenvector arising from decomposition are used to establish a new coordinate system where the crossover operation of the differential improved algorithm is performed. The experiment shows that such crossover pattern offers higher convergence rate and search accuracy compared to traditional ways.

## 2 Related Researches

### 2.1 Classical DE Algorithm

DE algorithm includes three basic operations: mutation, crossover and selection. During the evolution, DE will generate a test vector for each target vector through mutation and crossover operations. The controlled parameters of DE algorithm mainly include: number of population (NP), scaling factor (F) and crossover rate (CR). NP reflects the number of population information in the algorithm. CR reflects the number of information exchanged among offspring, parent and intermediate mutant in the process of crossover. Higher CR value leads to higher degree of information content exchange. In turn, smaller CR will enable population diversity to decline quickly, which is unfavorable for global optimization. Scaling factor F has more impact on the performance of algorithm compared to CR. F mainly affects the global optimization capability of the algorithm. The implementation steps of DE are as follows [11]:

For the global optimization problem under which  $n$  continuous variable(s) should be solved, such global optimization problem may be converted to the minimum problem under which the following functions should be solved:

$$\begin{cases} \min & f(x), x = [x_{[1]}, x_{[2]}, \dots, x_{[n]}] \\ \text{s.t.} & a_j \leq x_j \leq b_j, j = 1, 2, \dots, D \end{cases} \quad (1)$$

where:  $D$  denotes the dimensionality of the problem space solution.  $b_j$  and  $a_j$  denote the upper and lower limits of  $x_j$ , respectively.

**Initial Population.** The initial population is randomly generated:

$$\{x_i(0)|x_i(0) = [x_{i1}, x_{i2}, x_{i3}, \dots, x_{iD}], i = 1, 2, \dots, NP\} \quad (2)$$

$$\begin{cases} x_{ij} = a_j + \text{rand} * (b_j - a_j) \\ i = 1, 2, \dots, NP, j = 1, 2, \dots, D \end{cases} \quad (3)$$

where:  $NP$  denotes the number of population.  $x_i(0)$  denotes individual in the initial population.  $x_{ij}$  denotes component  $j$  of individual  $i$ .  $\text{rand}$  denotes the random number evenly distributed in  $(0, 1)$  interval.

**Mutation.** Mutation operation plays a vital role in DE. The idea is that mutation operation will be implemented by the differential mode. Literature [12, 13] has come up with many mutation operation modes. To make a distinction among the characteristics of different mutation operations in DE algorithm, we describe it on a “DE/a/b” basis. DE denotes differential evolution algorithm.  $a$  denotes the vector subject to mutation operation.  $b$  denotes the number of differential vector(s). Some common mutation operations in DE algorithm are as follows:

“DE/rand/1”

$$v_i(g+1) = x_{r1}(g) + F \cdot (x_{r2}(g) - x_{r3}(g)) \quad (4)$$

“DE/rand/2”

$$v_i(g+1) = x_{r1}(g) + F \cdot (x_{r2}(g) - x_{r3}(g)) + F \cdot (x_{r4}(g) - x_{r5}(g)) \quad (5)$$

“DE/current-to-best/1”

$$v_i(g+1) = x_i(g) + F \cdot (x_{best}(g) - x_i(g)) + F \cdot (x_{r2}(g) - x_{r3}(g)) \quad (6)$$

“DE/rand-to-best/1”

$$v_i(g+1) = x_{r1}(g) + F \cdot (x_{best}(g) - x_{r1}(g)) + F \cdot (x_{r2}(g) - x_{r3}(g)) \quad (7)$$

where:  $x_{best}(g)$  denotes the optimal individual in present population.  $r1 \neq r2 \neq r3$   
 $r4 \neq r5 \neq 1$ ,  $i = 1, 2, \dots, NP$ ,  $r1, r2, r3, r4$  and  $r5$  refer to random integers in  $[1, NP]$  interval.

**Crossover.** Crossover operation is performed for the individuals of population generation  $g$   $\{x_i(g), i = 1, 2, \dots, NP\}$  and its mutant intermediate population  $\{v_i(g+1), i = 1, 2, \dots, NP\}$ :

$$u_{ij}(g+1) = \begin{cases} v_{ij}(g+1), & \text{if } rand \leq CR \text{ or } j = j_{rand} \\ x_{ij}(g), & \text{otherwise} \end{cases} \quad (9)$$

where:  $i = 1, 2, \dots, NP$ ,  $j = 1, 2, \dots, D$ ,  $rand$  denotes random number evenly distributed in  $(0, 1)$  interval,  $u_i(g+1) = [u_{i1}, u_{i2}, u_{i3}, \dots, u_{iD}]$  denotes individual  $i$  in the new population of generation  $g+1$ ,  $u_{ij}(g+1)$  and  $v_{ij}(g+1)$  denote component  $j$  of  $u_{ij}(g+1)$  and  $v_{ij}(g+1)$ , respectively,  $CR$  denotes crossover rate, and  $j_{rand}$  is the random integer in  $[1, D]$  interval. Such crossover strategy ensures that at least one component of  $u_i(g+1)$  is provided by the component of  $v_i(g+1)$ .

**Selection Operation.** DE algorithm selects the individuals that will be added to new population with the greedy strategy according to the size of target function:

$$x_i(g+1) = \begin{cases} u_i(g+1), & \text{if } f(u_i(g+1)) \leq f(x_i(g)) \\ x_i(g), & \text{otherwise} \end{cases} \quad (10)$$

where:  $i = 1, 2, \dots, NP$ .

## 2.2 Current Situation of Researches

With researches for nearly two decades, DE algorithm have been greatly improved and optimized. The research findings of the last several years are briefed as follows:

Zhou et al. [14] came up with a kind of crossover and mutation operation with which the individuals in the population are divided superior and inferior populations according to the fitness function; the focus is not provision of reasonable parameters. Hu et al. [15] proposed subspace clustering mutation algorithm. Such mutation mode was combined with five common mutation modes of DE algorithm, and an optimal individual was selected from the individuals generated from mutation to be used as the center of disturbance. The differential vectors of two edge individuals randomly generated were used as the disturbance amplitude. The results show that such mutation

mode offers higher convergence rate than traditional ways. Gong and Cai [16] present a kind of mutation operation based on sorting philosophy. The selection was made by individual sorting order. This means that the individual with higher sorting order will be more probably selected. However, the probability computing way of the linear mode in mutation operation is not an optimal algorithm.

Gong et al. [17] put forward the crossover rate adjustment method based on feasibility parameter for the adaptive DE algorithm by analyzing crossover operation. Such new method was combined with different DE improved algorithms. The experimental results show that such method improves the performance of the original DE improved algorithm. Sarker et al. [8] defined parameters F, CR and NP (number of population). The optimal values of the three parameters in DE were selected according to the changing new mechanism in the process of evolution. However, the most leading-edge DE algorithms rarely deal with population distribution information. Guo [18] first add population distribution information to the research on DE algorithm in 2015 and used covariance matrix to calculate the distribution information of the population. Then, the eigenvector obtained from eigen decomposition was used for establishment of the characteristic coordinate system. Lastly, DE crossover operation was performed in the characteristic coordinate system to generate the test vector. However, such method only allowed for the population distribution information of a generation in the process of evolution. Inspired by this, we come up with an improved algorithm as follows.

### 3 Improved DE Algorithm

According to the development of present research, our focus on the research is to optimize and modify mutation and population multi-generation cumulative distribution information in DE algorithm. Firstly, we present a new cosine computational model according to vector characteristics. The selection probability of the vector is calculated with such model. Secondly, we select the individual in order of probability values. Higher sorting place will result in higher probability of selection. In addition, existing differential evolution algorithms rarely deal with population distribution information, and they only allow for the distribution information of a generation in population. The result will be that population information is inadequate. For this problem, we add multi-generation cumulative distribution information of the population in the algorithm and update the covariance matrix with the cumulated distribution information. The eigenvector so generated is used to establish the characteristic coordinate system. Crossover operation of the algorithm is performed in the new coordinate system.

#### 3.1 Mutation Operation Based on Cosine Function Sorting

Mutation operation plays a key role in DE performance. Mutation operation generate variable vector. Generally, the parent individual in mutation operation is randomly



selected from present population. For example, for the classical “DE/rand/1” evolution model, three parent vectors  $X_{r_1}$ ,  $X_{r_2}$  and  $X_{r_3}$  are selected from present population. The footnotes  $r_1$ ,  $r_2$ ,  $r_3$  should be  $r_1, r_2, r_3 \in \{1, NP\}$  and  $r_1 \neq r_2 \neq r_3 \neq i$ . NP denotes the number of population. Given the randomly selected parent population, DE will be favorable for global search but offers low convergence rate, and it may sometimes be caught in locally optimal solution.

Therefore, enhanced search capability of DE algorithm may improve the performance of DE algorithm. In practice, excellent individual contains more favorable information and will be more prone to be selected for the generation of offspring individual.

**Improved Mutation Operation.** In the improved mutation operation based on sorting presented in this paper, the vectors in the population will be proportionately selected according to their sorting position. Each vector will be ranked according to the fitness function. First, the population is sorted on an ascending basis according to the fitness value of each vector. Vector sorting may be set as follows:

$$R_i = NP - i, \quad i = 1, 2, \dots, NP \quad (11)$$

where: NP denotes the number of population. According to Formula (11), the optimal vector in present population will get the maximum sorting value.

After setting of the sorting value of each vector, the selection probability value of vector  $i$  ( $x_i$ ) may be calculated as follows:

$$p_i = 0.5 \cdot \left( 1.0 - \cos\left(\frac{R_i \cdot \pi}{NP}\right) \right), \quad i = 1, 2, \dots, NP \quad (12)$$

The terminal vectors of base vector and difference vector will be selected according to the sorting order of the probability value, while other vectors in mutation operation will be randomly selected according to the traditional DE algorithm. Obviously, in the improved and optimized mutation operation, vectors with higher sorting position are more prone to be selected as the terminal vector of either base vector or difference vector. We do not select the starting vector according to the probability value for the search amplitude of difference vector will take a sharp decline and even lead to premature convergence if two vectors behind and after the difference vector are selected from optimum vectors.

**Algorithm Implementation.** “DE/rand/1” mutation model is taken as an example. The ideas of algorithm implementation are described with pseudo-code as follows:

```

1: Generate the initial population randomly
2: Evaluate the fitness for each individual in the population
3: while the stop criterion is not satisfied do
4: Sort the population based on the fitness of each individual
5: Calculate the selection probability for each individual according to (14)
6: for i=1 to NP do
7: Randomly select  $r_1 \in \{1, NP\}$  {base vector index}
8: while  $\text{rndreal}[0,1) > p_{r1}$  or  $r_1 == i$  do
9: Randomly select  $r_1 \in \{1, NP\}$ 
10: end while
11: Randomly select  $r_2 \in \{1, NP\}$  {terminal vector index}
12: while  $\text{rndreal}[0,1) > p_{r2}$  or  $r_2 == r_1$  or  $r_2 == i$  do
13: Randomly select  $r_2 \in \{1, NP\}$ 
14: end while
15: Randomly select  $r_3 \in \{1, NP\}$ 
16: while  $r_3 == r_2$  or  $r_3 == r_1$  or  $r_3 == i$  do
17: Randomly select  $r_3 \in \{1, NP\}$ 
18: end while
19:  $j_{\text{rand}} = \text{rndint}(1, D)$ 
20: for j=1 to D do
21: if  $\text{rndreal}_j[0,1) < C_r$  or j is equal to  $j_{\text{rand}}$  then
22:  $u_{i,j} = x_{r1,j} + F \cdot (x_{r2,j} - x_{r3,j})$ 
23: else  $u_{i,j} = x_{i,j}$ 
24: end if end for end for
25: for i=1 to NP do
26: Evaluate the offspring  $u_i$ 
27: if  $f(u_i)$  is better than or equal to  $f(x_i)$  then
28: Replace  $x_i$  with  $u_i$ 
29: end if end for end while

```

According to pseudo-code analysis, the time complexity of population sorting  $O(NP \cdot \log(NP))$  and of probability computation  $O(NP)$  may be obtained.  $O(G \cdot NP \cdot D)$  is the time complexity of DE algorithm, where  $G$  denotes the maximal of evolution algebra.  $O(G \cdot NP \cdot (D + \log(NP) + 1))$  is the time complexity of mutation operation based on cosine function sorting. However, in DE algorithm, the number of population  $NP$  is generally set proportionately according to problem dimensionality  $D$ . On this account, the time complexity of mutation operation based on cosine function sorting may approximate to  $O(G \cdot D^2)$ . Such time complexity does not increase time overhead as that of traditional DE algorithms and many DE improved algorithms.

Apparently, the mutation operation algorithm based on cosine function sorting is advantageous in simple structure and easy to use as traditional DE algorithms. In addition, better individual will be more prone to be selected as some vectors in mutation operation are selected based on its own sorting. With this, the search capability of DE algorithm enhances greatly. Further, the mutation operation algorithm based on cosine function sorting may be also integrated with other improved DE algorithms, making it highly universal.

### 3.2 Crossover

**The application of population multi-generation cumulative distribution information in DE.** Inspired by CMA-ES [21] model, we improve the probability of successful information search in subsequent evolution process by updating the covariance matrix with population multi-generation cumulative distribution information so as to generate more rational search behavior.

In CMA-ES model, an individual of the population may be generated with Formula (13):

$$\vec{x}_i^{(g+1)} = \vec{m}^{(g)} + \sigma^{(g)} N\left(0, C^{(g)}\right), i = 1, \dots, \lambda \quad (13)$$

where:  $\vec{m}^{(g)}$  is the mean vector of the search area, and  $C^{(g)}$  is covariance matrix.

CMA-ES model is mainly to calculate the parameter values of  $\vec{m}^{(g+1)}$ ,  $\sigma^{(g+1)}$ ,  $C^{(g+1)}$  of generation  $g + 1$ . However, ES and DE have different search modes. In ES, offspring is generated at pre-defined probability. In DE, however, individual of the next generation is generated according to the arithmetic operation of base vector and difference vector and the target vector and variable vector. It is therefore that there is no necessary to update parameter  $\sigma^{(g)}$  in Formula (13). Our work is only to update and adjust the covariance matrix.

According to the characteristics of the updating algorithm and DE algorithm in CMA-ES model, we use rank-u-update as the update strategy to adapt to covariance matrix. The covariance matrix  $C^{(g)}$  is initialized is  $C^{(0)} = I$ , where  $I \in \mathbb{R}^{D \times D}$ . In addition, the mean vector  $\vec{m}^{(g)}$  of distribution searched is initialized to a random point in the search space. In generation  $g + 1$ ,  $\vec{m}^{(g+1)}$  is updated by Formula (14):

$$\vec{m}^{(g+1)} = \sum_{i=1}^{NP} w_i \vec{x}_{i:2*NP}^{(g+1)} \quad (14)$$

where:  $\vec{x}_{i:2*NP}^{(g+1)}$  is the optimal individual of offspring population, which means

$$f\left(\vec{x}_{1:2*NP}^{(g+1)}\right) \leq f\left(\vec{x}_{2:2*NP}^{(g+1)}\right) \leq \dots \leq f\left(\vec{x}_{NP:2*NP}^{(g+1)}\right), w_i \text{ is the positive weight coefficient } i, \text{ and } \sum_{i=1}^{NP} w_i = 1. \text{ Obviously, } \vec{m}^{(g+1)}$$

is the weighted mean of NP optimal individual of offspring population. To introduce the search shift tending to the feasible domain, the weight coefficient is made to rely on a single individual, i.e.  $w_1 \geq w_2 \geq \dots \geq w_{NP} > 0$ , where:

$$w_i = \frac{w'_i}{\sum_{j=1}^{NP} w'_j}, i \in 1, \dots, NP \quad (15)$$

$$w'_i = \ln(NP + 0.5) - \ln(i), i \in 1, \dots, NP \quad (16)$$

$$C_{NP}^{(g+1)} = \sum_{i=1}^{NP} w_i \left( \vec{x}_{i:2*NP}^{(g+1)} - \vec{m}^{(g)} \right) \left( \vec{x}_{i:2*NP}^{(g+1)} - \vec{m}^{(g)} \right)^T \quad (17)$$

$$C^{(g+1)} = (1 - c_{NP})C^{(g)} + c_{NP} \left( \sigma^{(g)^2} \right)^{-1} C_{NP}^{(g+1)} \quad (18)$$

where:  $c_{NP} \approx \min(1, NP_{eff}/D^2)$  is learning rate, and  $NP_{eff} = \left( \sum_{i=1}^{NP} w_i^2 \right)^{-1}$  is the valid selection domain of variance. In Formula (18), it allows for the distribution information of present population and also takes into account the population cumulative distribution information of prior generators. Parameter  $\sigma^{(g)}$  updating is of no significance. To simplify the operation,  $\sigma^{(g)} = 1$  is set in this paper. Such setting also shows that the weights of each generation for covariance matrix are the same.

**Eigen Decomposition of Covariance Matrix.** Eigen decomposition is a form under which the matrix is decomposed to the matrix product expressed in eigenvalue and eigenvector. Crossover operation is performed in the characteristic coordinate system. Firstly, eigenvalue is decomposed with the covariance matrix to generate a set of orthonormal base of eigenvector. Such orthonormal base is used to establish the characteristic coordinate system. The target vector and the variable vector are converted and adjusted so that they will adapt to the new characteristic coordinate system. Secondly, crossover operation is performed for the target vector and the variable vector generated by conversion and adjustment in the new coordinate system. In this way, a new test vector is generated in the established characteristic coordinate system. Finally, such test vector is converted and adjusted to the initial coordinate system.

Table 1. Improved experimental data table based on DE/best/1/bin

Cec2014 test functions	DE/best/1/bin		CDI-DE/best/1/bin		DE/best/1/bin		CDI + FD-DE/best/1/bin		DE/best/1/bin		CDI + FD + VCO-DE/best/1/bin	
	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.
F <sub>1</sub>	2.15E + 07 $\pm$ 7.73E + 06	1.42E + 07 $\pm$ 7.73E + 06	1.42E + 07 $\pm$ 7.73E + 06	1.42E + 07 $\pm$ 7.73E + 06	2.15E + 07 $\pm$ 7.73E + 06	2.15E + 07 $\pm$ 7.73E + 06	1.42E + 07 $\pm$ 7.73E + 06	1.42E + 07 $\pm$ 7.73E + 06	2.15E + 07 $\pm$ 7.73E + 06	2.15E + 07 $\pm$ 7.73E + 06	1.42E + 07 $\pm$ 7.73E + 06	1.42E + 07 $\pm$ 7.73E + 06
F <sub>2</sub>	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14	2.27E - 14 $\pm$ 1.27E - 14
F <sub>3</sub>	6.82E - 13 $\pm$ 3.19E - 13	3.41E - 14 $\pm$ 3.11E - 14	3.41E - 14 $\pm$ 3.11E - 14	6.82E - 13 $\pm$ 3.19E - 13	6.82E - 13 $\pm$ 3.19E - 13	6.82E - 13 $\pm$ 3.19E - 13	4.54E - 14 $\pm$ 2.54E - 14	4.54E - 14 $\pm$ 2.54E - 14	6.82E - 13 $\pm$ 3.19E - 13	6.82E - 13 $\pm$ 3.19E - 13	5.68E - 14 $\pm$ 0.00E + 00	5.68E - 14 $\pm$ 0.00E + 00
F <sub>4</sub>	2.73E + 01 $\pm$ 3.29E + 01	3.41E - 14 $\pm$ 3.11E - 14	3.41E - 14 $\pm$ 3.11E - 14	2.73E + 01 $\pm$ 3.29E + 01	2.73E + 01 $\pm$ 3.29E + 01	2.73E + 01 $\pm$ 3.29E + 01	4.54E - 14 $\pm$ 2.54E - 14	4.54E - 14 $\pm$ 2.54E - 14	2.73E + 01 $\pm$ 3.29E + 01	2.73E + 01 $\pm$ 3.29E + 01	3.41E - 14 $\pm$ 3.11E - 14	3.41E - 14 $\pm$ 3.11E - 14
F <sub>5</sub>	2.08E + 01 $\pm$ 4.78E - 02	2.09E + 01 $\pm$ 6.34E - 02	2.09E + 01 $\pm$ 6.34E - 02	2.08E + 01 $\pm$ 4.78E - 02	2.08E + 01 $\pm$ 4.78E - 02	2.08E + 01 $\pm$ 4.78E - 02	2.09E + 01 $\pm$ 4.08E - 02	2.09E + 01 $\pm$ 4.08E - 02	2.08E + 01 $\pm$ 4.78E - 02	2.08E + 01 $\pm$ 4.78E - 02	2.09E + 01 $\pm$ 3.35E - 02	2.09E + 01 $\pm$ 3.35E - 02
F <sub>6</sub>	8.18E - 01 $\pm$ 5.92E - 01	1.37E + 00 $\pm$ 8.95E - 01	1.37E + 00 $\pm$ 8.95E - 01	8.18E - 01 $\pm$ 5.92E - 01	8.18E - 01 $\pm$ 5.92E - 01	8.18E - 01 $\pm$ 5.92E - 01	4.19E - 01 $\pm$ 5.75E - 01	4.19E - 01 $\pm$ 5.75E - 01	8.18E - 01 $\pm$ 5.92E - 01	8.18E - 01 $\pm$ 5.92E - 01	4.79E - 01 $\pm$ 4.71E - 01	4.79E - 01 $\pm$ 4.71E - 01
F <sub>7</sub>	2.46E - 03 $\pm$ 5.50E - 03	2.27E - 14 $\pm$ 5.08E - 14	2.27E - 14 $\pm$ 5.08E - 14	2.46E - 03 $\pm$ 5.50E - 03	2.46E - 03 $\pm$ 5.50E - 03	2.46E - 03 $\pm$ 5.50E - 03	9.09E - 14 $\pm$ 5.08E - 14	9.09E - 14 $\pm$ 5.08E - 14	2.46E - 03 $\pm$ 5.50E - 03	2.46E - 03 $\pm$ 5.50E - 03	9.09E - 14 $\pm$ 5.08E - 14	9.09E - 14 $\pm$ 5.08E - 14
F <sub>8</sub>	6.16E + 00 $\pm$ 4.06E + 00	6.76E + 00 $\pm$ 1.91E + 00	6.76E + 00 $\pm$ 1.91E + 00	6.16E + 00 $\pm$ 4.06E + 00	6.16E + 00 $\pm$ 4.06E + 00	6.16E + 00 $\pm$ 4.06E + 00	7.38E + 00 $\pm$ 1.76E + 00	7.38E + 00 $\pm$ 1.76E + 00	6.16E + 00 $\pm$ 4.06E + 00	6.16E + 00 $\pm$ 4.06E + 00	5.96E + 00 $\pm$ 2.39E + 00	5.96E + 00 $\pm$ 2.39E + 00
F <sub>9</sub>	1.66E + 02 $\pm$ 1.29E + 01	1.75E + 02 $\pm$ 4.17E + 00	1.75E + 02 $\pm$ 4.17E + 00	1.66E + 02 $\pm$ 1.29E + 01	1.66E + 02 $\pm$ 1.29E + 01	1.66E + 02 $\pm$ 1.29E + 01	1.67E + 02 $\pm$ 5.88E + 00	1.67E + 02 $\pm$ 5.88E + 00	1.66E + 02 $\pm$ 1.29E + 01	1.66E + 02 $\pm$ 1.29E + 01	1.18E + 02 $\pm$ 1.91E + 01	1.18E + 02 $\pm$ 1.91E + 01
F <sub>10</sub>	1.54E + 02 $\pm$ 1.37E + 02	1.28E + 03 $\pm$ 1.00E + 03	1.28E + 03 $\pm$ 1.00E + 03	1.54E + 02 $\pm$ 1.37E + 02	1.54E + 02 $\pm$ 1.37E + 02	1.54E + 02 $\pm$ 1.37E + 02	7.60E + 02 $\pm$ 7.56E + 02	7.60E + 02 $\pm$ 7.56E + 02	1.54E + 02 $\pm$ 1.37E + 02	1.54E + 02 $\pm$ 1.37E + 02	8.77E + 02 $\pm$ 5.13E + 02	8.77E + 02 $\pm$ 5.13E + 02
F <sub>11</sub>	6.40E + 03 $\pm$ 1.94E + 02	6.43E + 03 $\pm$ 2.75E + 02	6.43E + 03 $\pm$ 2.75E + 02	6.40E + 03 $\pm$ 1.94E + 02	6.40E + 03 $\pm$ 1.94E + 02	6.40E + 03 $\pm$ 1.94E + 02	6.47E + 03 $\pm$ 2.30E + 02	6.47E + 03 $\pm$ 2.30E + 02	6.40E + 03 $\pm$ 1.94E + 02	6.40E + 03 $\pm$ 1.94E + 02	6.46E + 03 $\pm$ 1.43E + 02	6.46E + 03 $\pm$ 1.43E + 02
F <sub>12</sub>	1.86E + 00 $\pm$ 1.33E - 01	2.12E + 00 $\pm$ 1.57E - 01	2.12E + 00 $\pm$ 1.57E - 01	1.86E + 00 $\pm$ 1.33E - 01	1.86E + 00 $\pm$ 1.33E - 01	1.86E + 00 $\pm$ 1.33E - 01	2.17E + 00 $\pm$ 2.24E - 01	2.17E + 00 $\pm$ 2.24E - 01	1.86E + 00 $\pm$ 1.33E - 01	1.86E + 00 $\pm$ 1.33E - 01	2.11E + 00 $\pm$ 1.75E - 01	2.11E + 00 $\pm$ 1.75E - 01
F <sub>13</sub>	4.04E - 01 $\pm$ 4.80E - 02	3.10E - 01 $\pm$ 3.40E - 02	3.10E - 01 $\pm$ 3.40E - 02	4.04E - 01 $\pm$ 4.80E - 02	4.04E - 01 $\pm$ 4.80E - 02	4.04E - 01 $\pm$ 4.80E - 02	3.25E - 01 $\pm$ 7.09E - 02	3.25E - 01 $\pm$ 7.09E - 02	4.04E - 01 $\pm$ 4.80E - 02	4.04E - 01 $\pm$ 4.80E - 02	3.28E - 01 $\pm$ 2.34E - 02	3.28E - 01 $\pm$ 2.34E - 02
F <sub>14</sub>	4.20E - 01 $\pm$ 1.95E - 01	2.49E - 01 $\pm$ 3.22E - 02	2.49E - 01 $\pm$ 3.22E - 02	4.20E - 01 $\pm$ 1.95E - 01	4.20E - 01 $\pm$ 1.95E - 01	4.20E - 01 $\pm$ 1.95E - 01	2.48E - 01 $\pm$ 4.83E - 02	2.48E - 01 $\pm$ 4.83E - 02	4.20E - 01 $\pm$ 1.95E - 01	4.20E - 01 $\pm$ 1.95E - 01	2.69E - 01 $\pm$ 3.38E - 02	2.69E - 01 $\pm$ 3.38E - 02
F <sub>15</sub>	1.64E + 01 $\pm$ 1.15E + 00	1.51E + 01 $\pm$ 7.36E - 01	1.51E + 01 $\pm$ 7.36E - 01	1.64E + 01 $\pm$ 1.15E + 00	1.64E + 01 $\pm$ 1.15E + 00	1.64E + 01 $\pm$ 1.15E + 00	1.51E + 01 $\pm$ 1.28E + 00	1.51E + 01 $\pm$ 1.28E + 00	1.64E + 01 $\pm$ 1.15E + 00	1.64E + 01 $\pm$ 1.15E + 00	1.52E + 01 $\pm$ 6.42E - 01	1.52E + 01 $\pm$ 6.42E - 01
F <sub>16</sub>	1.22E + 01 $\pm$ 3.38E - 01	1.21E + 01 $\pm$ 1.26E - 01	1.21E + 01 $\pm$ 1.26E - 01	1.22E + 01 $\pm$ 3.38E - 01	1.22E + 01 $\pm$ 3.38E - 01	1.22E + 01 $\pm$ 3.38E - 01	1.22E + 01 $\pm$ 2.19E - 01	1.22E + 01 $\pm$ 2.19E - 01	1.22E + 01 $\pm$ 3.38E - 01	1.22E + 01 $\pm$ 3.38E - 01	1.23E + 01 $\pm$ 2.49E - 01	1.23E + 01 $\pm$ 2.49E - 01
F <sub>17</sub>	4.74E + 05 $\pm$ 1.78E + 05	4.98E + 02 $\pm$ 2.09E + 02	4.98E + 02 $\pm$ 2.09E + 02	4.74E + 05 $\pm$ 1.78E + 05	4.74E + 05 $\pm$ 1.78E + 05	4.74E + 05 $\pm$ 1.78E + 05	5.94E + 02 $\pm$ 3.02E + 02	5.94E + 02 $\pm$ 3.02E + 02	4.74E + 05 $\pm$ 1.78E + 05	4.74E + 05 $\pm$ 1.78E + 05	5.27E + 02 $\pm$ 5.00E + 02	5.27E + 02 $\pm$ 5.00E + 02
F <sub>18</sub>	3.69E + 02 $\pm$ 2.63E + 02	4.15E + 01 $\pm$ 3.70E + 01	4.15E + 01 $\pm$ 3.70E + 01	3.69E + 02 $\pm$ 2.63E + 02	3.69E + 02 $\pm$ 2.63E + 02	3.69E + 02 $\pm$ 2.63E + 02	4.05E + 01 $\pm$ 2.20E + 01	4.05E + 01 $\pm$ 2.20E + 01	3.69E + 02 $\pm$ 2.63E + 02	3.69E + 02 $\pm$ 2.63E + 02	4.73E + 01 $\pm$ 2.61E + 01	4.73E + 01 $\pm$ 2.61E + 01
F <sub>19</sub>	6.30E + 00 $\pm$ 1.00E + 00	4.52E + 00 $\pm$ 9.03E - 01	4.52E + 00 $\pm$ 9.03E - 01	6.30E + 00 $\pm$ 1.00E + 00	6.30E + 00 $\pm$ 1.00E + 00	6.30E + 00 $\pm$ 1.00E + 00	5.13E + 00 $\pm$ 9.45E - 01	5.13E + 00 $\pm$ 9.45E - 01	6.30E + 00 $\pm$ 1.00E + 00	6.30E + 00 $\pm$ 1.00E + 00	4.70E + 00 $\pm$ 9.35E - 01	4.70E + 00 $\pm$ 9.35E - 01
F <sub>20</sub>	7.74E + 01 $\pm$ 7.93E + 00	3.14E + 01 $\pm$ 1.29E + 01	3.14E + 01 $\pm$ 1.29E + 01	7.74E + 01 $\pm$ 7.93E + 00	7.74E + 01 $\pm$ 7.93E + 00	7.74E + 01 $\pm$ 7.93E + 00	2.13E + 01 $\pm$ 9.83E + 00	2.13E + 01 $\pm$ 9.83E + 00	7.74E + 01 $\pm$ 7.93E + 00	7.74E + 01 $\pm$ 7.93E + 00	2.90E + 01 $\pm$ 1.19E + 01	2.90E + 01 $\pm$ 1.19E + 01
+	6	4	4	6	4	4	3	3	6	4	3	3
-	15	17	17	15	17	17	20	20	15	17	20	20
$\approx$	9	9	9	9	9	9	7	7	9	9	7	7

Table 2. Improved experimental data table based on JADE

CEC2014 test functions	JADE		CDI-JADE		JADE		CDI + FD-JADE		JADE		CDI + FD + VCO-JADE	
	Mean error	Std. Dev.	Mean error	Std. Dev.	Mean error	Std. Dev.	Mean error	Std. Dev.	Mean error	Std. Dev.	Mean error	Std. Dev.
F <sub>1</sub>	1.44E + 02 ± 1.55E + 02	1.42E - 14 ± 0.00E + 00	1.44E + 02 ± 1.55E + 02	1.44E + 02 ± 1.55E + 02	1.13E - 14 ± 6.35E - 15	1.13E - 14 ± 6.35E - 15	1.44E + 02 ± 1.55E + 02	1.44E + 02 ± 1.55E + 02	1.06E - 14 ± 6.31E - 15	1.06E - 14 ± 6.31E - 15	1.44E + 02 ± 1.55E + 02	1.44E + 02 ± 1.55E + 02
F <sub>2</sub>	2.27E - 14 ± 1.27E - 14	2.84E - 14 ± 0.00E + 00	2.27E - 14 ± 1.27E - 14	2.27E - 14 ± 1.27E - 14	1.70E - 14 ± 1.55E - 14	1.70E - 14 ± 1.55E - 14	2.27E - 14 ± 1.27E - 14	2.27E - 14 ± 1.27E - 14	1.98E - 14 ± 1.33E - 14	1.98E - 14 ± 1.33E - 14	2.27E - 14 ± 1.27E - 14	2.27E - 14 ± 1.27E - 14
F <sub>3</sub>	4.27E - 08 ± 9.53E - 08	3.41E - 14 ± 3.11E - 14	4.27E - 08 ± 9.53E - 08	4.27E - 08 ± 9.53E - 08	1.13E - 14 ± 2.54E - 14	1.13E - 14 ± 2.54E - 14	4.27E - 08 ± 9.53E - 08	4.27E - 08 ± 9.53E - 08	2.27E - 14 ± 2.85E - 14	2.27E - 14 ± 2.85E - 14	4.27E - 08 ± 9.53E - 08	4.27E - 08 ± 9.53E - 08
F <sub>4</sub>	2.84E - 13 ± 3.87E - 13	4.54E - 14 ± 2.54E - 14	2.84E - 13 ± 3.87E - 13	2.84E - 13 ± 3.87E - 13	1.26E + 01 ± 2.83E + 01	1.26E + 01 ± 2.83E + 01	2.84E - 13 ± 3.87E - 13	2.84E - 13 ± 3.87E - 13	5.40E - 14 ± 1.27E - 14	5.40E - 14 ± 1.27E - 14	2.84E - 13 ± 3.87E - 13	2.84E - 13 ± 3.87E - 13
F <sub>5</sub>	2.02E + 01 ± 2.65E - 02	2.03E + 01 ± 3.16E - 02	2.02E + 01 ± 2.65E - 02	2.02E + 01 ± 2.65E - 02	1.09E + 01 ± 2.65E - 02	1.09E + 01 ± 2.65E - 02	2.02E + 01 ± 2.65E - 02	2.02E + 01 ± 2.65E - 02	2.09E + 01 ± 1.44E - 01	2.09E + 01 ± 1.44E - 01	2.02E + 01 ± 2.65E - 02	2.02E + 01 ± 2.65E - 02
F <sub>6</sub>	1.11E + 01 ± 1.19E + 00	3.81E + 00 ± 3.17E + 00	1.11E + 01 ± 1.19E + 00	1.11E + 01 ± 1.19E + 00	1.11E + 01 ± 1.19E + 00	1.11E + 01 ± 1.19E + 00	1.11E + 01 ± 1.19E + 00	1.11E + 01 ± 1.19E + 00	5.43E + 00 ± 5.09E + 00	5.43E + 00 ± 5.09E + 00	1.11E + 01 ± 1.19E + 00	1.11E + 01 ± 1.19E + 00
F <sub>7</sub>	4.54E - 14 ± 6.22E - 14	0.00E + 00 ± 0.00E + 00	4.54E - 14 ± 6.22E - 14	4.54E - 14 ± 6.22E - 14	6.82E - 14 ± 6.22E - 14	6.82E - 14 ± 6.22E - 14	4.54E - 14 ± 6.22E - 14	4.54E - 14 ± 6.22E - 14	6.25E - 14 ± 5.80E - 14	6.25E - 14 ± 5.80E - 14	4.54E - 14 ± 6.22E - 14	4.54E - 14 ± 6.22E - 14
F <sub>8</sub>	0.00E + 00 ± 0.00E + 00	4.54E - 14 ± 6.22E - 14	0.00E + 00 ± 0.00E + 00	0.00E + 00 ± 0.00E + 00	0.00E + 00 ± 0.00E + 00	0.00E + 00 ± 0.00E + 00	0.00E + 00 ± 0.00E + 00	0.00E + 00 ± 0.00E + 00	3.74E + 01 ± 6.73E + 00	3.74E + 01 ± 6.73E + 00	0.00E + 00 ± 0.00E + 00	0.00E + 00 ± 0.00E + 00
F <sub>9</sub>	2.38E + 01 ± 5.52E + 00	2.17E + 01 ± 7.77E + 00	2.38E + 01 ± 5.52E + 00	2.38E + 01 ± 5.52E + 00	1.11E + 01 ± 5.52E + 00	1.11E + 01 ± 5.52E + 00	2.38E + 01 ± 5.52E + 00	2.38E + 01 ± 5.52E + 00	1.39E + 01 ± 7.80E + 00	1.39E + 01 ± 7.80E + 00	2.38E + 01 ± 5.52E + 00	2.38E + 01 ± 5.52E + 00
F <sub>10</sub>	4.16E - 03 ± 9.31E - 03	5.35E - 01 ± 1.48E - 01	4.16E - 03 ± 9.31E - 03	4.16E - 03 ± 9.31E - 03	2.91E - 03 ± 1.89E - 02	2.91E - 03 ± 1.89E - 02	4.16E - 03 ± 9.31E - 03	4.16E - 03 ± 9.31E - 03	2.85E - 03 ± 3.96E - 02	2.85E - 03 ± 3.96E - 02	4.16E - 03 ± 9.31E - 03	4.16E - 03 ± 9.31E - 03
F <sub>11</sub>	1.58E + 03 ± 2.13E + 02	2.16E + 03 ± 6.04E + 01	1.58E + 03 ± 2.13E + 02	1.58E + 03 ± 2.13E + 02	3.15E + 03 ± 3.51E + 02	3.15E + 03 ± 3.51E + 02	1.58E + 03 ± 2.13E + 02	1.58E + 03 ± 2.13E + 02	1.44E + 03 ± 3.46E + 02	1.44E + 03 ± 3.46E + 02	1.58E + 03 ± 2.13E + 02	1.58E + 03 ± 2.13E + 02
F <sub>12</sub>	2.62E - 01 ± 2.23E - 02	1.67E - 01 ± 6.83E - 02	2.62E - 01 ± 2.23E - 02	2.62E - 01 ± 2.23E - 02	1.65E - 01 ± 1.39E - 01	1.65E - 01 ± 1.39E - 01	2.62E - 01 ± 2.23E - 02	2.62E - 01 ± 2.23E - 02	1.26E + 00 ± 5.05E - 01	1.26E + 00 ± 5.05E - 01	2.62E - 01 ± 2.23E - 02	2.62E - 01 ± 2.23E - 02
F <sub>13</sub>	2.04E - 01 ± 3.20E - 02	2.11E - 01 ± 3.70E - 02	2.04E - 01 ± 3.20E - 02	2.04E - 01 ± 3.20E - 02	2.52E - 01 ± 2.59E - 02	2.52E - 01 ± 2.59E - 02	2.04E - 01 ± 3.20E - 02	2.04E - 01 ± 3.20E - 02	1.74E - 01 ± 4.49E - 02	1.74E - 01 ± 4.49E - 02	2.04E - 01 ± 3.20E - 02	2.04E - 01 ± 3.20E - 02
F <sub>14</sub>	2.38E - 01 ± 8.08E - 03	2.08E - 01 ± 3.77E - 02	2.38E - 01 ± 8.08E - 03	2.38E - 01 ± 8.08E - 03	2.34E - 01 ± 5.53E - 02	2.34E - 01 ± 5.53E - 02	2.38E - 01 ± 8.08E - 03	2.38E - 01 ± 8.08E - 03	1.99E - 01 ± 3.07E - 02	1.99E - 01 ± 3.07E - 02	2.38E - 01 ± 8.08E - 03	2.38E - 01 ± 8.08E - 03
F <sub>15</sub>	3.36E + 00 ± 5.18E - 01	3.46E + 00 ± 5.37E - 01	3.36E + 00 ± 5.18E - 01	3.36E + 00 ± 5.18E - 01	1.32E + 00 ± 7.93E - 01	1.32E + 00 ± 7.93E - 01	3.36E + 00 ± 5.18E - 01	3.36E + 00 ± 5.18E - 01	2.37E + 00 ± 6.51E - 01	2.37E + 00 ± 6.51E - 01	3.36E + 00 ± 5.18E - 01	3.36E + 00 ± 5.18E - 01
F <sub>16</sub>	9.58E + 00 ± 2.68E - 01	9.68E + 00 ± 2.34E - 01	9.58E + 00 ± 2.68E - 01	9.58E + 00 ± 2.68E - 01	1.11E + 01 ± 2.46E - 01	1.11E + 01 ± 2.46E - 01	9.58E + 00 ± 2.68E - 01	9.58E + 00 ± 2.68E - 01	1.12E + 00 ± 3.83E - 01	1.12E + 00 ± 3.83E - 01	9.58E + 00 ± 2.68E - 01	9.58E + 00 ± 2.68E - 01
F <sub>17</sub>	1.08E + 03 ± 2.94E + 02	1.35E + 03 ± 3.24E + 02	1.08E + 03 ± 2.94E + 02	1.08E + 03 ± 2.94E + 02	0.85E + 03 ± 5.16E + 02	0.85E + 03 ± 5.16E + 02	1.08E + 03 ± 2.94E + 02	1.08E + 03 ± 2.94E + 02	1.19E + 02 ± 3.99E + 02	1.19E + 02 ± 3.99E + 02	1.08E + 03 ± 2.94E + 02	1.08E + 03 ± 2.94E + 02
F <sub>18</sub>	9.25E + 01 ± 4.85E + 01	1.10E + 02 ± 3.85E + 01	9.25E + 01 ± 4.85E + 01	9.25E + 01 ± 4.85E + 01	7.33E + 01 ± 3.20E + 01	7.33E + 01 ± 3.20E + 01	9.25E + 01 ± 4.85E + 01	9.25E + 01 ± 4.85E + 01	1.10E + 01 ± 3.97E + 01	1.10E + 01 ± 3.97E + 01	9.25E + 01 ± 4.85E + 01	9.25E + 01 ± 4.85E + 01
F <sub>19</sub>	4.43E + 00 ± 8.24E - 01	5.11E + 00 ± 1.35E + 00	4.43E + 00 ± 8.24E - 01	4.43E + 00 ± 8.24E - 01	5.58E + 00 ± 2.77E - 01	5.58E + 00 ± 2.77E - 01	4.43E + 00 ± 8.24E - 01	4.43E + 00 ± 8.24E - 01	2.73E + 00 ± 1.17E + 00	2.73E + 00 ± 1.17E + 00	4.43E + 00 ± 8.24E - 01	4.43E + 00 ± 8.24E - 01
F <sub>20</sub>	4.54E + 03 ± 3.24E + 03	1.58E + 01 ± 5.68E + 00	4.54E + 03 ± 3.24E + 03	4.54E + 03 ± 3.24E + 03	6.85E + 01 ± 1.31E + 01	6.85E + 01 ± 1.31E + 01	4.54E + 03 ± 3.24E + 03	4.54E + 03 ± 3.24E + 03	8.23E + 01 ± 3.82E + 01	8.23E + 01 ± 3.82E + 01	4.54E + 03 ± 3.24E + 03	4.54E + 03 ± 3.24E + 03
+	10		9						4			
-	14		17						23			
≈	6		4						3			

Table 3. Improved experimental data table based on jDE

CEC2014 Test functions	jDE		CDI-jDE		jDE		CDI + FD-jDE		jDE		CDI + FD + VCO-jDE	
	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.	Mean error $\pm$ Std. Dev.
F <sub>1</sub>	3.28E + 04 $\pm$ 1.80E + 04	3.95E - 12 $\pm$ 8.34E - 12	3.95E + 00 $\pm$ 0.00E + 00	3.28E + 04 $\pm$ 1.80E + 04	5.46E - 05 $\pm$ 1.22E - 04	3.28E + 04 $\pm$ 1.80E + 04	3.28E + 04 $\pm$ 1.80E + 04	5.46E - 05 $\pm$ 1.22E - 04	3.28E + 04 $\pm$ 1.80E + 04	3.28E + 04 $\pm$ 1.80E + 04	2.26E - 05 $\pm$ 3.22E - 04	2.26E - 05 $\pm$ 3.22E - 04
F <sub>2</sub>	5.68E - 15 $\pm$ 1.27E - 14	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	5.68E - 15 $\pm$ 1.27E - 14	0.00E + 00 $\pm$ 0.00E + 00	5.68E - 15 $\pm$ 1.27E - 14	5.68E - 15 $\pm$ 1.27E - 14	0.00E + 00 $\pm$ 0.00E + 00	5.68E - 15 $\pm$ 1.27E - 14	5.68E - 15 $\pm$ 1.27E - 14	2.17E - 15 $\pm$ 3.34E - 14	2.17E - 15 $\pm$ 3.34E - 14
F <sub>3</sub>	2.27E - 14 $\pm$ 3.11E - 14	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	2.27E - 14 $\pm$ 3.11E - 14	1.06E - 14 $\pm$ 2.74E - 14	2.27E - 14 $\pm$ 3.11E - 14	2.27E - 14 $\pm$ 3.11E - 14	1.06E - 14 $\pm$ 2.74E - 14	2.27E - 14 $\pm$ 3.11E - 14	2.27E - 14 $\pm$ 3.11E - 14	2.04E - 14 $\pm$ 1.36E - 14	2.04E - 14 $\pm$ 1.36E - 14
F <sub>4</sub>	1.41E + 00 $\pm$ 9.16E - 01	1.39E - 01 $\pm$ 1.86E - 01	1.39E - 01 $\pm$ 1.86E - 01	1.41E + 00 $\pm$ 9.16E - 01	4.08E - 02 $\pm$ 7.84E - 02	1.41E + 00 $\pm$ 9.16E - 01	1.41E + 00 $\pm$ 9.16E - 01	4.08E - 02 $\pm$ 7.84E - 02	1.41E + 00 $\pm$ 9.16E - 01	1.41E + 00 $\pm$ 9.16E - 01	5.14E - 02 $\pm$ 6.27E - 03	5.14E - 02 $\pm$ 6.27E - 03
F <sub>5</sub>	2.03E + 01 $\pm$ 3.64E - 02	2.04E + 01 $\pm$ 4.89E - 02	2.04E + 01 $\pm$ 4.89E - 02	2.03E + 01 $\pm$ 3.64E - 02	2.03E + 01 $\pm$ 3.64E - 02	2.03E + 01 $\pm$ 3.64E - 02	2.03E + 01 $\pm$ 3.64E - 02	2.03E + 01 $\pm$ 3.64E - 02	2.03E + 01 $\pm$ 3.64E - 02	2.03E + 01 $\pm$ 3.64E - 02	7.45E + 00 $\pm$ 3.64E - 02	7.45E + 00 $\pm$ 3.64E - 02
F <sub>6</sub>	3.76E + 00 $\pm$ 3.83E + 00	1.24E + 00 $\pm$ 8.45E - 01	1.24E + 00 $\pm$ 8.45E - 01	3.76E + 00 $\pm$ 3.83E + 00	3.76E + 00 $\pm$ 3.83E + 00	3.76E + 00 $\pm$ 3.83E + 00	3.76E + 00 $\pm$ 3.83E + 00	1.08E + 00 $\pm$ 1.15E + 00	3.76E + 00 $\pm$ 3.83E + 00	3.76E + 00 $\pm$ 3.83E + 00	3.01E + 00 $\pm$ 2.43E + 00	3.01E + 00 $\pm$ 2.43E + 00
F <sub>7</sub>	4.54E - 14 $\pm$ 6.22E - 14	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	4.54E - 14 $\pm$ 6.22E - 14	2.83E - 14 $\pm$ 5.30E - 14	4.54E - 14 $\pm$ 6.22E - 14	4.54E - 14 $\pm$ 6.22E - 14	2.83E - 14 $\pm$ 5.30E - 14	4.54E - 14 $\pm$ 6.22E - 14	4.54E - 14 $\pm$ 6.22E - 14	2.11E - 14 $\pm$ 6.27E - 14	2.11E - 14 $\pm$ 6.27E - 14
F <sub>8</sub>	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	0.00E + 00 $\pm$ 0.00E + 00	7.09E - 15 $\pm$ 3.96E - 14	7.09E - 15 $\pm$ 3.96E - 14
F <sub>9</sub>	4.55E + 01 $\pm$ 4.73E + 00	4.04E + 01 $\pm$ 3.31E + 00	4.04E + 01 $\pm$ 3.31E + 00	4.55E + 01 $\pm$ 4.73E + 00	4.55E + 01 $\pm$ 4.73E + 00	4.55E + 01 $\pm$ 4.73E + 00	4.55E + 01 $\pm$ 4.73E + 00	4.09E + 01 $\pm$ 3.90E + 00	4.55E + 01 $\pm$ 4.73E + 00	4.55E + 01 $\pm$ 4.73E + 00	3.13E + 01 $\pm$ 2.75E + 00	3.13E + 01 $\pm$ 2.75E + 00
F <sub>10</sub>	4.20E - 03 $\pm$ 9.30E - 03	2.90E + 00 $\pm$ 7.58E - 01	2.90E + 00 $\pm$ 7.58E - 01	4.20E - 03 $\pm$ 9.30E - 03	4.20E - 03 $\pm$ 9.30E - 03	4.20E - 03 $\pm$ 9.30E - 03	4.20E - 03 $\pm$ 9.30E - 03	3.14E + 00 $\pm$ 8.22E - 01	4.20E - 03 $\pm$ 9.30E - 03	4.20E - 03 $\pm$ 9.30E - 03	4.01E - 03 $\pm$ 6.73E - 01	4.01E - 03 $\pm$ 6.73E - 01
F <sub>11</sub>	2.46E + 03 $\pm$ 2.81E + 02	2.57E + 03 $\pm$ 2.96E + 02	2.57E + 03 $\pm$ 2.96E + 02	2.46E + 03 $\pm$ 2.81E + 02	2.46E + 03 $\pm$ 2.81E + 02	2.46E + 03 $\pm$ 2.81E + 02	2.46E + 03 $\pm$ 2.81E + 02	2.75E + 03 $\pm$ 2.81E + 02	2.46E + 03 $\pm$ 2.81E + 02	2.46E + 03 $\pm$ 2.81E + 02	3.55E + 03 $\pm$ 1.73E + 02	3.55E + 03 $\pm$ 1.73E + 02
F <sub>12</sub>	5.25E - 01 $\pm$ 7.26E - 02	5.60E - 01 $\pm$ 8.04E - 02	5.60E - 01 $\pm$ 8.04E - 02	5.25E - 01 $\pm$ 7.26E - 02	5.25E - 01 $\pm$ 7.26E - 02	5.25E - 01 $\pm$ 7.26E - 02	5.25E - 01 $\pm$ 7.26E - 02	5.60E - 01 $\pm$ 8.47E - 02	5.25E - 01 $\pm$ 7.26E - 02	5.25E - 01 $\pm$ 7.26E - 02	4.06E - 01 $\pm$ 6.22E - 02	4.06E - 01 $\pm$ 6.22E - 02
F <sub>13</sub>	2.85E - 01 $\pm$ 2.38E - 02	2.76E - 01 $\pm$ 4.36E - 02	2.76E - 01 $\pm$ 4.36E - 02	2.85E - 01 $\pm$ 2.38E - 02	2.85E - 01 $\pm$ 2.38E - 02	2.85E - 01 $\pm$ 2.38E - 02	2.85E - 01 $\pm$ 2.38E - 02	2.48E - 01 $\pm$ 3.29E - 02	2.85E - 01 $\pm$ 2.38E - 02	2.85E - 01 $\pm$ 2.38E - 02	1.77E - 01 $\pm$ 4.17E - 03	1.77E - 01 $\pm$ 4.17E - 03
F <sub>14</sub>	3.08E - 01 $\pm$ 2.86E - 02	2.55E - 01 $\pm$ 4.20E - 02	2.55E - 01 $\pm$ 4.20E - 02	3.08E - 01 $\pm$ 2.86E - 02	3.08E - 01 $\pm$ 2.86E - 02	3.08E - 01 $\pm$ 2.86E - 02	3.08E - 01 $\pm$ 2.86E - 02	2.39E - 01 $\pm$ 5.02E - 02	3.08E - 01 $\pm$ 2.86E - 02	3.08E - 01 $\pm$ 2.86E - 02	2.74E - 01 $\pm$ 4.63E - 02	2.74E - 01 $\pm$ 4.63E - 02
F <sub>15</sub>	6.10E + 00 $\pm$ 7.39E - 01	5.98E + 00 $\pm$ 5.71E - 01	5.98E + 00 $\pm$ 5.71E - 01	6.10E + 00 $\pm$ 7.39E - 01	6.10E + 00 $\pm$ 7.39E - 01	6.10E + 00 $\pm$ 7.39E - 01	6.10E + 00 $\pm$ 7.39E - 01	5.59E + 00 $\pm$ 6.46E - 01	6.10E + 00 $\pm$ 7.39E - 01	6.10E + 00 $\pm$ 7.39E - 01	4.73E + 00 $\pm$ 5.44E - 02	4.73E + 00 $\pm$ 5.44E - 02
F <sub>16</sub>	9.80E + 00 $\pm$ 5.34E - 01	1.03E + 01 $\pm$ 2.98E - 01	1.03E + 01 $\pm$ 2.98E - 01	9.80E + 00 $\pm$ 5.34E - 01	9.80E + 00 $\pm$ 5.34E - 01	9.80E + 00 $\pm$ 5.34E - 01	9.80E + 00 $\pm$ 5.34E - 01	1.03E + 01 $\pm$ 3.29E - 01	9.80E + 00 $\pm$ 5.34E - 01	9.80E + 00 $\pm$ 5.34E - 01	7.12E + 00 $\pm$ 3.29E - 01	7.12E + 00 $\pm$ 3.29E - 01
F <sub>17</sub>	4.72E + 02 $\pm$ 4.11E + 02	1.43E + 02 $\pm$ 8.35E + 01	1.43E + 02 $\pm$ 8.35E + 01	4.72E + 02 $\pm$ 4.11E + 02	4.72E + 02 $\pm$ 4.11E + 02	4.72E + 02 $\pm$ 4.11E + 02	4.72E + 02 $\pm$ 4.11E + 02	1.85E + 02 $\pm$ 1.54E + 02	4.72E + 02 $\pm$ 4.11E + 02	4.72E + 02 $\pm$ 4.11E + 02	2.64E + 02 $\pm$ 3.17E + 02	2.64E + 02 $\pm$ 3.17E + 02
F <sub>18</sub>	1.67E + 01 $\pm$ 8.40E + 00	9.73E + 00 $\pm$ 2.56E + 00	9.73E + 00 $\pm$ 2.56E + 00	1.67E + 01 $\pm$ 8.40E + 00	1.67E + 01 $\pm$ 8.40E + 00	1.67E + 01 $\pm$ 8.40E + 00	1.67E + 01 $\pm$ 8.40E + 00	1.02E + 01 $\pm$ 1.82E + 00	1.67E + 01 $\pm$ 8.40E + 00	1.67E + 01 $\pm$ 8.40E + 00	1.19E + 01 $\pm$ 2.32E + 00	1.19E + 01 $\pm$ 2.32E + 00
F <sub>19</sub>	4.62E + 00 $\pm$ 3.00E - 01	4.78E + 00 $\pm$ 5.13E - 01	4.78E + 00 $\pm$ 5.13E - 01	4.62E + 00 $\pm$ 3.00E - 01	4.62E + 00 $\pm$ 3.00E - 01	4.62E + 00 $\pm$ 3.00E - 01	4.62E + 00 $\pm$ 3.00E - 01	3.77E + 00 $\pm$ 5.93E - 01	4.62E + 00 $\pm$ 3.00E - 01	4.62E + 00 $\pm$ 3.00E - 01	2.34E + 00 $\pm$ 4.37E - 01	2.34E + 00 $\pm$ 4.37E - 01
+	7			6		6			3			
-	17			20		20			25			
$\approx$	6			4		4			2			

Decomposition of the eigenvalue of covariance matrix  $C^{(g)}$  of generation may be described as follows:

$$C^{(g)} = B^{(g)} D^{(g)^2} B^{(g)T} \quad (19)$$

In Formula (19), each column of orthogonal matrix  $B^{(g)}$  is the corresponding eigenvector of  $C^{(g)}$ . Each diagonal element of diagonal matrix  $D^{(g)}$  corresponds to the eigenvalue of  $C^{(g)}$ . In this way,  $B^{(g)}$  contains a set of orthogonal base of eigenvector.

$B^{(g)T}$  may transpose a vector to the characteristic coordinate system. After transposition, the elements of the resultant vector will correlate to the projection of eigenvector. Based on these characteristics, target vector  $x_i^{-(g)}$  and its corresponding variable vector  $V_i^{-(g)}$  may be converted with Formulas (20) and (21):

$$x_i^{-(g)} = B^{(g)T} x_i^{(g)} \quad (20)$$

$$V_i^{-(g)} = B^{(g)T} V_i^{(g)} \quad (21)$$

The crossover operation of DE algorithm is performed in the characteristic coordinate system. Test vector  $u_i^{-(g)} = (u_{i,1}^{-(g)}, \dots, u_{i,D}^{-(g)})$  is obtained with Formula (22):

$$u_{i,j}^{-(g)} = \begin{cases} V_{i,j}^{-(g)}, & \text{if } \text{rand}(0, 1) \leq CR \text{ or } j = j_{rand}, j = 1, \dots, D \\ x_{i,j}^{-(g)}, & \text{otherwise} \end{cases} \quad (22)$$

$B^{(g)}$  can transpose and transforms the resulting value to the initial coordinate system. Therefore, the final test vector  $u_i^{-(g)}$  in the initial coordinate system may be converted with Formula (23):

$$u_i^{-(g)} = B^{(g)} u_i^{-(g)} \quad (23)$$

In this paper, the advantage of implementing DE crossover operation in the characteristic coordinate system is described as follows:

In the initial stage of differential evolution, the population keeps its diversity at a higher level. With the cumulative distribution information of the population, the covariance matrix will quickly provide a reasonable search domain and leads the population to keep evolving towards a feasible direction.

In the middle and late stage of evolution, the population undergoes reduced diversity. The search may concentrate on a relatively small area. With covariance matrix, the fitness function will be determined. Implementation of the crossover operation in characteristic coordinate system may enhance the search capability of the population.



## 4 Experimental Analysis

In this paper, firstly, the mutation mode based on cosine function sorting was used in mutation operation. Secondly, the covariance matrix was decomposed based on the consideration of population multi-generation cumulative distribution information. Crossover operation was performed in the newly established characteristic coordinate system. In the following experiment, we used IEEE CEC 2004 as the test function. The classical differential evolution model DE/best/1/bin and recent DE improved algorithms JADE, jDE were used as the representative. The population cumulative distribution information of multi-generations, eigen decomposition and performance of improved algorithms after mutation operation based on cosine function sorting were gradually added based on the experimental analysis of existing algorithm.

During the experiment, the strengths and weakness of the algorithm were measured with the error mean and standard deviation of multiple operation results of the test function. Wilcoxon rank sum test method was used to compare the differences among different algorithms in statistical data. IEEE CEC2014  $F_1$ – $F_3$  are unimodal function,  $F_4$ – $F_{16}$  are simple unimodal function and  $F_{17}$ – $F_{22}$  are mixed function and  $F_{23}$ – $F_{30}$  are compound function. It is highly universal to use CEC2014 as the test function. Symbols “+”, “-”, “ $\approx$ ” were used in the experimental data table to denote the size of relevant experimental data of two comparative algorithms (Tables 1, 2 and 3).

According to the analysis of the experimental data, in unimodal function  $F_1$ – $F_3$ , the improved algorithm had better performance than the original algorithm. In simple unimodal function  $F_4$ – $F_{16}$ , the degree of performance improvement of jDE and JADE algorithms would be higher than that of the classical algorithm DE/best/1/bin provided that it only allowed for the population multi-generation cumulative distribution information. However, with consideration of eigen decomposition, there was no improved performance whether of jDE, JADE or DE/best/1/bin. If the mutation operation based on cosine function sorting was added, the algorithm performance would be qualitatively improved. In compound function  $F_{17}$ – $F_{22}$ , performance of DE/best/1/bin and jDE algorithms improved greatly with consideration of the multi-generation cumulative distribution of the population, which almost approximated to performance optimum limit of algorithm. However, JADE algorithm performance did not gain much improvement, and was beyond improvement even of another two improvement strategies were added on this basis. In compound function, the degree of optimization of original algorithms by the improvement method was lower than that of function  $F_1$ – $F_{22}$ . However, the experimental data showed that function performance would reach its optimal state with consideration of population multi-generation cumulative distribution information, eigen decomposition and mutation operation based on cosine function sorting. This is the evidence that our research field and optimization strategy are right.

An analysis of these experimental data showed that based on the classical DE/best/1/bin and existing improved algorithms jDE, JADE, the performance of algorithms would gradually improve with the involvement of population multi-generation cumulative distribution information, eigen decomposition and mutation operation based on cosine function sorting. In particular, when the three improvement strategies were allowed for, algorithm performance would reach its optimal state.

## 5 Conclusion

This paper presents an optimization and improvement strategy for DE consisting of improvement for both mutation and crossover operations. For mutation operation, it presents the mutation operation based on cosine function sorting. This increases the accuracy of mutation operation while maintaining algorithm simplicity. For crossover operation, a new coordinate system is established in the support of eigen decomposition of covariance matrix. The multi-generation cumulative distribution information of the population is introduced so that data information in the process of evolution will be more plentiful and evolution accuracy will be enhanced. This model overcomes the deficiencies of existing model including insufficient essential data, and prone to be caught in local optimum. It offers much improvement to the evolution algorithm in both accuracy and convergence rate. In particular, it is applicable to the analysis and prediction of large volume of data.

## References

1. Storn, R., Price, K.: Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces. Technical report TR-95-012, Berkeley, CA (1995)
2. Storn, R., Price, K.V.: Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **11**(4), 341–359 (1997)
3. Zhang, J., Sanderson, A.C.: JADE: adaptive differential evolution with optional external archive. *IEEE Trans. Evol. Comput.* **13**(5), 945–958 (2009)
4. Brest, J., Greiner, S., Boskovic, B., Mernik, M., Zumer, V.: Self-adapting control parameters in differential evolution: a comparative study on numerical benchmark problems. *IEEE Trans. Evol. Comput.* **10**(6), 646–657 (2006)
5. Fan, H.-Y., Lampinen, J.: A trigonometric mutation operation to differential evolution. *J. Glob. Optim.* **27**(1), 105–129 (2003)
6. Kaelo, P., Ali, M.M.: Differential evolution algorithms using hybrid mutation. *Comput. Optim. Appl.* **37**(2), 231–246 (2007)
7. Wang, Y., Liu, Z.-Z., Li, J., Li, H.-X., Yen, G.G.: Utilizing cumulative population distribution information in differential evolution. *Appl. Soft Comput.* **48**, 329–346 (2016)
8. Guo, S., Yang, C.: Enhancing differential evolution utilizing Eigenvector-based crossover operator. *IEEE Trans. Evol. Comput.* **19**(1), 31–49 (2015)
9. Wang, Y., Li, H.-X., Huang, T., Li, L.: Differential evolution based on covariance matrix learning and bimodal distribution parameter setting. *Appl. Soft Comput.* **18**, 232–247 (2014)
10. Hansen, N., Ostermeier, A.: Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* **9**(2), 159–195 (2001)
11. Price, K., Storn, R., Lampinen, J.: *Differential Evolution: A Practical Approach to Global Optimization*. Springer, Berlin (2005). <https://doi.org/10.1007/3-540-31306-0>
12. Storn, R., Price, K.: Home Page of Differential Evolution. International Computer Science Institute, Berkeley (2010)
13. Zhou, Y., Li, X., Gao, L.: A differential evolution algorithm with intersect mutation operator. *Appl. Soft Comput.* **13**(1), 390–401 (2013)
14. Hu, Z., Xiong, S., Wang, X., Su, Q., Liu, M., Chen, Z.: Subspace clustering mutation operator for developing convergent differential evolution algorithm. *Math. Probl. Eng.* **2014**, 18 (2014). (Article ID 154626)

15. Gong, W., Cai, Z.: Differential evolution with ranking-based mutation operators. *IEEE Trans. Cybern.* **43**(6), 2066–2081 (2013)
16. Gong, W., Cai, Z., Wang, Y.: Repairing the crossover rate in adaptive differential evolution. *Appl. Soft Comput.* **15**, 149–168 (2014)
17. Sarker, R., Elsayed, S., Ray, T.: Differential evolution with dynamic parameters selection for optimization problems. *IEEE Trans. Evol. Comput.* **18**(5), 689–707 (2014)
18. Li, Y.L., Zhan, Z.H., Gong, Y.J., Chen, W.N., Zhang, J., Li, Y.: Differential evolution with an evolution path: a deep evolutionary algorithm. *IEEE Trans. Cybern.* **45**(9), 1798–1810 (2015)

# Revisiting Localization Attacks in Mobile App People-Nearby Services

Jialin Wang<sup>1</sup>, Hanni Cheng<sup>1</sup>, Minhui Xue<sup>2,3</sup>, and Xiaojun Hei<sup>1</sup>(✉)

<sup>1</sup> Huazhong University of Science and Technology, Wuhan 430074, China  
{wangjialin,chenghn,heixj}@hust.edu.cn

<sup>2</sup> East China Normal University, Shanghai 200062, China

<sup>3</sup> NYU Shanghai, Shanghai 200122, China  
minhuixue@nyu.edu

**Abstract.** The widespread use of people-nearby services has spawned the development of social discovery applications that help users make new friends with nearby users (such as WeChat). Unfortunately, malicious third-parties can often deploy trilateration attacks to exploit people-nearby applications to determine the exact locations of target users, therefore compromising their privacy. In this paper, we revisit these localization attacks and propose a new two-step localization method that boosts the accuracy of the state of the art for the contemporary location-based social network (LBSN) services which have adopted the band-distance obfuscation to blur the location information. The basic idea is to first locate the target in a small circle with the radius of the band distance; then, refine the estimated location with sufficient queries which is driven by the required localization accuracy. We theoretically prove that our method is able to converge to pinpoint users with an upper bound of the complexity of our design. We also evaluate the performance of our model when considering different distribution errors, and finally show our localization method is robust with exciting accuracy and limited complexity through extensive simulation experiments. This attack can locate target users within 20m with over 95% accuracy in most cases while the query-time is a limited value and can be roughly computed.

**Keywords:** Privacy leakage · Localization attack

Two-step localization · Location-based social network · WeChat

## 1 Introduction

Location-based services (LBS) provide value-added applications for users based on their locations. LBS can be used in a variety of contexts, such as health, indoor object search, entertainment, work, personal life, etc. Internet applications such as Facebook and Yelp allow users to “check-in” at restaurants, bars, retail outlets, schools, and offices, thereby sharing their locations within their social networks. However, the appearance of “the Snowden incident” has risen the people’s

consciousness of the privacy protection. Internet users are being reminded frequently that their online behaviors have been under constant scrutiny by NASA and other third parties.

All these factors have contributed to the recent growth of privacy-preserving mobile application people-nearby services. Applications with people-nearby services utilize the users' geographical information to provide proximity-based social and message discovery. People-nearby services are being used to find dating partners, friends who live or work nearby, and for bulletin-board style messages that have been posted nearby. These recent-growing people-nearby services may be based on anonymous messaging or relatively close relationship: 4chan, Whisper, and Yik Yak that allow users to anonymously post their thoughts to a public audience; and WeChat that allows users to share content only visible to friends.

The above cases of privacy-preserving communications have brought forth a challenging privacy problem: **Can a malicious third-party determine the locations of those nearby users by any people-nearby services in mobile application?** News articles have reported that the Egyptian government used trilateration to locate and imprison users of gay dating applications [1, 2]. Recent academic work has also shown several general avenues of localization attacks to evaluate privacy of the following two types of mobile application nearby services with *exact-distance* and *band-distance*. (i) If a mobile application people-nearby service provides the exact distance to the nearby user or message, the attack can be launched by taking readings from at least three different vantage points to trilaterate or triangulate the user's location [3, 4]. The readings from different vantage points can be set by virtual probes using fake GPS locations. (ii) If the mobile application people-nearby services do not provide an exact distance to the nearby user or message but instead report distances of nearby users in concentric bands, such as bands of 100 m as used by WeChat. Other research can still infer user's location with high accuracy from theory to practice [3–10].

In this work, we revisit mobile app people-nearby services with band distances and adopt a new model that is different from the existing methods to localize a user in a two-dimensional plane. In theory, our method can localize the target user within a square of any size. We also derive a complexity upper-bound of our algorithm. In practice, we acknowledge that localization errors may occur because of the GPS measurement deviation or that errors are intentionally added by mobile apps for privacy protection. At the same time we evaluate the performance of our model when considering different distribution errors, and finally show our localization method is robust through extensive simulation experiments.

The paper is organized as follows. Section 2 reviews some related work. Section 3 demonstrates the localization method. In Sect. 5, we evaluate our model based on simulation experiments considering different error distributions. In Sect. 6, we present some discussions. Finally, Sect. 7 concludes the paper.

## 2 Related Work

There have been quite a few studies on localizing users using mobile app people-nearby services either with exact distances or band distances. In Euclidean geometry, trilateration is the process of determining relative locations of points by measurement of exact distances, using the geometry of circles; triangulation is the process of determining the location of a point by measuring angles to it from known points at either end of a fixed baseline, rather than measuring distances to the point directly (trilateration). To perform the trilateration attack, assume that when a target user is known to lie on three circles from known locations, the centers of the three circles with their exact radii provide sufficient information to pinpoint the location of the target user [11]; the triangulation attack is similar. Qin *et al.* demonstrated triangulation attacks against services with exact distances [12]. In other independent studies, Mascetti *et al.* [13] applied a distance-based clustering algorithm to formalize a location privacy attack to approximately localize the users. Li *et al.* [3] explored user discovery attacks and highlight the significance of this threat. However, the method has some limitations that it may require applications' information while it is sensitive to the noise introduced by nearby services. Polakis *et al.* [4] conducted a theoretical study and proved tight bounds on the number of queries required for carrying out the localization attacks, irrespective of machine learning techniques.

Recently, many measurement studies focus on online social networks such as Whisper [5, 14], WeChat [6–10, 15, 16], and Yik Yak [17]. These online social networks have stored large volumes of sensitive data about users (e.g., controversial discussion information, user profiling, activity traces), all of which pose potential privacy risks.

## 3 A Two-Step Localization Model

Most locating models proposed in recent literature focus on the practical effects on the application people-nearby services while few of them have the theory analysis for the localization errors and the complexity of their models. In addition, the experiment results show that there are still room for improving the localization accuracy. In this section, we propose a new localization attack model in which localization errors can be quantified. Based on the theoretical analysis, we find an efficient way to locate the target user and quantify the query complexity of our model. Table 1 tabulates the notations that will be used throughout the rest of this paper.

### 3.1 A Two-Step Localization Algorithm

Since the location information provided is not specific enough to locate the target user easily with high precision, we divide the locating procedure into the following two steps: (i) Coarsely locate the target to a small circle whose radius is  $r$ , which needs to traverse the whole target distributing ring. This step is called

**Table 1.** Notations

Notation	Description
$dist(A, V)$	The Euclidean distance between point A and V
$r$	The band distance of the application
$V$	The target point
$A_i$	The attacker
$d_r$	The reporting distance of the application

*LocateToR*; (ii) Restrict the user within a small square precisely and efficiently when constraints to the horizontal and vertical coordinates are added. The step is called *LocateAccurate*. Algorithm 1 describes the entire locating procedure.

---

**Algorithm 1.** Two-step localization

---

**Input:**

$r$ : the band distance of the mobile app  
 $\epsilon$ : the tolerant error

**Output:**

$p_{est}$ : estimating target position  
 % the whole locating procedure  
 1:  $p = \text{LocateToR}(r)$   
 2:  $p_{est} = \text{LocateAccurate}(r, p, \epsilon)$   
 3: **return**  $p_{est}$

---

### 3.2 Coarse Location Estimation

*LocateToR* is aimed at coarsely restricting the target to a small circle whose radius is  $r$ . The basic idea is to cover the target distribution ring with the circle one by one until the reporting distance is  $r$  which indicates that the target is inside the circle, and then record the center of the circle for the following accurate localization.

The whole implementation is described in Algorithm 2. Suppose that the band distance of the application is  $r$  and the target is  $V$ , the reporting distance from  $V$  to initial attacker location  $A_0$  is  $d_r$ . To traverse the target circular band, we set all the covering circle centers distributing along the circle whose radius is  $d_r - \frac{r}{2}$ . We can easily take  $A_1$  that is  $d_r - \frac{r}{2}$  far above the  $A_0$  as shown in Fig. 1. The coordinates of  $A_i$  can be computed according to the angle  $\angle A_1 A_0 A_i$  and the coordinate of  $A_0$ . Besides, the  $\theta$  in the figure can be calculated by the cosine formula (take the advantage of triangle  $A_1 A_0 D$  in the illustration). As for *LocateToR*, we prove the following theorem.

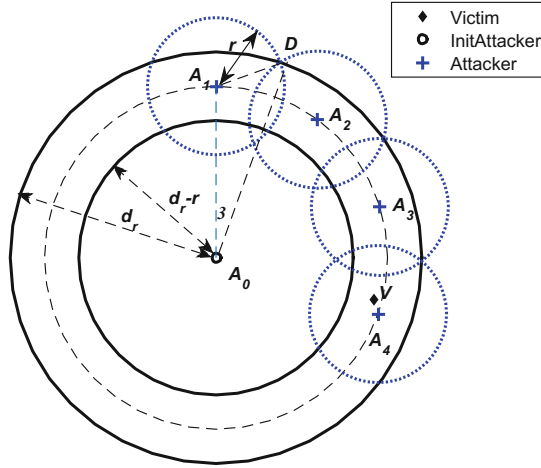
**Algorithm 2.** LocateToR**Input:**

$r$ : the band distance of the app  
 $p$ : the initial attacker position

**Output:**

$p$ : the coordinates of point which is less than  $r$  far from target  
 % restrict the target to a circle of radius  $r$   
 1:  $d_r = \text{AppDist}(p)$   
 2:  $p = \text{AttackerInit}(p)$   
 3: **while**  $d_r > r$  **do**  
 4:    $p = \text{NextAttacker}(p, r)$   
 5:    $d_r = \text{AppDist}(p)$   
 6: **end while**  
 7: **return**  $p$

**Theorem 1.** In a two-dimensional space, given a point  $V$  existing in a ring with internal diameter  $d_A - r$  and external diameter  $d_A$ , if we cover the ring with the disk whose radius is  $r$ . There must be a disk and the distance between its center and  $V$  is less than  $r$ , and it takes at most  $\Omega$  queries to find the disk, which satisfies  $\Omega = \frac{1}{2} \cdot \frac{2\pi}{2\theta} = \frac{\pi}{2\arccos\left(\frac{(d^2 + (d - \frac{r}{2})^2 - r^2)}{2d(d - \frac{r}{2})}\right)}$ .



**Fig. 1.** An illustration of the LocateToR algorithm

### 3.3 Location Refinement

In *LocateToR*, we have restricted the target point  $V$  in a small circle of radius  $r$ , now we turn our attention to locating  $V$  accurately. In Fig. 2,  $P \in \{W, E, N, S\}$



is  $r$  away from the target  $V$ . We reach our goal by finding these four points  $P$  that satisfy  $dist(\hat{P}, P) < \epsilon$  on the line which is horizontal or vertical.  $\epsilon$  is the tolerance set by us. It seems that  $V$  is pulled by  $P$  from four directions. Thus, the target  $V$  is restricted in a small square  $C_0C_1C_2C_3$  whose length of side is  $\epsilon$ . We take the center (i.e.,  $V_{est}$ ) of the square to be the estimated position of the target  $V$ . The algorithm details are shown in Algorithm 3.

Now we concentrate on how to determine  $\hat{W}, \hat{E}, \hat{N}, \hat{S}$  very fast. We first prove Theorem 2, Corollaries 1 and 2. *BiSecSearch* (i.e., Algorithm 4) is used to search the eligible points. Starting from the position which is returned by *LocateToR*, we hunt for  $\hat{W}, \hat{E}$  first. The moving distance is  $r$  at the beginning and reduces by half every time when moving forward to the actual point  $P$ . Once the reporting distance changes, moving direction will be changed. From Corollary 2, we make sure that  $dist(P, \hat{P})$  decreases exponentially. Then, taking the central point of  $\hat{W}, \hat{E}$  as the starting point, we obtain  $\hat{N}, \hat{S}$  in the similar way of determining  $\hat{W}, \hat{E}$ .

---

### Algorithm 3. LocateAccurate

---

**Input:**

- $r$ : the band distance of the mobile app
- $p$ : the point returned by *LocateToR*
- $\epsilon$ : the tolerant error

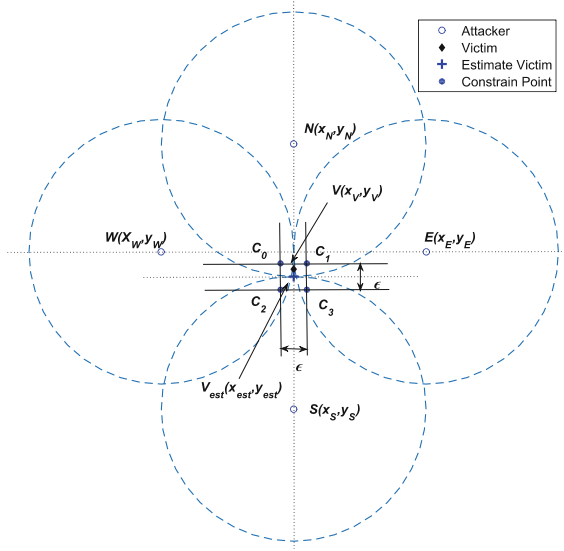
**Output:**

- $p_{est}$ : estimating target position
  - % locate the target accurately
  - 1:  $[p_{X1}, p_{X2}] = BiSecSearch(r, p, X, \epsilon)$
  - 2:  $p = \frac{p_{X1} + p_{X2}}{2}$
  - 3:  $[p_{Y1}, p_{Y2}] = BiSecSearch(r, p, Y, \epsilon)$
  - 4:  $p_{est,x} = \frac{p_{X1,x} + p_{X2,x}}{2}$
  - 5:  $p_{est,y} = \frac{p_{Y1,y} + p_{Y2,y}}{2}$
  - 6: **return**  $p_{est}$
- 

**Theorem 2.** For any two points  $A, V$  in a two-dimensional space, the distance between them satisfies  $dist(A, V) < r$ , and the two different points  $A_1, A_2$  can be found on any straight line crossing the  $A$  point, which satisfies the following equations:

$$\begin{aligned} dist(A_1, V) &= r, \\ dist(A_2, V) &= r. \end{aligned}$$

**Corollary 1.** For any two points  $A(x_A, y_A), V(x_V, y_V)$  in a two-dimensional space, the distance between two points satisfies  $dist(A, V) < r$ . Then two different points  $A_1$  and  $A_2$  can be found on the line  $x = x_A$  or  $y = y_A$  that is parallel to the axis through the point  $A$ , which satisfies:



**Fig. 2.** An illustration of the location refinement algorithm

$$\begin{aligned} \text{dist}(A_1, V) &= r, \\ \text{dist}(A_2, V) &= r, \\ \frac{x_{A_1} + x_{A_2}}{2} &= x_A \text{ or } \frac{y_{A_1} + y_{A_2}}{2} = y_A. \end{aligned}$$

**Corollary 2.** For any two points  $A, V$  in a two-dimensional space, the distance between two points satisfies  $\text{dist}(A, V) < r$ , there must exist  $A_1, A_2$  satisfies the Corollary 1. Starting from  $A$ , move the point along the any line that crosses the  $A$  point with the moving sequence  $l = \{l_i | l_i = \frac{r}{2^i}, i = 0, 1, \dots, N - 1\}$  and the direction is uncertain. For any small  $\epsilon$ , after  $N$  times of movement, we can find  $\hat{A}_1$  and  $\hat{A}_2$  respectively, which satisfies:

$$\begin{aligned} \text{dist}(A_1, \hat{A}_1) &< \epsilon, \\ \text{dist}(A_2, \hat{A}_2) &< \epsilon, \end{aligned}$$

where

$$N = \left\lceil \log_2 \frac{\sqrt{2}r}{\epsilon} \right\rceil + 1.$$

## 4 Theoretical Analysis

In Fig. 2, the target  $V$  is at  $(x_{victim}, y_{victim})$ , and we denote the estimation point with  $V_{est}(x_{est}, y_{est})$ . On the line  $y = y_{victim}$ ,  $W(x_W, y_W)$  and  $E(x_E, y_E)$

**Algorithm 4.** BiSecSearch**Input:**

- $r$ : the band distance of the mobile app
- $p$ : the initial point
- $dim$ : X or Y
- $\epsilon$ : the tolerant error when locate the point

**Output:**

- $P$ : List of the points of estimated position along the desired direction
- % look for the point who is  $r$  away from the target

```

1:  $IterNum = \lceil \log_2 \frac{\sqrt{2}r}{\epsilon} \rceil + 1$ 
2:  $P$  is a empty list
3: for  $sgn=-1,1$  do
4:    $p_{e,dim} = p_{dim} - sgn * r$ 
5:    $dim_2 = \{X,Y\} \setminus dim$ 
6:    $p_{e,dim_2} = p_{dim_2}$ 
7:   for  $j = 0, 1, \dots, IterNum$  do
8:      $d_r = AppDist(p_e)$ 
9:     if  $r > d_r$  then
10:        $p_{e,dim} = p_{e,dim} + sgn * r$ 
11:     else
12:        $p_{e,dim} = p_{e,dim} - sgn * r$ 
13:     end if
14:   end for
15:   insert  $p_e$  into  $P$ 
16: end for
17: return  $P$ 

```

are the points satisfying Corollary 1,  $\hat{W}(x_{\hat{W}}, y_{\hat{W}})$  and  $\hat{E}(x_{\hat{E}}, y_{\hat{E}})$  are the points satisfying Corollary 2,  $\epsilon$  is the tolerance parameter in *LocateAccurate*.

$$\sqrt{(x_W - x_{victim})^2 + (y_W - y_{victim})^2} = r, \quad (1)$$

$$\sqrt{(x_E - x_{victim})^2 + (y_E - y_{victim})^2} = r. \quad (2)$$

According to Corollary 2,

$$\sqrt{(x_W - x_{\hat{W}})^2 + (y_W - y_{\hat{W}})^2} < \epsilon.$$

Since  $y_W = y_E = y_I = y_{\hat{W}} = y_{\hat{E}}$ , then

$$|x_W - x_{\hat{W}}| < \epsilon, \quad (3a)$$

$$|x_E - x_{\hat{E}}| < \epsilon, \quad (3b)$$

in which

$$N = \left\lceil \log_2 \frac{\sqrt{2}r}{\epsilon} \right\rceil + 1.$$

Suppose  $\epsilon_W, \epsilon_E$  is the true error when estimating  $x_W, x_E$ , i.e.,

$$\epsilon_W = |x_W - \hat{x}_W| < \epsilon, \quad (4a)$$

$$\epsilon_E = |x_E - \hat{x}_E| < \epsilon. \quad (4b)$$

According to Corollary 1,

$$\left| \frac{x_W + x_E}{2} - x_{victim} \right| = 0. \quad (5)$$

Then, we have

$$\begin{aligned} |x_{est} - x_{victim}| &= \left| \frac{\hat{x}_W + \hat{x}_E}{2} - x_{victim} \right| \\ &= \left| \frac{(x_W - \epsilon_W) + (x_E - \epsilon_E)}{2} - x_{victim} \right| \\ &= \left| \frac{x_W + x_E}{2} - x_{victim} + \frac{\epsilon_W + \epsilon_E}{2} \right| \\ &= \left| \frac{\epsilon_W + \epsilon_E}{2} \right| < \frac{|\epsilon_W| + |\epsilon_E|}{2}. \end{aligned} \quad (6)$$

If  $\epsilon < \frac{r}{2^{N-1}}$ , we have

$$|x_{est} - x_{victim}| < \frac{r}{2^{N-1}}.$$

Similarly,

$$|y_{est} - y_{victim}| < \frac{r}{2^{N-1}}. \quad (7)$$

The final localization error between the target actual position and the estimated position is

$$\begin{aligned} err &= \sqrt{(x - x_{victim})^2 + (y - y_{victim})^2} \\ &< \frac{\sqrt{2}r}{2^{N-1}}. \end{aligned} \quad (8)$$

Thus, our method can achieve any small localization accuracy with more queries.

## 5 Experiments

In the sections before, we discussed that our algorithm can reach any precision in theory. However, there are always some differences between the actual distance and the measured distance. Nowadays the popular locating ways of mobile phone include GPS, Network Locating and the blending of the both. There will be inevitable errors regardless of locating ways. At the same time, the application would like to add errors to the location information to protect the privacy of the users. As a result, it is important for locating model to be robust when there exist errors that can not be negligible. Here we simulate our model with different error distribution settings and different band distance  $r$ .

## 5.1 Model Settings

The error model settings refer to [10]. Considering that errors exist even for short distance, we change the error to non-zero when the actual distance is less than 100 m for different models.

$$err = \begin{cases} \text{exprnd}(1), \text{Exponential}; \\ \text{unirnd}(0, 5), \text{Uniform}; \\ \text{raylrnd}(\frac{1}{1.253}), \text{Rayleigh}; \\ \text{normrnd}(1, \frac{1}{3}), \text{Gaussian}. \end{cases} \quad (9)$$

In our simulation, we set the error tolerance  $\epsilon = 1$  and suppose that the first measurement (i.e.,  $|A_0V|$  in Fig. 1) is accurate. The adding error is  $e\hat{r}r$  in the simulation which satisfies  $e\hat{r}r = n \cdot err, n \in Z^+$  (error is generated from the above model).

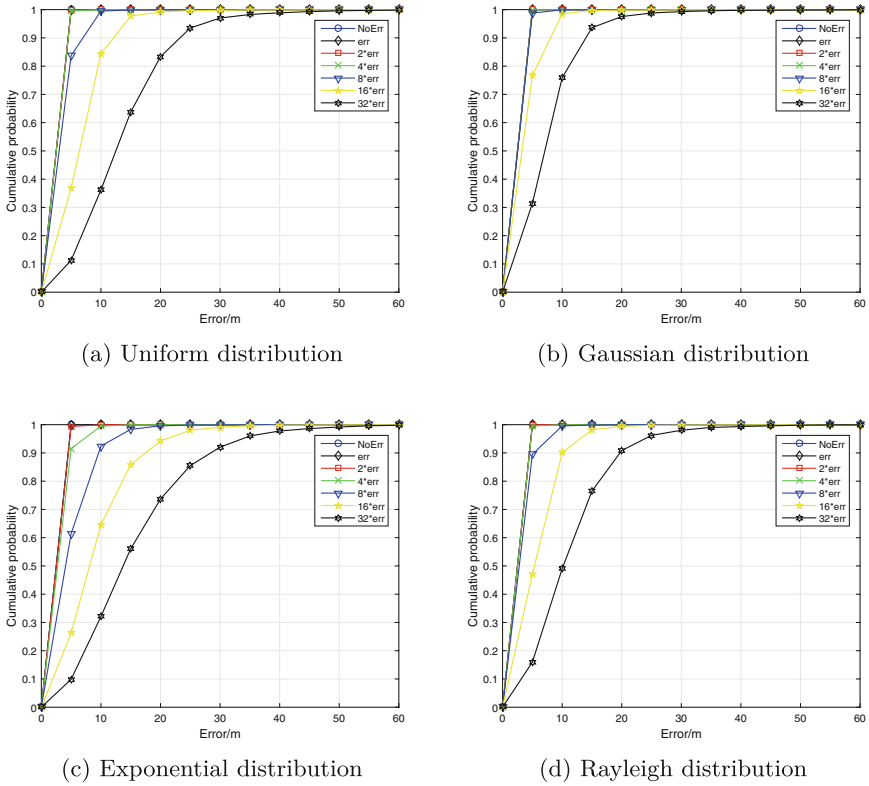
## 5.2 Simulation

**Effects of error distribution models.** Assume that  $r = 100$  m, our model only leverage the information of small distances. When the reporting distance is bigger than  $2r$ , it is useless. To exploit the model robustness well, we enlarge the original errors generated by above settings by timing  $n$ , and  $n \in \{1, 2, 4, 8, 16, 32\}$  (i.e.,  $e\hat{r}r = n \cdot err$ ).

Figure 3 shows the localization error distribution under different models with different mean. We could find that our model works well even if the error is very big (such as  $e\hat{r}r = 32err$ ). The worst case whose incorporating error distribution is exponential demonstrates that the localization error cumulative probability can reach 70% when the corresponding error is within 20 m. The remain cases increase approximately 10% higher, and the model works best for Gaussian distributions.

**Effects of band distance  $r$ .** We set  $e\hat{r}r = err, r \in \{100 \text{ m}, 200 \text{ m}, 400 \text{ m}, 800 \text{ m}\}$ . Figure 4 shows the result with different  $r$ . We find that the performance is worst when incorporating exponential distribution error. When  $r$  is small like less than 400 m, the localization error cumulative probability is close to 100% within 20 m for exponent distribution and close to 100% within 10 m for other error models; even  $r$  is bigger such as  $r = 800$  m, the cumulative probability is close to 80% within 40 m for exponential distribution and close to 100% within 40 m for other error models. Still the model performs best for gaussian distribution.

The above results show our localization methodology can efficiently get rid of the errors. Taking the locating details into consideration, we should know that the localization precision is decided by the refinement step *LocateToR* of our model. The following reasons may lead to our algorithm robustness: (i) In *LocateAccurate*, we locate the victim by adding constraints to the horizontal



**Fig. 3.** Localization error distributions under various error models with different average value settings

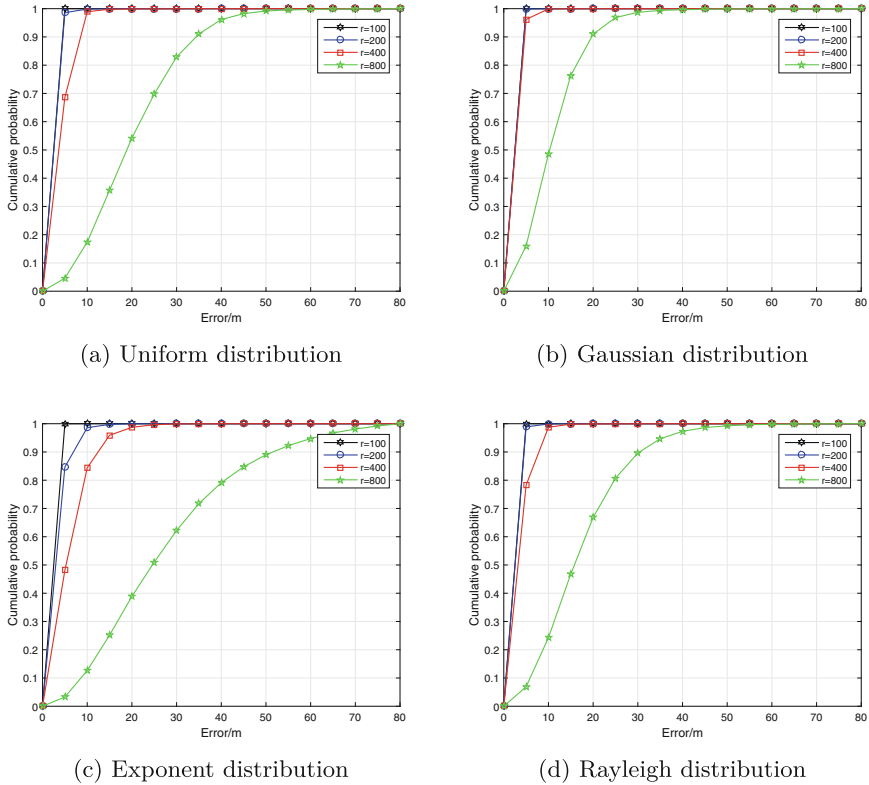
coordinates and vertical coordinates then estimate the position by taking average of the constraint points. Taking average may cancel the errors because all the errors are positive. (ii) When determining the constraint points, the movement step size is determined. And the resulting error must be caused by the previous wrong movement. As our algorithm is iterative, the previous error can be corrected by the following movement.

The model performs best when incorporating Gaussian error models. We conjecture that Gaussian distributions are symmetric distributions and the error may be cancelled by taking the average.

## 6 Discussions

### 6.1 Complexity Analysis

In this section, we analyze the query complexity of the proposed localization algorithm. It consists of the two parts: Complexity of *LocateToR* and Complexity of *LocateAccurate*. The first part is determined by the  $\theta$  in Fig. 1.



**Fig. 4.** Localization error distribution with different  $r$  for different error models

$$\begin{aligned}
 \Omega_1 &= \frac{1}{2} \cdot \frac{2\pi}{2\theta} \\
 &= \frac{\pi}{2 \arccos\left(\frac{(d^2 + (d - \frac{r}{2})^2 - r^2)}{2d(d - \frac{r}{2})}\right)}.
 \end{aligned} \tag{10}$$

The second part is decided by the tolerance  $\epsilon$ , the smaller  $\epsilon$  is, the more queries will be required.

$$\begin{aligned}
 \Omega_2 &= 4N \\
 &= 4 \cdot \left( \left\lceil \log_2 \frac{\sqrt{2}r}{\epsilon} \right\rceil + 1 \right).
 \end{aligned} \tag{11}$$

As a result, the query complexity of the entire model is

$$\begin{aligned}
 \Omega &= \Omega_1 + \Omega_2 \\
 &= \frac{\pi}{2 \arccos\left(\frac{(d^2 + (d - \frac{r}{2})^2 - r^2)}{2d(d - \frac{r}{2})}\right)} + 4 \cdot \left( \left\lceil \log_2 \frac{\sqrt{2}r}{\epsilon} \right\rceil + 1 \right).
 \end{aligned} \tag{12}$$

## 6.2 Comparison Remarks

The recent research progress of the privacy problems have driven popular LBSN service providers to enhance their apps to track more user abnormal behaviors for privacy protection. It has become more difficult to conduct experiments in real-world systems. Nevertheless, we believe that the battle between location attacks and protection will be still continuing. In this subsection, we compare this paper with a few representative studies. [7,9,10] focus more on using a number of probes in order to decrease the obfuscation. These methods require significant probe cost while the localization results are less accurate than this work. [4] proposed a similar spatial iteration attack approach yet with 3 – 5 s response time. Our method is less time-consuming. [17] requires to train a supervised or unsupervised model. The results are promising yet the data-collection procedure is expensive. In summary, our two-step method provides a practical attack method with high accuracy and small time complexity.

## 7 Conclusion

Mobile app people-nearby services have often been providing band distances instead of exact distances in order to protect privacy. In this paper, we revisit localization attacks in band-based distances. We proposed a new model to launch localization attacks and proved that our method can pinpoint target users accurately with the theoretically settings. When we incorporated different error distributions into our model, the simulation results showed that our model can efficiently combat against the impact the errors. This proposed model is robust with all most all band distances. In this paper, we continuously investigate the privacy leakage problem from end-systems. In the emerging software defined wireless networks, the network infrastructure may provide more data functions [18]. We envision that the privacy leakage problem may become more severe in the coming software defined edge computing and networking era.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China (No. 61370231), and in part by the Fundamental Research Funds for the Central Universities (No. HUST:2016YXMS303).

## References

1. Noack, R.: Could using gay dating app Grindr get you arrested in Egypt? The Washington Post, 12 September 2014
2. Paton, C.: Grindr urges LGBT community to hide their identities as Egypt persecutes nation’s gay community. The Independent, 26 September 2014
3. Li, M., Zhu, H., Gao, Z., Chen, S., Yu, L., Hu, S., Ren, K.: All your location are belong to us: breaking mobile social networks for automated user location tracking. In: 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing, pp. 43–52 (2014)



4. Polakis, I., Argyros, G., Petsios, T., Sivakorn, S., Keromytis, A.D.: Where's wally?: Precise user discovery attacks in location proximity services. In: ACM SIGSAC CCS, pp. 817–828 (2015)
5. Wang, G., Wang, B., Wang, T., Nika, A., Zheng, H., Zhao, B.Y.: Whispers in the dark: analysis of an anonymous social network. In: ACM Internet Measurement Conference, pp. 137–150 (2014)
6. Ding, Y., Peddinti, S.T., Ross, K.W.: Stalking Beijing from Timbuktu: a generic measurement approach for exploiting location-based social discovery. In: ACM Workshop on Security and Privacy in Smartphones & Mobile Devices (2014)
7. Xue, M., Liu, Y., Ross, K.W., Qian, H.: I know where you are: thwarting privacy protection in location-based social discovery services. In: IEEE Conference on Computer Communications Workshops (2015)
8. Xue, M., Liu, Y., Ross, K., Qian, H.: Thwarting location privacy protection in location-based social discovery services. *Secur. Commun. Netw.* **9**(11), 1496–1508 (2016)
9. Peng, J., Meng, Y., Xue, M., Hei, X., Ross, K.W.: Attacks and defenses in location-based social networks: a heuristic number theory approach. In: International Symposium on Security and Privacy in Social Networks and Big Data (SocialSec), pp. 64–71 (2015)
10. Cheng, H., Mao, S., Xue, M., Hei, X.: On the impact of location errors on localization attacks in location-based social network services. In: Wang, G., Ray, I., Alcaraz Calero, J.M., Thampi, S.M. (eds.) SpaCCS 2016. LNCS, vol. 10066, pp. 343–357. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-49148-6\\_29](https://doi.org/10.1007/978-3-319-49148-6_29)
11. Liu, J., Zhang, Y., Zhao, F.: Robust distributed node localization with error management. In: Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing, pp. 250–261 (2006)
12. Qin, G., Patsakis, C., Bourouche, M.: Playing hide and seek with mobile dating applications. In: Cuppens-Boulahia, N., Cuppens, F., Jajodia, S., Abou El Kalam, A., Sans, T. (eds.) SEC 2014. IAICT, vol. 428, pp. 185–196. Springer, Heidelberg (2014). [https://doi.org/10.1007/978-3-642-55415-5\\_15](https://doi.org/10.1007/978-3-642-55415-5_15)
13. Mascetti, S., Bertolaja, L., Bettini, C.: A practical location privacy attack in proximity services. In: IEEE 14th International Conference on Mobile Data Management (MDM), vol. 1, pp. 87–96 (2013)
14. Correa, D., Silva, L.A., Mondal, M., Benevenuto, F., Gummadi, K.P.: The many shades of anonymity: characterizing anonymous social media content. In: International AAAI Conference on Web and Social Media (2015)
15. Xue, M., Yang, L., Ross, K.W., Qian, H.: Characterizing user behaviors in location-based find-and-flirt services: anonymity and demographics. *Peer-to-Peer Netw. Appl.* **10**(2), 357–367 (2017)
16. Wang, R., Xue, M., Liu, K., Qian, H.: Data-driven privacy analytics: a WeChat case study in location-based social networks. In: Xu, K., Zhu, H. (eds.) WASA 2015. LNCS, vol. 9204, pp. 561–570. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-21837-3\\_55](https://doi.org/10.1007/978-3-319-21837-3_55)
17. Xue, M., Ballard, C., Liu, K., Nemelka, C., Wu, Y., Ross, K., Qian, H.: You can yak but you can't hide: localizing anonymous social network users. In: ACM IMC, pp. 25–31 (2016)
18. Chen, Z., Fu, D., Gao, Y., Hei, X.: Performance evaluation for software defined WiFi DCF networks from theory to testbed. In: 16th IEEE International Conference on Ubiquitous Computing and Communications (IUCC) (2017)

# Lengthening Unidimensional Continuous-Variable Quantum Key Distribution with Noiseless Linear Amplifier

Yu Cao, Jianwu Liang, and Ying Guo<sup>(✉)</sup>

School of Information Science and Engineering,  
Central South University, Changsha 410083, China  
yingguo@csu.edu.cn

**Abstract.** In order to increase the maximum transmission distance and simplified implementation, we propose a continuous-variable quantum key distribution (CVQKD) protocol of entanglement in the middle, which the protocol is based on gaussian modulation of a single quadrature of the coherent states of light and the noiseless linear amplifier (NLA). This protocol uses the simplified unidimensional protocol to simplified implementation, and uses the NLA to increase the maximum transmission distance and tolerable excess noise in the presence of gaussian lossy and noisy channel. The simulation results show that the proposed unidimensional CVQKD protocol of entanglement in the middle with NLA obvious increased the maximum transmission distance.

**Keywords:** Unidimensional · CVQKD · Source in the middle  
Noiseless linear amplifier

## 1 Introduction

A principal practical application of quantum-information science is the quantum key distribution (QKD), which allows two remote parties, normally referred to as Alice and Bob, to generate a secure key through an insecure quantum channel [1–6]. Quantum key distribution (QKD), an important application of the quantum information sciences, enables two correspondents to share a secret key, which in turn allows them to communicate with full security [7]. Its unconditional security is guaranteed by the Heisenberg’s uncertainty principle and the quantum no-cloning theorem. In the first QKD protocol, the Bennett and Brassard 1984 (BB84) QKD protocol [6], information is carried by single-photon signals. Due to the channel loss and other implementation imperfections, those single-photon signals are detected in a probabilistic fashion with a relatively high quantum bit error rate (QBER) on the order of  $10^{-2}$  [8]. This is in sharp contrast to a classical optical communication system, where information can be transmitted deterministically in an almost error-free fashion. Furthermore, specialized devices, such as single-photon detectors, are typically required in QKD. This makes QKD a very different communication modality in comparison with classical communication.

Most QKD protocols encode information on discrete variables such as the phase or the polarization of single photons and are currently facing technological challenges, especially the limited performances of photodetectors in terms of speed and efficiency in the single-photon regime. A way to relieve this constraint is to encode information on continuous variables such as the quadratures of coherent states [9] which are easily generated and measured with remarkable precision by standard optical telecommunication components. Compared with the discrete variable QKD, it has been shown theoretically that the continuous variable (CV) QKD has potential advantages, e.g., higher secret key rate and better compatibility with the current optical networks [10]. Since it was proposed, the CVQKD has achieved great progress both in theory and experiment. The most valuable achievement is the Gaussian-modulated coherent state (GMCS) CVQKD protocol, which has been proven to be secure against general collective attacks [11–14] and coherent attacks [15, 17]. Recently, long secure distance [18, 19] and high speed [20, 21] GMCS CVQKD have been experimentally demonstrated by using commercial components.

Continuous-variable key distribution (CVQKD) has been intensively developed these years [22]. The CVQKD approach has two main advantages: first, it avoids the limitations associated with single photon counting, and second, it offers the prospect of very high rate secure key distribution. There are several kinds of the Gaussian CVQKD protocols, which are based on the sender Alice's states (squeezed or coherent states), the receiver Bob's measurements (homodyne or heterodyne detection), and reconciliation approaches (direct or reverse reconciliation) [23, 24]. In a practical implementation, it does not require the singlephoton detector any longer and hence can be made to be much compatible with the modern optical telecommunication networks. Although it theoretically demonstrates that a secure key can be generated for a pure loss channel in long distance, the practical CVQKD is still limited to scores of kilometers by noise or loss in unreliable source monitoring and imperfect data processing [18].

The usual way to analyze CVQKD protocols is based on the typical assumption that one of the legitimate partners created the source of the QKD protocol, e.g., Alice created the source of the QKD protocol, Eve being the source of the entanglement has not been considered in CVQKD [25, 26]. In a recent paper [27], Christian proposed a protocol where the entangled source originates not from one of the trusted parties, Alice and Bob, but from the malicious eavesdropper in the middle.

Here we are interested in the problem of extending CVQKD with entanglement in the middle over longer distances but with proven security, and simplified implementation compared to the symmetrically modulated gaussian coherent-state protocols. The main idea is as follows: First of all, we apply an improved CVQKD protocol where Eve is placed in the middle between Alice and Bob and is given power to control the creation of the gaussian entangled source. Secondly, we propose protocol by using NLA before heterodyne detection of Bob. Thirdly, we propose protocol based on the gaussian modulation of a single quadrature of the coherent states of light aimed to simplified implementation.

The remaining of the paper is organized as follows. In Sect. 2, application of NLA in CVQKD is described particularly. In Sect. 3, a modified CVQKD protocol with entanglement in the middle based on the gaussian modulation of a single quadrature of the coherent states of light is proposed to improve the system by adding a NLA before Bob's measurement. In Sect. 4, we derive the secret key rates of the protocol in this paper. In Sect. 5, we present the performance of this study. Then, the conclusion is drawn in Sect. 6.

## 2 Application of NLA in CVQKD

The quantum state preparation is an important link of CVQKD protocol. In the GG02 protocol widely used currently, Alice uses coherent state as the optical source and translates the central position of the coherent state through the modulator to implement the Gaussian modulation. However, during the quantum state preparation, due to diversified non-ideal characteristics existing in the laser and the modulator, the prepared quantum state will also include some noise. For instance, in the actual system, the laser does not emerge non-ideal coherent state, but includes the amplitude noise and phase noise; the modulator is limited by its own performance and the voltage signal as excitation, and it cannot implement the modulation result with perfect Gaussian distribution, and those factors will result in the noise in the preparation of quantum state, and those noises are collectively called optical source noises.

The optical source noise will have great negative influence on the system performance. As for the earlier research on the optical source noises, it is considered that the optical source is generated within the information sender and thus it is not controlled by the eavesdropper, and therefore it reduces the mutual information between two communication parties, and it will not result in the additional leakage of the key information. For this characteristic, researchers try to theoretically reduce the negative influence of the optical source noises on the system performance through the reasonable modeling of the optical source noises.

Ralph et al. first proposed the concept of NLA [28] in 2009. In recent years, researches related to NLA have had abundant achievements theoretically and experimentally. In 2013, Walk et al. introduced the LNA to the phase space representation, gave the transformation form of the noiseless amplification operation in the phase space to greatly simplify the mathematical analysis difficulty in the NLA [29]. McMahan et al. in this group proposed the optical architecture of the NLA which is based on the unitary transformation, the projection measurement and such operation and explored the upper limit to magnify the probability of success [30]; in the same year, Bernu et al. in this group provided the theoretical analysis of entanglement purification for the EPR state by utilizing NLA and the quantum scissor structure to make a conclusion that the entanglement purification and the probability of success are required to be compromised [31].

Under CVQKD application scenarios, the effect of the NLA is reflected in the improvement in the correlation between Alice and Bob through the entanglement purification. Before the quantum state distribution, Alice and Bob share

partial correlation through the quantum entanglement; upon the quantum state transmitted through the quantum channel, the eavesdropping behavior by the Eve makes her respectively establish the correlation between Alice and Bob and at the same time weaken the correlation between Alice and Bob; before detection, the NLA is used to simultaneously improve the correlation between two parties among Alice, Bob and Eve. When the correlation between Bob and Alice is stronger than the correlation between Bob and Eve, the objective of NLA is used to improve CVQKD performance is achieved.

Quantitatively, the role of NLA is to update the parameters of the original protocol. In the EB protocol, Alice performs a heterodyne detection of one mode of the EPR state and sends the other mode  $B_0$  to Bob. When Bob does not know the measurement results of Alice, the quantum state of mode  $B_0$  corresponds to a thermal field, that is:

$$\rho_{B_0} = (1 - \lambda^2) \sum_{n=0}^{\infty} \lambda^{2n} |n\rangle\langle n|. \quad (1)$$

Accordingly, the mean value of the orthogonal component of the thermal field is 0, and the variance is  $V = (1 + \lambda^2)/(1 - \lambda^2)$ . After the transmission of the quantum channel with the parameter  $(T, \varepsilon)$ , the mode  $B_0$  becomes the mode  $B_1$ , and the variance of the orthogonal component also changes to  $T(V + \varepsilon) + 1 - T$ . Thus, we can write the expression of the thermal field of mode  $B_1$ :

$$\rho_{B_1} = (1 - \lambda_1^2) \sum_{n=0}^{\infty} \lambda_1^{2n} |n\rangle\langle n|. \quad (2)$$

Using the mode  $B_1$ , the variance of orthogonal components can be calculated:

$$\lambda_1^2 = \frac{T[\lambda^2(2 - \varepsilon) + \varepsilon]}{2 - \lambda^2[2 + T(\varepsilon - 2)] + T\varepsilon}. \quad (3)$$

Since the thermal field  $\rho_{B_1}$  can also be written in the form of follow function:

$$\rho_{B_1}(\lambda_1) = \int \frac{1 - \lambda_1^2}{\pi \lambda_1^2} e^{\frac{1 - \lambda_1^2}{\lambda_1^2} |\alpha|^2} |\alpha\rangle\langle \alpha| d\alpha. \quad (4)$$

At the receiving end, Bob uses NLA to amplify the received thermal field state  $\rho_{B_1}$ . After amplification, quantum state of the mode B of the NLA output can be written:

$$\rho_B^g(\lambda_1) = g^{\hat{n}} \rho_{B_1}(\lambda_1) g^{\hat{n}}. \quad (5)$$

Combination of formula (4) and (5) can be obtained:

$$\rho_B^g(\lambda_1) = C(1 - g^2 \lambda_1^2) \sum_{n=0}^{\infty} (g \lambda_1)^{2n} |n\rangle\langle n|, \quad (6)$$

where C is the normalized coefficient in the upper form, independent of the integral variables. It can be seen from the above that the quantum state of NLA

amplification is a thermal field with a parameter  $g\lambda_1$ . Conclusion combined with the formula: when Bob does not know the measurement results  $\alpha_A$  of Alice, NLA be going to thermal field state of the Bob device with the incident side parameter  $\lambda_1$  amplification to the thermal states of Bob device with output parameter  $g\lambda_1$ , the formula can be expressed as:

$$\frac{T[\lambda^2(2-\varepsilon)+\varepsilon]}{2-\lambda^2[2+T(\varepsilon-2)]+T\varepsilon} \rightarrow g^2 \frac{T[\lambda^2(2-\varepsilon)+\varepsilon]}{2-\lambda^2[2+T(\varepsilon-2)]+T\varepsilon}. \quad (7)$$

When Bob has known measured result of Alice, the quantum state of mode  $B0$  corresponds to a coherent state. After through the quantum channel transmission with the parameter  $(T, \varepsilon)$ , the coherent state of mode  $B0$  is transformed into the translational thermal field of mode  $B1$ , which can be expressed by the translation operator  $D(\beta)$  on the thermal field  $\rho_{th}(\lambda_2)$  with the coefficient of  $\lambda_2$ :

$$\rho_{B1|A} = D(\beta)\rho_{th}(\lambda_2)D(-\beta) = \int \frac{1-\lambda_2^2}{\pi\lambda_2^2} e^{\frac{1-\lambda_2^2}{\lambda_2^2}|\alpha-\beta|^2} |\alpha\rangle\langle\alpha| d\alpha, \quad (8)$$

where,  $\beta = \sqrt{T}\lambda\alpha_A$ ,  $\lambda_2^2 = T\varepsilon/(T\varepsilon+2)$ . Using the NLA to magnify the translation thermal field  $\rho_{B1|A}$  can get:

$$\rho_{B1|A}^g = \int \frac{1-\lambda_2^2}{\pi\lambda_2^2} e^{\frac{1-\lambda_2^2}{\lambda_2^2}|\alpha-\beta|^2} g^{\hat{n}} |\alpha\rangle\langle\alpha| g^{\hat{n}} d\alpha. \quad (9)$$

That is

$$\rho_{B1|A}^g = C \int \frac{1-g^2\lambda_2^2}{\pi g^2\lambda_2^2} e^{\frac{1-g^2\lambda_2^2}{g^2\lambda_2^2}|u-g\frac{1-\lambda_2^2}{1-g^2\lambda_2^2}\beta|^2} |u\rangle\langle u| du, \quad (10)$$

$C$  is the normalized coefficient in the upper form, independent of the integral variables. It can be seen from the above that the translational thermal field is still the translational thermal field after NLA amplification, but the mean value and the variance are changed:

$$\sqrt{T}\lambda\alpha_A \rightarrow g\frac{1-\lambda_2^2}{1-g^2\lambda_2^2}\sqrt{T}\lambda\alpha_A, \lambda_2^2 \rightarrow g^2\lambda_2^2. \quad (11)$$

EPR states  $|\Phi_{AB_0}(\lambda)\rangle$  through transmission the quantum channel with parameters  $(T, \varepsilon)$  then be successfully amplified by NLA, the covariance matrix can be equivalent to another EPR state  $|\Phi_{AB_0}(\zeta)\rangle$  which through transmission quantum channel with parameters  $(\eta, \varepsilon^g)$  but without amplification by NLA [32]. For the above equation can be

$$\zeta = \lambda\sqrt{\frac{(g^2-1)(\varepsilon-2)T-2}{(g^2-1)\varepsilon T-2}}. \quad (12)$$

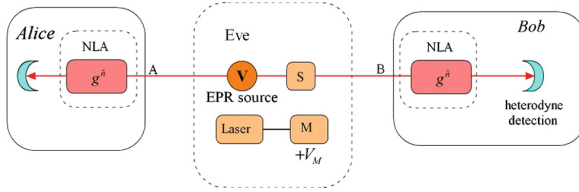
$$\eta = \frac{g^2T}{(g^2-1)T[(g^2-1)(\varepsilon-2)\varepsilon T/4-\varepsilon+1]+1}. \quad (13)$$

$$\varepsilon^g = \varepsilon - \frac{1}{2}(g^2-1)(\varepsilon-2)\varepsilon T. \quad (14)$$

The improvement of the CVQKD protocol utilization NLA can be calculated.

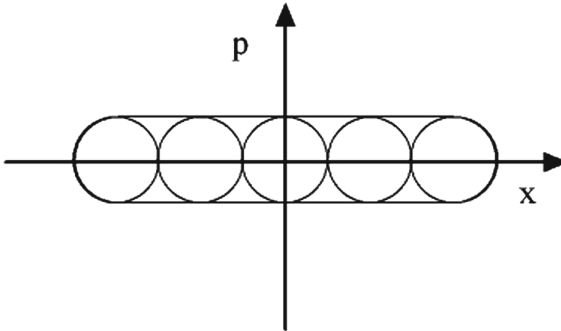
### 3 Modified CVQKD Protocol

Inspired by the method in [27,33], we propose a modified CVQKD protocol based on the gaussian modulation of a single quadrature of the coherent states of light by using a NLA and entanglement source in the middle. The scheme of the protocol is given in Fig. 1. Eve produces coherent states with a laser source. Then she applies modulation in one of the quadratures (denoted as  $x$ ), using modulator  $M$ , and displaces each coherent state according to a random Gaussian variable with displacement variance  $V_M$ . In this protocol, Eve prepares a coherent state using a laser source and then modulates the state by displacing it along the modulated quadrature using modulator  $M$  so that the modulation variance is  $V_M$ . The states travel through an untrusted generally phase-sensitive channel to a remote party Bob, who performs heterodyne measurement of the modulated quadrature, and we insert a NLA at the output of the quantum channel and inside Bob's apparatus.



**Fig. 1.** The scheme of unidimensional CVQKD protocol with Entanglement in the Middle by Using NLA.

The modulation scheme is given in Fig. 2. Mixture of modulated coherent states on a phase space (assuming  $x$  quadrature was modulated).



**Fig. 2.** Modulation scheme of unidimensional.

## 4 Secret Key Rates

In order to simplify the security analysis, we focus on the direct reconciliation, and the protocols based on Gaussian modulation of coherent states and heterodyne detection. The secret key rate for the modified protocol is defined as:

$$K = S_{AB}(g) - S_{AE}(g), \quad (15)$$

where  $S_{AB}$  and  $S_{AE}$  are the mutual information between Alice and Bob and Alice and Eve, respectively.

Reference the equivalent entanglement-based (EPR) scheme [34], which allows the explicit description of trusted modes and their correlations. we taking a two-mode squeezed vacuum state of variance  $V$  and squeezing one of its modes with the squeezing parameter  $-\log\sqrt{V}$ , resulting in the covariance matrix:

$$\gamma_{AB} = \begin{pmatrix} V & 0 & \sqrt{V(V^2-1)} & 0 \\ 0 & V & 0 & -\sqrt{\frac{V^2-1}{V}} \\ \sqrt{V(V^2-1)} & 0 & V^2 & 0 \\ 0 & -\sqrt{\frac{V^2-1}{V}} & 0 & 1 \end{pmatrix}. \quad (16)$$

Since the EPR scheme is then equivalent to the Gaussian displacement of coherent states along the  $x$  quadrature with the variance  $V_M = V^2 - 1$ , so covariance matrix after the modulation variance  $V_M$  is shown below,

$$\gamma'_{AB} = \begin{pmatrix} \sqrt{1+V_M} & 0 & \sqrt{\eta_x V_M}(1+V_M)^{\frac{1}{4}} & 0 \\ 0 & \sqrt{1+V_M} & 0 & C_p \\ \sqrt{\eta_x V_M}(1+V_M)^{\frac{1}{4}} & 0 & 1 + \eta_x(V_M + \varepsilon_x) & 0 \\ 0 & C_p & 0 & V_p^B \end{pmatrix}, \quad (17)$$

where  $\eta_x$  is the channel transmittance and  $\varepsilon_x$  is the excess noise of estimated by the trusted parties through the measurement of the  $x$  quadrature.  $C_p$  is the correlation between trusted modes in the  $p$  quadrature.  $V_p^B$  is the output variance of the mode  $B$  in the  $p$  quadrature. The mutual information between Alice and Bob can be written as:

$$S_{AB}(g) = \frac{1}{2} \log\left(1 + \frac{\eta_x V_M}{1 + \eta_x \varepsilon_x}\right). \quad (18)$$

As for the mutual information between Eve and Alice, it can be written as:

$$S_{AE}(g) = S(E) - S(E|x_A). \quad (19)$$

Since Eve provides a purification of Alice and Bob's density matrix, we can write  $S(E) = S(AB)$ . So we can get:

$$S(E) = G\left(\frac{\lambda_1 - 1}{2}\right) + G\left(\frac{\lambda_2 - 1}{2}\right), \quad (20)$$

where  $G(x) = (x+1)\log(x+1) - x\log(x)$  is the bosonic entropic function [35].  $S(AB)$  can be calculated from the symplectic eigenvalues  $\lambda_{1,2}$  that are given by the square roots of the solutions of equation,



$$z^2 - \Delta z + \det\gamma'_{AB} = 0, \quad (21)$$

where  $\Delta = \det\gamma_A + \det\gamma_B + 2\sigma_{AB}$ .  $\gamma_A$  is the covariance matrix of the modes A, and  $\gamma_B$  is the covariance matrix of the modes B, both are the submatrices of the covariance matrix  $\gamma'_{AB}$ .  $\sigma_{AB}$  is the submatrix of  $\gamma'_{AB}$ , which denotes correlation between modes A and B. We can write  $S(E|A) = S(B|A)$  using the same purification argument. So we calculate Bobs correlation matrix conditioned on Alices measurement outcome  $x_a$ , which is calculated using,

$$\gamma_B^{x_a} = \gamma_B - \sigma_{AB}(X\gamma_A X)^{-1}\sigma_{AB}, \quad (22)$$

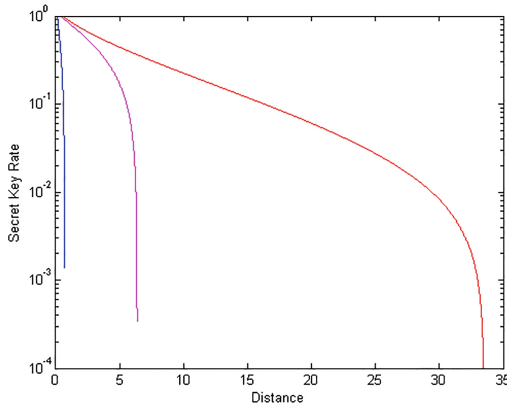
where the inverse is a pseudoinverse and the matrix  $X$  is given by,

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (23)$$

Since  $S(B|A) = G[\frac{\lambda_3-1}{2}]$ , we can calculate  $S(B|A)$  from  $\lambda_3 = \sqrt{\det\gamma_B^{x_a}}$ . Finally, the Shannon secret key generation rate can be calculated by the above formula.

## 5 Numerical Simulation and Discussion

In this section, we apply the results derived in Sect. 4 to practical QKD systems in order to compare their performance for different protocol. We compare the modified CVQKD protocol with the standard symmetrical modulation protocol used over the same channel. We compare the gaussian coherent-state CVQKD protocol, Unidimensional CVQKD protocol [33], and our proposed unidimensional CVQKD protocol of entanglement in the middle with NLA.



**Fig. 3.** Secret key generation rate as a function of distance for three protocols. (Blue solid line) Gaussian coherent-state CVQKD protocol; (Fuchsia solid line) Unidimensional CVQKD protocol; (Red solid line) Unidimensional CVQKD protocol of entanglement in the middle with NLA. Modulation variance  $V_M = 100$ . (Color figure online)

Normally, various protocols are present the channel noise and reduces the security of the protocol. The results of the calculations in this case are given in Fig. 3 in terms of the key rate upon fixed channel excess noise  $\varepsilon = 0.05$ , modulation variance  $V_M = 100$ , and the simulation parameters are the attenuation coefficient  $\eta = 0.95$  of the light source. As can be seen from the figure after the use of NLA, Bob side using heterodyne detection, the safety distance has been improved; the simulation results show that the safety of the NLA system has brought a significant increase in the distance.

## 6 Conclusion

We have proposed a modified unidimensional CVQKD protocol of entanglement in the middle with NLA based on gaussian modulation of a single quadrature of the coherent states of light. In the protocol, Eve controls the source of the protocol originates from the malicious eavesdropper who is in the middle between Alice and Bob, and allows simpler technical realization, and by utilizing the NLA to obviously increase the farthest transmission distance. The simulation results show that our proposed unidimensional CVQKD protocol of entanglement in the middle with NLA obvious increased the maximum transmission distance compared to the gaussian coherent-state CVQKD protocol and the unidimensional CVQKD protocol.

**Acknowledgment.** This work was supported by the National Natural Science Foundation of China (Grant Nos. 61379153, 61572529).

## References

1. Ekert, A.K.: Quantum cryptography theorem. *Phys. Rev. Lett.* **67**, 661 (1991)
2. Gisin, N., Ribordy, G., Tittel, W., Zbinden, H.: Quantum cryptography. *Rev. Mod. Phys.* **74**, 145 (2002)
3. Scarani, V., Bechmann-Pasquinucci, H., Cerf, N.J.: The security of practical quantum key distribution. *Rev. Mod. Phys.* **81**, 1301 (2009)
4. Lo, H.-K., Curty, M., Tamaki, K.: Secure quantum key distribution. *Nat. Photonics* **8**, 595–604 (2014)
5. Diamanti, E., Lo, H.-K., Qi, B., Yuan, Z.: Practical challenges in quantum key distribution. arXiv preprint [arXiv:1606.05853](https://arxiv.org/abs/1606.05853) (2016)
6. Bennett, C.H., Brassard, G.: An update on quantum cryptography. In: Blakley, G.R., Chaum, D. (eds.) *CRYPTO 1984*. LNCS, vol. 196, pp. 475–480. Springer, Heidelberg (1985). [https://doi.org/10.1007/3-540-39568-7\\_39](https://doi.org/10.1007/3-540-39568-7_39)
7. Garca-Patrñ, R., Cerf, N.J.: Continuous-variable quantum key distribution protocols over noisy channels. *Phys. Rev. Lett.* **102**, 130501 (2009)
8. Zhao, Y., Qi, B., Ma, X., Lo, H.-K., Qian, L.: Experimental quantum key distribution with decoy states. *Phys. Rev. Lett.* **96**, 070502 (2006)
9. Grosshans, F., Van Assche, G., Wenger, J., Brouri, R., Cerf, N.J., Grangier, P.: Quantum key distribution using Gaussian-modulated coherent states. arXiv preprint [quant-ph/0312016](https://arxiv.org/abs/quant-ph/0312016) (2003)

10. Grosshans, F., Grangier, P.: Continuous variable quantum cryptography using coherent states. *Phys. Rev. Lett.* **88**, 057902 (2002)
11. Grosshans, F.: Optimality of Gaussian attacks in continuous-variable quantum cryptography. *Phys. Rev. Lett.* **97**, 190502 (2006)
12. Cerf, N.J.: Unconditional optimality of Gaussian attacks against continuous-variable quantum key distribution. *Phys. Rev. Lett.* **97**, 190503 (2006)
13. Leverrier, A., Grosshans, F., Grangier, P.: Finite-size analysis of a continuous-variable quantum key distribution. *Phys. Rev. A* **81**, 062343 (2010)
14. Leverrier, A.: Composable security proof for continuous-variable quantum key distribution with coherent states. *Phys. Rev. Lett.* **114**, 070501 (2015)
15. Renner, R., Cirac, J.I.: de Finetti representation theorem for infinite-dimensional quantum systems and applications to quantum cryptography. *Phys. Rev. Lett.* **102**, 110504 (2009)
16. Furrer, F., Franz, T., Berta, M., Leverrier, A., Scholz, V.B., Tomamichel, M., et al.: Continuous variable quantum key distribution: finite-key analysis of composable security against coherent attacks. *Phys. Rev. Lett.* **109**, 100502 (2012)
17. Leverrier, A., Renner, R., Cerf, N.J.: Security of continuous-variable quantum key distribution against general attacks. *Phys. Rev. Lett.* **110**, 030502 (2013)
18. Jouguet, P., Kunz-Jacques, S., Leverrier, A., Grangier, P., Diamanti, E.: Experimental demonstration of long-distance continuous-variable quantum key distribution. *Nat. Photonics* **7**, 378–381 (2013)
19. Huang, D., Huang, P., Lin, D., Zeng, G.: Long-distance continuous-variable quantum key distribution by controlling excess noise. *Sci. Rep.* **6** (2016)
20. Huang, D., Lin, D., Wang, C., Liu, W., Fang, S., Peng, J., et al.: Continuous-variable quantum key distribution with 1 Mbps secure key rate. *Opt. Express* **23**, 17511–17519 (2015)
21. Wang, C., Huang, D., Huang, P., Lin, D., Peng, J., Zeng, G.: 25 MHz clock continuous-variable quantum key distribution system over 50 km fiber channel. *Sci. Rep.* **5** (2015)
22. Weedbrook, C., Pirandola, S., Garca-Patrón, R., Cerf, N.J., Ralph, T.C., Shapiro, J.H., et al.: Gaussian quantum information. *Rev. Mod. Phys.* **84**, 621 (2012)
23. Gerhardt, I., Liu, Q., Lamas-Linares, A., Skaar, J., Kurtsiefer, C., Makarov, V.: Full-field implementation of a perfect eavesdropper on a quantum cryptography system. arXiv preprint [arXiv:1011.0105](https://arxiv.org/abs/1011.0105) (2010)
24. Jain, N., Wittmann, C., Lydersen, L., Wiechers, C., Elser, D., Marquardt, C., et al.: Device calibration impacts security of quantum key distribution. *Phys. Rev. Lett.* **107**, 110501 (2011)
25. Waks, E., Zeevi, A., Yamamoto, Y.: Security of quantum key distribution with entangled photons against individual attacks. *Phys. Rev. A* **65**, 052310 (2002)
26. Ma, X., Fung, C.-H.F., Lo, H.-K.: Quantum key distribution with entangled photon sources. *Phys. Rev. A* **76**, 012307 (2007)
27. Weedbrook, C.: Continuous-variable quantum key distribution with entanglement in the middle. *Phys. Rev. A* **87**, 022308 (2013)
28. Ralph, T., Lund, A.: Nondeterministic noiseless linear amplification of quantum systems. In: AIP Conference Proceedings, pp. 155–160 (2009)
29. Walk, N., Lund, A.P., Ralph, T.C.: Nondeterministic noiseless amplification via non-symplectic phase space transformations. *New J. Phys.* **15**, 073014 (2013)
30. McMahon, N., Lund, A., Ralph, T.: Optimal architecture for a nondeterministic noiseless linear amplifier. *Phys. Rev. A* **89**, 023846 (2014)

31. Bernu, J., Armstrong, S., Symul, T., Ralph, T.C., Lam, P.K.: Theoretical analysis of an ideal noiseless linear amplifier for Einstein-Podolsky-Rosen entanglement distillation. *J. Phys. B: At. Mol. Opt. Phys.* **47**, 215503 (2014)
32. Blandino, R., Leverrier, A., Barbieri, M., Etesse, J., Grangier, P., Tualle-Brouri, R.: Improving the maximum transmission distance of continuous-variable quantum key distribution using a noiseless amplifier. *Phys. Rev. A* **86**, 012327 (2012)
33. Usenko, V.C., Grosshans, F.: Unidimensional continuous-variable quantum key distribution. *Phys. Rev. A* **92**, 062337 (2015)
34. Grosshans, F., Cerf, N.J., Wenger, J., Tualle-Brouri, R., Grangier, P.: Virtual entanglement and reconciliation protocols for quantum cryptography with continuous variables. arXiv preprint quant-ph/0306141 (2003)
35. Serafini, A., Paris, M., Illuminati, F., De Siena, S.: Quantifying decoherence in continuous variable systems. *J. Optics B Quant. Semiclassical Opt.* **7**, R19 (2005)

# Research on Internet of Vehicles' Privacy Protection Based on Tamper-Proof with Ciphertext

Qifan Wang<sup>1</sup>, Guihua Duan<sup>1</sup>, Entao Luo<sup>2</sup>, and Guojun Wang<sup>3</sup>(✉)

<sup>1</sup> School of Information Science and Engineering, Central South University, Changsha 410083, China

<sup>2</sup> School of Electronics Information and Engineering, Hunan University of Science Engineering, Yongzhou 425199, China

<sup>3</sup> School of Computer Science and Educational Software, Guangzhou University, Guangzhou 510006, China  
csgjwang@gmail.com

**Abstract.** With the development of the internet of vehicles, the safety of the wireless network communication in the internet of vehicles and the privacy protection of users are in need of an immediate solution. In order to address the interception and tampering by an attacker of the contents among the vehicle mobile terminals, and their consequence of the privacy leakage of the location, and the user's activity tracks and also the consequence of the user's receiving the wrong information due to the tampering of the message, this essay comes up with a solution of tamper-proof ciphertext of the revised ciphertext policy attribute-based encryption, which to some degree has reduced the risk of the leakage of the user's information. This solution realize the encryption protection of the contents of the vehicle mobile terminals by destroying the structure extension of the ciphertext, achieves the goal of non-extensibility security.

**Keywords:** Internet of vehicles · Privacy protection  
Attribute-based encryption · Tamper-proof ciphertext  
Ciphertext-policy access control

## 1 Introduction

### 1.1 Background

Recently, Intelligent Traffic System (ITS) has become the guiding direction of the future transportation. As the extension of the internet of things in the Intelligent Traffic System area, the internet of vehicles is the key part of the Intelligent Traffic System. The internet of vehicles allows real-time control of the traffics on the road and improves the efficiency and safety of the traffic by fully sensing the roads and the traffic conditions and realizing the large-scale and high-capacity communication and interaction among multiple traffic systems [1].

However, the internet of vehicles communicate by using wireless channels, which inevitably face many threats and attacks, for example, the attacker may put wrong or error information in the internet system of the internet of vehicles or he may revise or repeat previous information. These threats and attacks would result in severe consequences for the transmission process has something to do with privacy information such as location and identity in the internet of vehicles.

The communication among the nodes of the vehicles in the internet of vehicles takes DSRC (Dedicated Short Range Communication) [2]. But in the communication process, the services provider, *SP*, cannot be trusted totally so that the data of the nodes stored in the cloud server may be at safety risk, for example, *SP* may provide the data of nodes of the vehicles to the third party without the authorization of the vehicle nodes. The third party may take advantage of the data to grasp the location and the range of activity of the nodes of the vehicles, then possibly make some harmful actions such as selling information to the advertisement company. So, in this case, it is usually necessary to encrypt the location, the ID of the vehicles, the user's identity and the user's outing habits.

## 1.2 Related Work

Many scholars have come up with their solutions of the safety problems of the internet of vehicles. Reference [3] puts forward a pseudonym-changing strategy based on the social points and the anonymous analysis model. References [4,5] put forward the anonymous authenticated and key agreement protocol which solves the problem of the protection of the location information of the vehicles, but in this strategy, the pseudonym doesn't need to be changed, which adds to the total cost. References [6,7] put forward an effective privacy protection model for the identity privacy, data privacy and location privacy, but a solution of the data privacy protection in the process of communication among the nodes of the vehicles is not mentioned. Reference [8] puts forward a new frame ACPN based on the public key cryptography system and pseudonym in terms of the privacy protection and the non-repudiation authentication, and effectively solve the problem of non-repudiation of the vehicle's identity in the internet of vehicles, but this frame lacks the fine grained access control and user doesn't have easy access control authority. Reference [9] puts forward a safe communication protocol of the internet of vehicles based on the key agreement protocol, symmetric encryption and message authentication. Reference [10] puts forward a portable electronic money scheme based on a centralized AAA structure, which address the identity privacy and data privacy in the privacy protection of the internet of vehicles, but it doesn't solve the problem of attacker's tampering of the communication contents in the communication process of the nodes of vehicles.

This essay puts forward a strategy of tamper-proof with ciphertext of the ciphertext policy attribute-based encryption based on the fine grained access control. This strategy is based on Ciphertext Policy Attribute-Based Encryption, CP-ABE, and the vehicle nodes makes its own access policy according to the attribute of them that receive the message. Besides, it is very flexible

for the decoding process happens only when the attribute of the vehicle nodes that receive the message meet the requirement. At the same time, to avoid the message is intercepted and then tampered, the trusted authority use the hash function to combine the cipher text and the related parameters into mapping, so that the receiver will first confirm whether the message is tampered or not before decrypting the message. This strategy ensures the safety of the communication process among vehicle nodes and improves the interaction efficiency among the vehicle nodes. The contributions of this paper are as follows:

- (1) **Fine-grained Access Control:** This strategy takes the policy of CP-ABE, the phase of encryption is controlled by access policy, and the access policy is contained in the ciphertext, only the *OBU* that meets the request of attributes could receive and decrypt the ciphertext.
- (2) **Flexible Key Management:** Each *OBU* has a set of personal attribute which is associated with the *OBU*'s private key in our scheme. So different *OBU* has an independent private key, which can avoid the leakage risk of key sharing.
- (3) **Tamper-Proof:** In the encryption and decryption process, the hash map guarantees the integrity of the ciphertext. The ciphertext can be prevented from being tampered with by the hash map.

## 2 Preliminaries

In this section, some preliminaries related to bilinear maps, complexity assumptions and access structure are presented.

### 2.1 Bilinear Maps

Let  $\mathbb{G}_1$  and  $\mathbb{G}_2$  be two multiplicative cyclic groups with big prime order  $p$ . Let  $g$  be a generator of  $\mathbb{G}_1$ . Let  $e$  be a bilinear map  $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$  with the following properties [11, 12]:

- (a) **Bilinearity:** For all  $g, h \in \mathbb{G}_1$  and  $a, b \in \mathbb{Z}_q$ , the equation  $e(g^a, h^b) = e(g, h)^{ab}$  holds.
- (b) **Non-degeneracy:** There exists  $g, h \in \mathbb{G}_1$ ,  $e(g, h) \neq 1$ .
- (c) **Computability:** There exists an efficient algorithm to compute bilinear map  $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ .

### 2.2 Access Structure and Access Tree

**Definition 1 (Access Structure [13, 14]):** Let  $\{P_1, P_2, \dots, P_n\}$  be a set of parties. A collection  $A \subseteq 2^{\{P_1, P_2, \dots, P_n\}}$  is monotone if  $\forall B, C$ : if  $B \in A$  and  $B \subseteq C$ . An access structure (respectively, monotonic access structure) is a collection (respectively, monotone collection)  $A$  of non-empty subsets of  $\{P_1, P_2, \dots, P_n\}$ , i.e.,  $A \subseteq 2^{\{P_1, P_2, \dots, P_n\}} \setminus \{0\}$ . The sets in  $A$  are called the authorized sets, and the sets not in  $A$  are called the unauthorized sets.

**Definition 2 (Access Policy Tree):** Access policy trees are often used to describe access policies. Each internal node of the access policy tree is a threshold, such as “and”, “or” and “n of m”. Each leaf node is represented by an attribute value. From the secret sharing idea, each node in the access policy tree is a secret, and the closer it is to the root node, the greater the permissions [15].

### 2.3 Security Assumption

**Definition 3:** Bilinear Diffie-Hellman Inversion Assumption (Decisional BDH problem) [16].

The challenger randomly selects four parameters  $a, b, c, z$  and then throws out a fair binary coin  $\beta$ . If  $\beta = 1$ , output tuple  $(A = g^a, B = g^b, C = g^c, e(g, g)^{abc})$ ; If  $\beta = 0$ , output  $(A = g^a, B = g^b, C = g^c, e(g, g)^z)$ , adversary  $A$  gives a guessing value  $\beta'$ . If the following formula is satisfied:

$$|Pr[B(g, g^a, g^b, g^c, e(g, g)^{abc}) = 1] - Pr[B(g, g^a, g^b, g^c, e(g, g)^z) = 1]| \geq \varepsilon$$

Then the algorithm  $B$  has the advantage  $\varepsilon$  of solving the Decisional BDH assumption. If the algorithm does not exist, it can solve the above games in a polynomial time  $t$  with a non-negligible advantage  $\varepsilon$ : let's say the  $(t, \varepsilon)$  DBDH assumption holds.

### 2.4 CP-ABE Security Model

In security definition, it has the following provisions: the access policy in ciphertext  $\Psi$  will be challenged by the adversary. During the challenge, the adversary has the access of any of the private keys  $SK$ . The security model is defined as follows [17, 18]:

**Setup:** The adversary selects a challenging access policy  $\Psi$ .

**System Establishment:** The challenger runs the Setup algorithm in the scheme, generates the system public key  $PK$  and the system master key  $MK$ , the challenger is reserved  $MK$ , and  $PK$  will be sent to the adversary.

**Phase 1:** The adversary requests the corresponding private key to construct attribute set  $\{S_1, S_2, \dots, S_q\}$ , but any of the attributes in the attribute set does not match the access policy  $\Psi$ . To generate the corresponding private key  $SK_i$ ,  $1 \leq i \leq q$ , the challenger use the construction method  $KenGen()$  in the private key scheme, and sent to the adversary.

**Challenge:** The adversary selects two plaintext of equal length  $m_0, m_1$  and sends them to the challenger. The challenger randomly selects a bit value  $b \in \{0, 1\}$ , encrypts the plaintext message  $m_b$  by using the access policy  $\Psi$ , and sends the ciphertext  $CT_b$  to the adversary.

**Phase 2:** Repeat phase 1 operation.

**Guess:** The adversary outputs a guess of  $b$ , to guess a value of  $b$ .



Above the challenge game mentioned, the adversary’s advantage definition:  $Pr[b' = b] = -1/2$ . If a polynomial time algorithm with non-negligible advantage in solving the problem does not exist in the game, then the CP-ABE scheme under chosen plaintext attacks is safe.

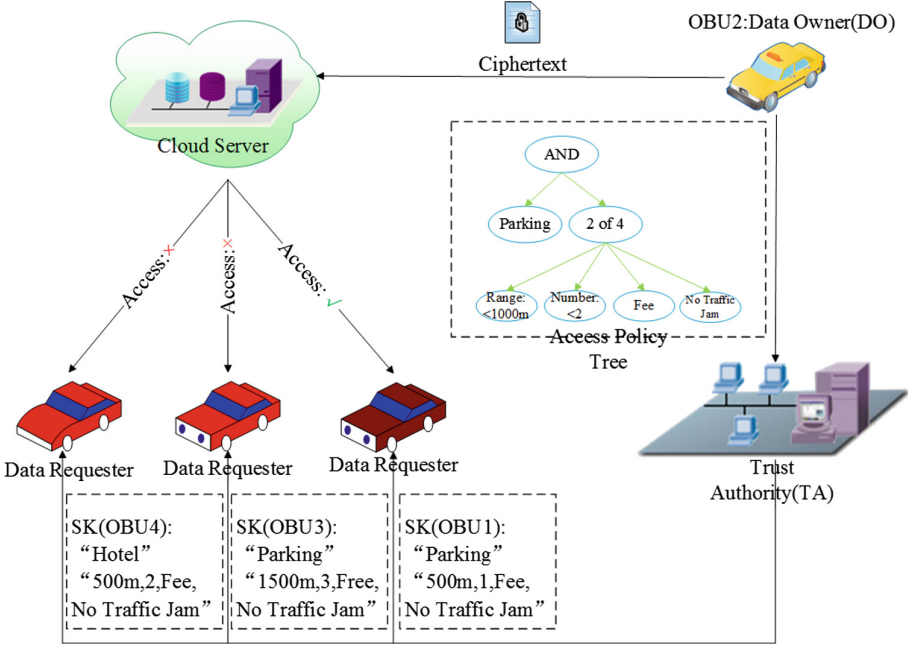


Fig. 1. System model diagram.

## 2.5 Structure Definition

The model of this essay consists of the following parts: the Server Provider (*SP*), the Trust Authority (*TA*), the Data Owner (*DO*) and the Data Requester (*DR*).

*TA* is totally reliable and *SP* is honest but curious: *SP* will surely deal with the data according to the related algorithm and the protocol, and it may also pry into the data stored in the cloud end (Fig. 1).

*TA* is to initialize the system and create and then give out the key.

*SP* is to store the encrypted cipher text sent by the owner, which includes the location, the vehicle number, the parking lot information etc.

*DO* is to make the access policy and revise and delete the message. In this essay, *OBU2* is the owner of the vehicle’s information, the ciphertext will only be encrypted correctly when the solicitant meets the access policy made by the *OBU2*.

*DR* is to issue the request of the surrounding information of the vehicles, in this essay, *OBU1* and other vehicle nodes are the requester.

This strategy uses for reference Bethencourt, Sahai and Waters' CP-ABE strategy [13] which is based on prime order group. Their strategy consists of four stages: the system initialization stage, the creation of the secret key stage, the message encryption stage and the message decryption stage.

The vehicle nodes *OBU2* uploads the self-defined access control policy to the *TA*, and then *TA* manages the attribute sets within the domain and gives out the corresponding user secret key for the attribute sets owned by the *OBU1*. *OBU2* will encrypt the message for fear that the location information may be unsafe, and then uploads *CT* to *SP*. *OBU1* downloads the ciphertext from the *SP* and meets the access control policy by its own *SK*. If the matching process is successful, the *OBU1* will get the message sent by *OBU2*; if not, *OBU1* will not get the message.

## 2.6 Scheme Definition

The scheme of this paper consists of the following four algorithms:

1.  $Setup(params, S_{user}) \rightarrow (PK, MK, GP)$ . Input the parameters  $params$  and the attributes set of the vehicle data requester  $S_{user}$ , then output the system public key  $PK$ , the system master key  $MK$  and the common parameters  $GP$ .
2.  $KeyGen(MK, S_{user}, GP) \rightarrow SK$ . Input the system master key  $MK$ , the attributes set of the vehicle data requester  $S_{user}$  and the common parameters  $GP$ , then output the private key  $SK$ .
3.  $Enc(PK, m, \Psi, GP) \rightarrow CT$ . Input the system public key  $PK$ , plaintext  $m$ , the access policy  $\Psi$  formulated by the vehicle data owner and the common parameters  $GP$ , then output ciphertext  $CT$ .
4.  $Dec(SK, CT, GP) \rightarrow m$ . Input the private key  $SK$ , ciphertext  $CT$  and the common parameters  $GP$ , output plaintext  $m$ .

## 3 Scheme Model

### 3.1 System Initialization Phase

In this scheme, *OBU2* chooses the security parameter  $\kappa$  which determines the scale of bilinear groups. *TA* chooses two cyclic groups  $G_0$  and  $G_1$ , whose order is the prime number  $p$ . It also randomly chooses the generator  $g$ . Let  $e : G_0 \times G_0 \rightarrow G_1$  denote a bilinear mapping, then it defines the hash function  $H : \{0, 1\}^* \rightarrow G_0$ , this function can translates the attributes of users which represented by some strings in random length into the random digits in bilinear groups. Randomly choose  $\alpha, \beta \in Z_q$ , then generate the system public key (Fig. 2):

$$PK = (G_0, g, h, h = g^\beta, e(g, g)^\alpha) \quad (1)$$

The system master key:

$$MK = (\beta, g^\alpha) \quad (2)$$

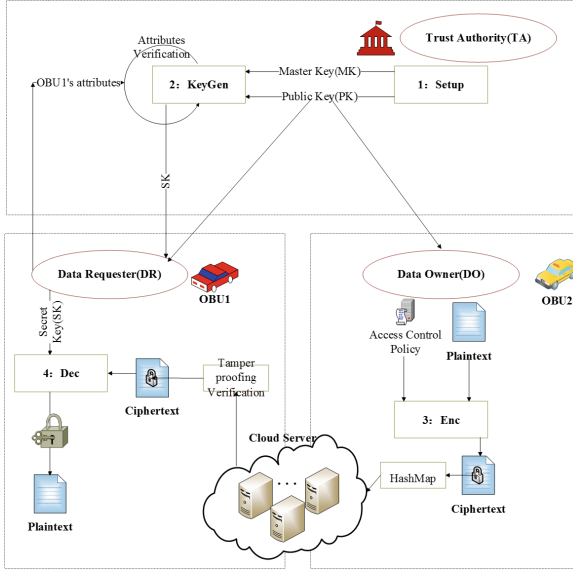


Fig. 2. Flow process diagram.

### 3.2 User Private Key Generation Phase

Suppose the *OBU1* asks for the information about parking spaces which learned by the *OBU2* of the internet of vehicles system. *OBU1* submits its own property sets to the trusted authority *TA*. The attribute sets of *OBU1* are matched with the *OBU2*'s own access policy by *TA*, when the attributes successfully matched with the nodes of the access policy, *TA* calculates the *OBU1*'s private key and sends it to *OBU1* on the basis of the master key *MK* and the system public key *PK* that it saves. The private key is kept by the vehicle node who has requested for the private key. *KenGen()* algorithm takes the user's attribute sets *S*, the master key *MK* and the system public key *PK* as input, then it randomly chooses  $ru \in Z_q$  represent for *S*,  $\forall i \in S$ , and chooses random number  $ru_i \in Z_q$ . Finally, the private key of *OBU1* is generated:

$$SK = (S, D = g^{(\alpha+ru)/\beta}, \{D_i = g^{ru} \cdot H(i)^{ru_i}, D'_i = g^{ru_i}\}_{\forall i \in S}) \quad (3)$$

### 3.3 Encryption Phase

*OBU2*, the information provider, will encrypt *m* before uploading *m* to the server. *OBU2* first makes the access policy  $\Psi$ , and together with the *PK* which is generated by *TA* to be encrypted, the ciphertext is generated. The leaf nodes represent attribute (the price of the hotel, the parking lot and the remaining parking spaces) and the non-leaf nodes represents threshold.

When the *OBU2* sends the ciphertext to the cloud, if there exists an attacker *Mallory*, he intercepts the ciphertext  $C_1 = me(g, g)^{\alpha\lambda}$  and then tampers it into

$C_1 = me(g, g)^{\alpha\lambda}m_1$  although he cannot decrypt the ciphertext, and then he acts as *OBU2* sends the revised ciphertext. When the legal *OBU1* receives the ciphertext and decrypts the ciphertext as  $m'$ , not  $m$ , and *OBU1* will take it as the message sent by *OBU2*. Thus, the attacker *Adversary* achieves his goal of tampering the ciphertext. Thus, *OBU2* combines the attribute sets  $S$  and  $C_1, C_2$  into a random number of mapping through hash function, and finally encrypt it together with ciphertext and then send them to the cloud.

Every leaf node and non-leaf node from the root node of the access policy tree should choose a corresponding polynomial  $q_x$ , and use  $d_x$  to represent the step of  $q_x$ . For the leaf node,  $d_x = 0$ ; For the non-leaf node,  $d_x = k_x - 1$ .  $k_x$  is the threshold of the nodes. For the OR,  $k_x = 1$ ; for the AND,  $k_x$  represents the number of the nodes. If the nodes are the root nodes, then the algorithm chooses random number  $\lambda \in Z_q$ ,  $q_R(0) = \lambda$ , and then define  $d_R$  random points as  $q_R$ , determine the threshold  $q_x$  by using Lagrange Polynomial; if nodes are the nodes other than the root nodes, then  $q_x(0) = q_{parent(x)}(index(x))$ , and  $index(x)$  returns the number of the node that corresponding the node  $x$ , then randomly choose  $d_x$  points to define  $q_x$ .

To calculate the following:

$$(C_1 = me(g, g)^{\alpha\lambda}, C_2 = h^\lambda, C_3 = H(S, C_1, C_2)^\lambda) \quad (4)$$

Use  $\Lambda$  to represent the sets of all the leaf nodes of the access policy tree  $\Psi$ , and then the ciphertext of the access policy tree  $\Psi$  in the  $TA$ :

$$\begin{aligned}
 &(\Psi, C_1 = me(g, g)^{\alpha\lambda}, C_2 = h^\lambda, C_3 = H(S, C_1, C_2)^\lambda, \\
 &\{C_\gamma = g^{q_\gamma(0)}, C'_\gamma = H(att(\gamma))^{q_\gamma(0)}\}_{\forall \gamma \in \Lambda})
 \end{aligned} \quad (5)$$

### 3.4 Decryption Phase

$SP$  sends the ciphertext  $CT$  which encrypted by *OBU2* to *OBU1* when *OBU1* requests access to the message. *OBU1* could use the private key that is previously generated to decrypt the message correctly if and only if the attributes of *OBU1* successfully matched with the access policy of  $TA$ . The decryption algorithm first calls the method  $Tree(S)$  to determine whether the private key  $SK$  meets the access policy tree  $\Psi$  in the ciphertext  $CT$ . For any node  $x$  in the access policy  $\Psi$ , if  $x$  is the non-leaf node, algorithm will calculate  $Tree(S)$  for all the child nodes of  $x$ , then  $Tree(S)$  returns a non-empty set  $S_x$  containing a subset number; if  $x$  is the leaf node, if and only if the attribute sets of  $S$  contain  $att(x)$ , then  $Tree(S)$  returns a non-empty set  $S_x$  containing a subset number.

The algorithm first confirm whether the ciphertext *OBU1* received is tampered or not, and then continue, if tampered, a null value will be returned. Calculate  $H' = H(S, C_1, C_2)$ , and confirm  $e(h, C_3) = e(H', C_2)$ , in the calculation,

$$e(h, C_3) = e(h, H(S, C_1, C_2)^\lambda) = e(h, H(S, C_1, C_2))^\lambda \quad (6)$$

$$e(H', C_2) = e(H(S, C_1, C_2), h^\lambda) = e(H(S, C_1, C_2), h)^\lambda \quad (7)$$

If the ciphertext received by the *OBU1* is not tampered, then the two hash value equals to each other, in reference of the bilinearity of the bilinear mapping:  $e(h, C_3) = e(H', C_2)$ .

If the confirmation succeeds, then continues, if not, a null value is returned.

$Tree(S)$  goes by recursive way, for any node  $x$  in  $\Psi$ ,  $Tree_x(S)$  will return a non-empty set  $S_x$  consisting of subset number, if  $S$  doesn't meet  $\Psi$ , then  $Tree(S)$  returns a null value. Otherwise, the algorithm goes recursive algorithm  $DecryptNode(CT, SK, x)$ , in this, the function takes ciphertext  $CT$ ,  $SK$ , any node  $x$  of  $\Psi$  as input. If  $x$  is leaf node, then  $j = att(x)$ ,  $H(j)$  is an element on  $G_0$ , we may suppose  $H(j) = g^n$  and when it meets  $j \in S$ , calculate:

$$\begin{aligned}
& DecryptNode(CT, SK, x) \\
&= \frac{e(D_j, C_x)}{e(D_{j'}, C_{x'})} = \frac{e(g^r \cdot H(j)^{r_j}, g^{q_x(0)})}{e(g^{r_j}, H(j)^{q_x(0)})} \\
&= \frac{e(g^r \cdot g^{\eta \cdot r_j}, g^{q_x(0)})}{e(g^{r_j}, g^{\eta \cdot q_x(0)})} = \frac{e(g, g)^{(r+\eta \cdot r_j) \cdot q_x(0)}}{e(g, g)^{r_j \cdot \eta \cdot q_x(0)}} \\
&= e(g, g)^{r \cdot q_x(0)}
\end{aligned} \tag{8}$$

If  $j \notin S$ ,  $DecryptNode(CT, SK, x) = \perp$ ; for non-leaf nodes  $x$ , use its returned value  $F_z$  of the sub node  $z$ . Then save the result in  $F_z$  :  $F_z = DecryptNode(CT, SK, z)$ .

Next, calculate  $F_z$  of every sub node of  $z \in S_x$ , and then use Lagrange Interpolation method to get the  $F_x$  of the node  $x$ ,  $F_x = \prod_{z \in S_x} F_z^{\Delta iz, S'_z}$ ,

( $iz = index(z), S'_z = \{index(z) : z \in S_x\}$ ), the lagrange coefficient is  $\Delta iz, S'_z(0) = \prod_{jz \in S'_z, jz \neq iz} \frac{0-jz}{iz-jz}$ , then the function  $F_x$  of the node  $x$  is:

$$\begin{aligned}
F_x &= \prod_{z \in S_x} F_z^{\Delta iz, S'_z} = \prod_{z \in S_x} (e(g, g)^{r \cdot q_z(0)})^{\Delta iz, S'_z} \\
&= \prod_{z \in S_x} (e(g, g)^{r \cdot q_{parent(z)}(index(z))})^{\Delta iz, S'_z} \\
&= \prod_{z \in S_x} e(g, g)^{r \cdot q_x(iz) \cdot \Delta iz, S'_z} \\
&= e(g, g)^{r \cdot q_x(0)}
\end{aligned} \tag{9}$$

Considering the above mentioned function  $DecryptNode()$ , run back upwards and then get the value of the function  $DecryptNode(CT, SK, R)$ :

$$F_R = e(g, g)^{r \cdot q_R(0)} \tag{10}$$

Considering the above thing,  $q_R(0) = \lambda$ , we can get:

$$F = e(g, g)^{r \cdot \lambda} \tag{11}$$

Then decrypt it, we can get:

$$\begin{aligned}
 m &= \frac{C_1}{F} = \frac{C_1}{e(g, g)^{\alpha \cdot \lambda}} \\
 &= \frac{m \cdot e(g, g)^{\alpha \cdot \lambda}}{e(g, g)^{\alpha \cdot \lambda}} \\
 &= m
 \end{aligned} \tag{12}$$

Then the decryption succeeds and we get the plaintext  $m$ . Next,  $OBU1$  can get all the information that  $OBU2$  have about the information of the parking lot space.

## 4 Security Analysis and Verification

### 4.1 Data and Communication Security

In this strategy, the message is first encrypted and then stored on  $SP$ , so the message largely remains safe. The vehicle nodes  $OBU2$  makes access control policy and then encrypts the message and then sends it to the cloud, so no matter the  $SP$  or any other illegal  $OBU$  cannot by any means obtain the encrypted message, thus unable to obtain any location information or any other information of  $OBU2$ . Due to the policy of CP-ABE, if illegal  $OBU$  or other attackers appear in the communication process and they cannot meet the access control policy, they will be unable to decrypt the message so that they cannot get the plain text.

### 4.2 Security Proof

**Definition 4:** If in the polynomial time, the adversary can win the above game with negligible advantage, so the proposed scheme can achieve CPA security [19–21].

**Proof:** If the adversary can break the security model, there is at least one polynomial time algorithm which can solve the DBDH problem without the negligible advantage.

A bilinear map  $e : \mathbb{G}_0 \times \mathbb{G}_0 \rightarrow \mathbb{G}_1$  defined on a multiplicative cyclic group  $\mathbb{G}_0$  whose prime number is a prime number  $p$  and a generator  $g$ . A challenger  $C$  is built here, and the challenger  $C$  throws a coin according to the fair coin toss agreement, then system set the corresponding tuple: randomly choose four parameters  $a, b, c, z \in \mathbb{Z}_q$ , if  $b = 0$ ,  $(g, A, B, C, Z) = (g, g^a, g^b, g^c, e(g, g)^{abc})$ ; if  $b = 1$ ,  $(g, A, B, C, Z) = (g, g^a, g^b, g^c, e(g, g)^z)$ . In the next DBDH game:

**Setup:** The adversary selects a challenging access policy  $\Psi$ .

**System Establishment:** The challenger  $C$  runs the Setup algorithm in the scheme, the challenger  $C$  randomly choose  $\alpha^0, \beta^0, a, b \in \mathbb{Z}_q'$ , the equation

$e(g, g)^{ab} = e(A, B)$  holds. Generates the system public key  $PK_0 = \{G_0, g, h, h = g^{\beta_0}, e(g, g)^{\alpha_0}\}$  and the system master key  $MK_0 = \{\beta_0, g^{\alpha_0}\}$ , the challenger is reserved  $MK_0$ , and  $PK_0$  will be sent to the adversary.

**Phase 1:** The adversary requests the corresponding private key to construct attribute set  $\{S_1, S_2, \dots, S_q\}$ , but any of the attributes in the attribute set does not match the access policy  $\Psi$ ,  $\Psi(S_i) = 0, 1 \leq i \leq q$ . To generate the corresponding private key  $SK_i, 1 \leq i \leq q$ , the challenger use the method  $KeyGen() : D = g^{(\alpha+ru)/\beta}, D_i = g^{ru} \cdot H(i)^{ru_i}, D_i' = g^{ru_i}$  in the private key scheme, and sent to the adversary.

**Challenge:** The adversary selects two plaintext of equal length  $m_0, m_1$  and sends them to the challenger. The challenger randomly selects a bit value  $\eta \in \{0, 1\}$ , encrypts the plaintext message  $m_b$  by using the access policy  $\Psi$ , and sends the ciphertext  $CT_b$  to the adversary. the ciphertext  $CT_b$  is:

$$CT_b = \{\Psi, C_1 = m_b \cdot Z, C_2 = h^\lambda, \{C_\gamma = g^{q_\gamma(0)}, C'_\gamma = H(att(y))^{q_\gamma(0)}\}_{\forall \gamma \in \Lambda}\}$$

If  $b = 0$ , from the definition previously available:  $Z = e(g, g)^{abc}, c = \lambda, C_1 = m \cdot Z = m \cdot e(g, g)^{abc} = m \cdot e(g, g)^{a\lambda}$ , then  $CT_b$  is a valid ciphertext.

If  $b = 1$ , from the definition previously available:  $Z = e(g, g)^z, C_1 = m \cdot Z = m \cdot e(g, g)^z$ , since  $z$  is randomly chosen,  $G_1$  must be a random element on  $G_0$  and it will not disclose any information about  $m$ .

**Phase 2:** Repeat phase 1 operation.

**Guess:** If the adversary gives the correct guess,  $\eta = \eta'$ , then the challenger outputs  $b' = 0$ , it indicates that the tuple received by challenger is  $(g, g^a, g^b, g^c, e(g, g)^{abc})$ ; otherwise, the challenger outputs  $b' = 1$ , it indicates that the tuple received by challenger is  $(g, g^a, g^b, g^c, e(g, g)^z)$ .

In the above DBDH game, if  $b = 0$ , the adversary can get the ciphertext  $m$ , according to the previous definition, the adversary can break our scheme with the non-negligible advantage  $\varepsilon$ , so  $Pr[\eta = \eta' | b = 0] = 1/2 + \varepsilon$ . When  $\eta' = \eta$ , the challenger will guess  $b' = 0$ , so  $Pr[b' = b | b = 0] = 1/2 + \varepsilon$ ; if  $b = 1$ , the adversary won't receive any information about  $m$ , so  $Pr[\eta' \neq \eta | b = 1] = 1/2$ . When  $\eta' \neq \eta$ , challenger guesses  $b' = 1$ , so  $Pr[b' = b | b = 1] = 1/2$ .

Given all that, the advantage of challenger correctly guessed  $b' = b$  as:

$$\begin{aligned} Adv_c &= Pr[b' = b] - 1/2 \\ &= Pr[b' = b, b = 0] + Pr[b' = b, b = 1] - 1/2 \\ &= 1/2 Pr[b' = b | b = 1] + 1/2 Pr[b' = b | b = 0] - 1/2 \\ &= 1/2 Pr[b' = 0 | b = 0] + 1/2 Pr[b' = 1 | b = 1] - 1/2 \\ &= 1/2 Pr[\eta' \neq \eta | b = 1] + 1/2 Pr[\eta' = \eta | b = 0] - 1/2 \\ &= 1/2 \times 1/2 + 1/2 \times (1/2 + \varepsilon) - 1/2 \\ &= \varepsilon/2 \end{aligned}$$

We can get the conclusion by  $Adv_c = \varepsilon/2$ : if the adversary break the scheme with the non-negligible advantage  $\varepsilon/2$  in the DBDH game, there is a polynomial time algorithm with non-negligible advantage in solving the problems of DBDH. This problem contradicts the result that DBDH problem is difficult, so there isn't a polynomial time algorithm to solve the DBDH problem with non-negligible advantage, namely: the adversary cannot break this scheme with non-negligible advantage, our scheme can achieve CPA security.

The security of the scheme is proved, then the ciphertext can't be tampered would be proved in the following scheme.

To conclude, we can see that the message  $m$  which is encrypted:

$$CT = \{\Psi, C_1 = me(g, g)^{\alpha\lambda}, C_2 = h^\lambda, C_3 = H(S, C_1, C_2)^\lambda, \\ \{C_\gamma = g^{q_\gamma(0)}, C'_\gamma = H(att(y))^{q_\gamma(0)}\}_{\forall \gamma \in \Lambda}\}$$

Among that,  $H' = H(S, C_1, C_2)$ . From the form of ciphertext, the ciphertext that has not been tampered must be satisfies with the following relations:  $e(h, C_3) = e(H', C_2)$ .  $C_1, C_2$  in ciphertext can be faked successfully, but  $C_3$  cannot be faked successfully, because the attacker does not know the  $\lambda$  which is randomly selected by the encryption agent. To sum up, if an attacker has tampered with  $m$ , the decryption agent will be able to detect and reject the decryption, so that it can resist the ciphertext tampering attack.

## 5 Performance Analysis

We will compare some security requirements in this section. From the Table 1 we can see that our scheme meets the fine-grained access control. Although the other three programs can also meet the requirements of message integrity and conditional privacy, only our scheme achieves the goal of ciphertext tampering. The following is the comparison between our scheme and related schemes.

**Table 1.** Comparison with related schemes

	Our scheme	Rajput et al.'s scheme [5]	Li et al.'s scheme [8]	Choi et al.'s scheme [22]
Fine-grained access	Yes	No	No	No
Message integrity	Yes	Yes	Yes	Yes
Conditional privacy	Yes	Yes	Yes	Yes
Tamper-proof	Yes	No	No	No
Scalability	Yes	Yes	No	No



## 6 Conclusion

Internet of vehicles is the perfect combination of traditional outing and high technology outing, in the same time, the country is fully supporting the intelligent traffic, the internet of vehicles thus has a more potential future. However, due to the many problems that it has, because of its combination with the internet, the communication among the vehicle nodes, the communication between the vehicle nodes and the facilities around the roads are the top choices of study for future researches. This strategy takes the cryptology as basis, having realized the privacy protection among vehicle nodes and making the owner of the vehicles able to discover cautiously the vehicle nodes information requester which meets their access policy that preciously set, thus ensure the safety of the communication process.

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation of China under Grants 61632009, 61472451 and 61772194, in part by the Guangdong Provincial Natural Science Foundation under Grant 2017A030308006 and High-Level Talents Program of Higher Education in Guangdong Province under Grant 2016ZJ01.

## References

1. Chang, W., Zheng, H., Wu, J.: On the RSU-based secure distinguishability among vehicular flows. In: *IEEE/ACM 25th International Symposium on Quality of Service (IWQoS)*, pp. 1–6. IEEE (2017)
2. Peng, X., Shen, Q., Li, W., et al.: Data dissemination based on trajectory for sparse VANETs. *J. Softw.* **27**(Suppl.(1)), 59–70 (2016)
3. Wang, D., Li, D., Li, X., et al.: An analysis of anonymity on capacity finite social spots based pseudonym changing for location privacy in VANETs. In: *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery, Zhangjiajie*, pp. 763–767 (2015)
4. Bttner, C., Bartels, F., Huss, S.A.: Real-world evaluation of an anonymous authenticated key agreement protocol for vehicular ad-hoc networks. In: *IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications, Abu Dhabi*, pp. 651–658 (2015)
5. Rajput, U., Abbas, F., Oh, H.A.: A hierarchical privacy preserving pseudonymous authentication protocol for VANET. *IEEE Access* **PP**(99), 1 (2016)
6. Tyagi, A.K., Sreenath, N.: Location privacy preserving techniques for location based services over road networks. In: *International Conference on Communications and Signal Processing, Melmaruvathur*, pp. 1319–1326 (2015)
7. Emara, K.: Location privacy in vehicular networks. In: *2013 IEEE 14th International Symposium on A World of Wireless, Mobile and Multimedia Networks, Madrid*, pp. 1–2 (2013)
8. Li, J., Lu, H., Guizani, M.: ACPN: a novel authentication framework with conditional privacy-preservation and non-repudiation for VANETs. *IEEE Trans. Parallel Distrib. Syst.* **26**(4), pp. 938–948 (2015)
9. Xie, Y., Wu, L., Zhang, Y., et al.: Anonymous mutual authentication and key agreement protocol in multi-server architecture for VANETs. *J. Comput. Res. Dev.* **10**, 2323–2333 (2016)

10. Yeh, L.Y., Huang, J.L.: PBS: a portable billing scheme with fine-grained access control for service-oriented vehicular networks. *IEEE Trans. Mob. Comput.* **13**(11), 2606–2619 (2014)
11. Wang, S., Liang, K., Liu, J.K., et al.: Attribute-based data sharing scheme revisited in cloud computing. *IEEE Trans. Inform. Forensics Secur.* **11**(8), 1 (2016)
12. Luo, E., Liu, Q., Wang, G.: Hierarchical multi-authority and attribute-based encryption friend discovery scheme in mobile social networks. *IEEE Commun. Lett.* **20**(9), 1772–1775 (2016)
13. Yadav, U.C., Ali, S.T.: Ciphertext policy-hiding attribute-based encryption. In: *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Kochi, pp. 2067–2071 (2015)
14. Ning, J., Dong, X., Cao, Z., et al.: White-box traceable ciphertext-policy attribute-based encryption supporting flexible attributes. *IEEE Trans. Inform. Forensics Secur.* **10**(6), 1 (2015)
15. Zhou, Z., Huang, D., Wang, Z.: Efficient privacy-preserving ciphertext-policy attribute-based encryption and broadcast encryption. *IEEE Trans. Comput.* **64**(1), 126–138 (2013)
16. Wei, J., Liu, W., Hu, X.: Provable secure attribute based authenticated key exchange protocols in the standard model. *J. Softw.* **25**(10), 2397–2408 (2014)
17. Zhang, S., Wang, G., Liu, Q., et al.: A trajectory privacy-preserving scheme based on query exchange in mobile social networks. *Soft. Comput.* (2017). <https://doi.org/10.1007/s00500-017-2676-6>
18. Phuong, T.V.X., Yang, G., Susilo, W.: Hidden ciphertext policy attribute-based encryption under standard assumptions. *IEEE Trans. Inform. Forensics Secur.* **11**(1), 1 (2015)
19. Luo, E., Liu, Q., Abawajy, J.H., et al.: Privacy-preserving multi-hop profile-matching protocol for proximity mobile social networks. *Future Gener. Comput. Syst.* **68**, 222–233 (2017)
20. Zhang, S., Choo Raymond, K.W., Liu, Q., et al.: Enhancing privacy through uniform grid and caching in location-based services. In: *Future Generation Computer Systems* (2017). <https://doi.org/10.1016/j.future.2017.06.022>
21. Xu, J., Wen, Q., Li, W., et al.: Circuit ciphertext-policy attribute-based hybrid encryption with verifiable delegation in cloud computing. *IEEE Trans. Parallel Distrib. Syst.* **27**(1), 119–129 (2016)
22. Choi, H., Kim, I., Yoo, J.: Secure and efficient protocol for vehicular ad hoc network with privacy preservation. *EURASIP J. Wirel. Commun. Netw.* **2011**(1), 1–15 (2011)

# An Attack to an Anonymous Certificateless Group Key Agreement Scheme and Its Improvement

Xuefei Cao<sup>(✉)</sup>, Lanjun Dang, Kai Fan, and Yulong Fu

School of Information and Cyber Security, Xidian University, Xi'an 710071, China  
xfcao@xidian.edu.cn

**Abstract.** In this paper, we demonstrate that Kumar-Tripathi's anonymous authenticated group key agreement protocol is insufficient in authenticity and unlinkability. Then the scheme is improved based on the Computational Diffie-Hellman (CDH) problem and Divisible Computational Diffie-Hellman (DCDH) problem. Compared with available schemes, the improved scheme satisfies strengthened security with lower computational overhead. The security is proven formally using AVISPA.

**Keywords:** Certificateless cryptography · Group key agreement  
Authentication · Anonymity · Bilinear pairing · Mobile communication

## 1 Introduction

The emergence of mobile communications facilitates our daily life greatly. Group-based mobile communications such as vehicular communications, teleconferences and remote learning attract customers to mobile communications and further promote its development. Authenticated Group Key Agreement (AGKA) protocols, which help to authenticate group members and establish secure group session key, provide a proper security solution for these applications [1, 2].

Compared with wired networks, user anonymity becomes an important issue in mobile networks [3]. Consider the scenario of a conversation in a group of vehicles along the road. The vehicles may want to share the information on their journey to the group members. They hope to keep the secrecy of their identity to the eavesdroppers on the road. When the conversation is over, the participants could keep on moving, and they always require that their movement be kept untraceable to others. Anonymity ensures that a user will not be identified. Unlinkability is an important aspect of user anonymity, it prevents anyone outside the protocol from linking two different protocol executions to the same user [4].

Certificateless Public Key Cryptography (CL-PKC) [5] removes the inborn issue of key escrow with the Identity-based cryptography [6] and the issue of certificate management with the traditional certificate-based public key cryptography. Since the first concrete certificateless key agreement protocol in [5],

many certificateless authenticated key agreement protocols have been proposed [7–9]. Certificateless Authenticated Group Key Agreement (CL-AGKA) protocols can be put forward based on the certificateless authenticated key agreement protocols [10–13]. The first CL-AGKA scheme was proposed by Heo in 2007 [14]. The protocol supports group key agreement and group membership change, i.e., users can join or leave the group dynamically. However, Lee *et al.* pointed out that Heo’s protocol didn’t satisfy the forward secrecy and improved the insufficiency in [15]. Teng *et al.* proposed a CL-AGKA with stronger security in [10]. Sun *et al.* proposed a CL-AGKA protocol in which the server doesn’t get the agreed group key at the end of a session [16]. The security of the protocol is proved with the random oracle model.

AGKA protocols with user anonymity are relatively new. In 2008, Wan *et al.* proposed an anonymous authenticated group key agreement protocol [17]. They employ the pairing-based signature and the identity-based anonymous encryption to provide authenticity and anonymity. Yang *et al.* proposed an anonymous CL-AGKA protocol without pairings for e-health systems [18]. Recently, Kumar and Tripathi proposed a CL-AGKA protocol, claiming that the protocol achieves user anonymity and unlinkability [19]. However, in this paper we demonstrate that the protocol fails to provide unlinkability and authenticity. We also improve the scheme and show the security and efficiency of the improved scheme.

**Contribution.** We demonstrate that Kumar-Tripathi’s scheme employs a flawed method. The same random number is used twice by a Key Generation Center (KGC) to generate the partial private key for a user, which may result in the break of the master private key, and it is the severest threat to certificateless cryptosystem. Secondly, a KGC couldn’t authenticate a user because the user’s public key is not verified. In addition, the scheme fails to satisfy unlinkability because a user’s partial private key component is transferred in plaintext. An attacker can link different protocols to the same user by the information. We then give a simple but adequate approach to refine the scheme. Our approach hashes a user’s public key into the user’s partial private key and modifies the KGC’s authentication towards a user.

The rest of the paper is arranged as follows: in Sect. 2 we introduce the preliminaries, Sect. 3 shows our attack, Sect. 4 improves the protocol. The improved protocol is analyzed in Sect. 5.1 while Sect. 6 concludes the paper.

## 2 Preliminaries

### 2.1 Elliptic Curve Cryptography

We denote by  $E/F_p$  an elliptic curve  $E$  over a prime finite field  $F_p$ . We define an elliptic curve by  $y^2 = x^3 + ax + b$  with  $a, b \in F_p$  and with the discriminant  $\Delta = 4a^3 + 27b^2 \neq 0$ . The points on  $E/F_p$  and an extra point  $\mathcal{O}$  called the *point at infinity* form a group  $G = \{(x, y) : x, y \in F_p; (x, y) \in E/F_p\} \cup \{\mathcal{O}\}$ .  $G$  is a cyclic additive group under the point addition “+” defined as follows: Let  $P, Q \in G$ ,  $l$  be a line connecting  $P$  and  $Q$  (tangent line to  $E/F_p$  if  $P = Q$ ),

and  $R$  is the third point of intersection of  $l$  with  $E/F_p$ . Then define a line  $l'$  connecting  $R$  and  $\mathcal{O}$ . Then  $P + Q$  is the point such that  $l'$  intersects  $E/F_p$  at  $R$ ,  $\mathcal{O}$  and  $P + Q$ . Scalar multiplication over  $E/F_p$  can be computed as follows:  $tP = P + P + \dots + P$  ( $t$  times).

There are two difficult problems defined over elliptic curve group:

**Computational Diffie-Hellman (CDH) Problem:** For  $a, b \in_R Z_p^*$  and  $P$  the generator of  $G$ , given  $(aP, bP)$ , computing  $abP$ .

**Divisible CDH (DCDH) Problem:** For  $a, b \in_R Z_p^*$  and  $P$  the generator of  $G$ , given  $(aP, bP)$ , computing  $ab^{-1}P$ .

It is assumed that CDH problem and DCDH problems are intractable in polynomial time in a security parameter used to define the problem instances.

## 2.2 Security Requirements

For a CL-AGKA protocol with user anonymity, the following security requirements should be satisfied [20, 21]:

- **Authenticity:** There are two folds in authenticity: (1) there should be mutual authentication between a user and a KGC when KGC is evolved in the group key establishment; and (2) the group members should be able to authenticate each other.
- **Group Forward/Backward Security:** A previous group member should not be able to access the current group conversation anymore, and a new group member should not be able to access the previous group conversation.
- **Forward Secrecy:** If the long term private keys of all the users are compromised, the previous group session keys should still be secure.
- **Perfect Forward Secrecy:** Even if the system master key is compromised, the previous group session keys are still secure.
- **Anonymity:** A group member's identity should not be transferred in the form of plaintext.
- **Unlinkability:** A group member's activity in two different executions in a protocol can't be linked by outside attackers in any way.

## 3 Attack to Kumar-Tripathi Scheme

Kumar-Tripathi scheme include three protocols, i.e., the group key establishment, user join and user leave. We proposes attacks to the group key establishment protocol only, the user join and the user leave are secure. The group key establish protocol has 6 algorithms. Due to the space limitation, we propose our attack without the description of the scheme. Readers are referred to [19] for the detailed information.

1. *Insider Attack to Master Key.* Assume that  $U_a$  is a legal system member with certificateless cryptography secret key  $SK_a = \langle x_a, s_a \rangle$ , then  $U_a$  can

obtain KGC's master key. To get the goal,  $U_a$  first completes the Algorithm 1 to Algorithm 5, and the *Member Registration for Anonymity* of Algorithm 6. By the end of the *Member Registration for Anonymity*,  $U_a$  receives  $\{s_{ta} \| h_{t1} \| Q_1 \| h_{t2} \| Q_2 \cdots \| h_{tn} \| Q_n\}_{K_{ax}}$  from KGC, and  $U_a$  could decrypt the message for  $s_{ta}$  with symmetric key  $K_{ax}$ . Then  $U_a$  has a CL-PKC secret key  $s_a$  and a temporary secret key  $s_{ta}$  and

$$\begin{cases} s_a &= (r_a + s \cdot h_a) \bmod p \\ s_{ta} &= (r_a + s \cdot h_{ta}) \bmod p \end{cases}$$

$U_a$  can obtain the system master key  $s$  by calculating:

$$s = (s_a - s_{ta})(h_a - h_{ta})^{-1} \bmod p$$

2. *Impersonation to KGC.* Suppose an attacker wants to impersonate a legal group member  $U_i$  to KGC,  $U_a^i$  denotes the attacker. To do this, the attacker execute the *Member Registration for Anonymity* in a different manner. It first chooses at random  $t_a \in_R Z_p^*$  and computes  $T_a = t_a \cdot P$ . The attacker then computes  $K_a = t_a \cdot P_{pub}$  and uses  $K_a$ 's  $x$ -coordinate  $K_{ax}$  as the encryption key. The attacker chooses a temporary identifier  $TID_a^i$  and sends the following message to KGC:

$$U_a^i \rightarrow KGC : ID_i, T_a, \{T_a, TID_a^i, Q_i\}_{K_{ax}}$$

On receiving  $U_a^i$ 's message, KGC computes  $\tilde{K}_a = sT_a$  and uses the  $x$ -component of  $\tilde{K}_a$  to decrypt  $\{T_a, TID_a^i, Q_i\}_{K_{ax}}$ . KGC could decrypt the message because  $\tilde{K}_a = sT_a = s t_a P = t_a P_{pub} = K_a$ . The  $T_a$  decrypted from the message is the same with  $T_a$  sent in plaintext, then according to the protocol, KGC would accept  $U_a^i$  as  $U_i$ .

3. *Outsider Attack to Unlinkability.* The authors claim that their protocol provides unlinkability. However, In the *Anonymous Key Establishment* of the same algorithm, the following message is sent from  $U_i$  to  $U_{i+1}$  and  $U_{i-1}$ :

$$U_i \rightarrow U_{i+1}, U_{i-1} : h_{ti}, R_i, W_i$$

In the settings of CL-PKC,  $R_i$  is a component of a user's partial private key which stays unchanged for every user  $U_i$  as long as the user's private key is unchanged. A passive attacker can collect  $R_i$  in different protocol executions and trace the user. The protocol fails to satisfy the unlinkability.

## 4 Our Improvement

In this section, we propose our improved scheme. Since the user join and user leave of the original scheme are secure, we only improve the group key establishment. The protocol consists of two phases, i.e., *initialization*, *key agreement with user anonymity*. The two phases are described as follows:

**Initialization:** This phase comprises five algorithms, i.e., *setup*, *set-secret-value*, *set-partial-secret-key*, *set-secret-key*, and *set-public-key* as follows:

1. *Setup*: This algorithm takes a security parameter  $k$  as input and outputs system parameters and a master key. Given  $k$ , KGC does the following:
  - (a) chooses a  $k$ -bit prime  $p$  and determines the tuple  $\{F_p, E/F_p, G, P\}$ .
  - (b) chooses the master private key  $s \in_R Z_p^*$  and sets  $P_{pub} = sP$  as the system's public key.
  - (c) selects two hash functions as  $H_1 : \{0, 1\}^* \rightarrow Z_p^*$  and  $H_2 : \{0, 1\}^* \rightarrow \{0, 1\}^k$ .
  - (d) publishes the system parameters  $param = \{F_p, E/F_p, G, P, P_{pub}, H_1, H_2\}$  and secretly keeps the master key  $s$ .
2. *Set-Secret-Value*: The user with identity  $ID_i$  picks randomly  $x_i \in_R Z_p^*$  and sets  $x_i$  as his secret value.
3. *Partial-Secret-Key-Extract*: The user with identity  $ID_i$  and secret value  $x_i$  computes  $Q_i = x_iP$ .  $U_i$  sends  $ID_i$  and  $Q_i$  to KGC, and KGC does the following:
  - (a) chooses  $r_i \in_R Z_p^*$  and computes  $R_i = r_iP$ .
  - (b) computes  $h_i = H_1(ID_i || R_i || Q_i)$  and  $s_i = r_i + s \cdot h_i \bmod p$ .
  - (c) sends the tuple  $\langle s_i, R_i \rangle$  to  $U_i$  via a secure channel.
 The user's partial secret key is the tuple  $D_i = \langle s_i, R_i \rangle$ .  $U_i$  can check the validity of  $D_i$  by checking if the equation  $s_iP = R_i + H_1(ID_i || R_i || Q_i)P_{pub}$  holds. The key is valid if it does and vice visa.
4. *Set-Secret-Key*: The user with identity  $ID_i$  sets his/her secret key as the pair

$$SK_i = \langle x_i, s_i \rangle$$

5. *Set-Public-Key*: The user with identity  $ID_i$  calculates  $P_i = (x_i + s_i)P$ .  $P_i$  is one component of the user's public key. The other component is  $Q_i = x_iP$  which is computed by the user him/herself using the user's secret value. The user's public key is the tuple

$$PK_i = \langle P_i, Q_i \rangle$$

**Key Agreement with User Anonymity:** Suppose a group of users  $\{U_1, U_2, \dots, U_n\}$  decide to agree an authenticated group session key with user anonymity. They do as follows:

1. For every user  $U_i$ , he/she randomly chooses a temporary identifier  $TID_i$ , a random  $t_i \in_R Z_p^*$  and calculates the following:

$$\begin{cases} T_i & = t_i P_{pub} \\ K_i & = t_i P = (K_{ix}, K_{iy}) \end{cases}$$

$K_i$  is a point in the elliptic curve group  $G$ , and  $(K_{ix}, K_{iy})$  are  $K_i$ 's  $x$ -coordinate and  $y$ -coordinate.

Then  $U_i$  sends the following message to KGC:

$$U_i \rightarrow KGC : T_i, \{ID_i, TID_i, R_i, Q_i\}_{K_{ix}}$$

The message is encrypted with  $K_{ix}$  using a symmetric encryption algorithm.

2. On receiving  $U_i$ 's message ( $1 \leq i \leq n$ ), KGC does the following:

(a) computes

$$K_i = s^{-1}T_i = \{K_{ix}, K_{iy}\}.$$

(b) decrypts  $U_i$ 's message with  $K_{ix}$  and records the decrypted message  $\{ID_i, TID_i, R_i, Q_i\}$ .

(c) chooses  $t'_i \in_R Z_p^*$  and calculates:

$$\begin{cases} T'_i &= t'_i(R_i + Q_i + H_1(ID_i \| R_i \| Q_i)P_{pub}) \\ K'_i &= t'_iP = (K'_{ix}, K'_{iy}) \end{cases}$$

(d) KGC waits until all the group members complete Step 1, and sends the following message to every member  $U_i$  ( $1 \leq i \leq n$ ).

$$KGC \rightarrow U_i : T'_i, \{ID_1, TID_1, R_1, Q_1 \| \cdots \| ID_n, TID_n, R_n, Q_n\}_{K'_{ix}}$$

3. After receiving KGC's message in Step 2,  $U_i$  does the following:

(a) computes  $K'_i = (x_i + s_i)^{-1}T'_i = (K'_{ix}, K'_{iy})$ .

(b) decrypts KGC's message in Step 2d with  $K'_{ix}$ , keeps the decrypted message.

(c) chooses  $w_i \in_R Z_p^*$  and computes  $W_i = w_iP$ .

(d) sends to its two neighbors  $U_{i+1}$  and  $U_{i-1}$  the following messages:

$$U_i \rightarrow U_{i+1}, U_{i-1} : TID_i, W_i$$

$U_i$  also receives the message from  $U_{i-1}$  and  $U_{i+1}$ .  $U_i$  searches  $(ID_{i+1}, R_{i+1}, Q_{i+1})$  and  $(ID_{i-1}, R_{i-1}, Q_{i-1})$  with  $TID_{i+1}$  and  $TID_{i-1}$  respectively from KGC's message. Then  $U_i$  computes the following:

(a)  $K_{i,i+1} = (x_i + s_i)W_{i+1} + w_i(Q_{i+1} + R_{i+1} + H_1(ID_{i+1} \| R_{i+1} \| Q_{i+1})P_{pub})$

(b)  $K'_{i,i+1} = w_iW_{i+1}$

(c)  $K_i^R = H_2(K_{i,i+1}, K'_{i,i+1})$

(d)  $K_{i,i-1} = (x_i + s_i)W_{i-1} + w_i(Q_{i-1} + R_{i-1} + H_1(ID_{i-1} \| R_{i-1} \| Q_{i-1})P_{pub})$

(e)  $K'_{i,i-1} = w_iW_{i-1}$

(f)  $K_i^L = H_2(K_{i,i-1}, K'_{i,i-1})$

(g)  $X_i = K_i^L \oplus K_i^R$

$U_i$  broadcasts  $X_i$  to all the group:

$$U_i \rightarrow * : X_i$$

After getting  $X_j$  from all  $U_j$  ( $j \neq i$ ),  $U_i$  first checks whether

$$X_1 \oplus X_2 \cdots \oplus X_i \cdots \oplus X_n = 0$$

holds.  $U_i$  quits the protocol if it doesn't, or else  $U_i$  can compute  $K_1^R, K_2^R, K_3^R, \dots, K_n^R$  by chaining XOR with  $K_i^R$  as follows:

$$\begin{cases} K_{i+1}^R &= X_{i+1} \oplus K_i^R \\ K_{i+2}^R &= X_{i+2} \oplus K_{i+1}^R \\ \dots &= \\ K_n^R &= X_n \oplus K_{n-1}^R \\ K_1^R &= X_1 \oplus K_n^R \\ \dots &= \\ K_{i-1}^R &= X_{i-1} \oplus K_{i-2}^R \end{cases}$$



The group session key is computed as

$$K = H_2(K_1^R \| K_2^R \| \dots \| K_n^R)$$

The improved scheme is illustrated in Fig. 1.

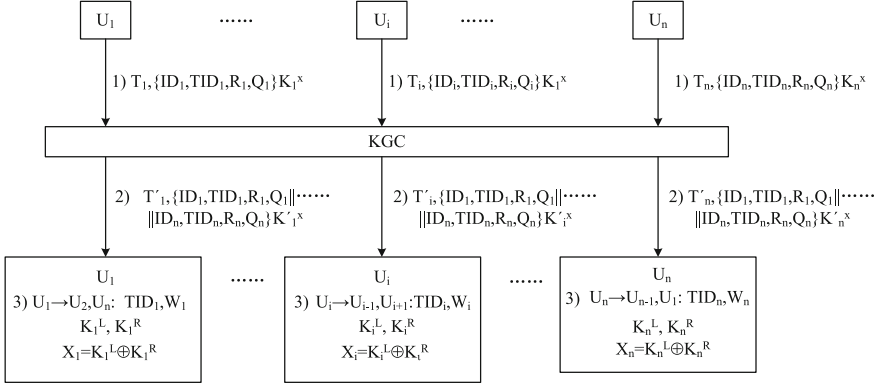


Fig. 1. Illustration of the improved scheme.

## 5 Security and Efficiency Analysis

### 5.1 Security Analysis

Now we consider the security strength of the improved scheme.

- S1 **Authenticity:** The improved scheme provides mutual authentication between KGC and user  $U_i$  ( $1 \leq i \leq n$ ).  $U_i$  could authenticate KGC because the message from  $U_i$  to KGC is encrypted by  $K_i$ . Only KGC could compute  $K_i = s^{-1}T_i$ . To compute  $K_i$  from  $T_i$  with  $P_{pub}$  and  $T_i$  is as hard as to solve the DCDH problem. KGC can authenticate  $U_i$  because the message to  $U_i$  is encrypted by the key  $K'_i$ , which could only be computed by  $U_i$  with its private key  $\langle x_i, s_i \rangle$ . Except for KGC and  $U_i$  the attacker who can compute  $K'_i$  should be able to solve the DCDH problem.  $U_i$  has to decrypt KGC's message correctly in order to proceed in the protocol. Thus KGC could authenticate the user. The group key is generated by the user's private key, every user in the system can authenticate the others. The system provides implicit key authentication.
- S2 **Group Forward/Backward Security:** The *join* and *leave* phases of Kumar-Tripathi scheme are secure, so these requirements are satisfied.
- S3 **Forward Secrecy:** In the improved scheme,  $K_i^R = H_2(K_{i,i+1}, K'_{i,i+1})$  with  $K'_{i,i+1} = w_i w_{i+1} P$ . Even an attacker can obtain all the users' private key pairs  $\langle x_i, s_i \rangle$  ( $1 \leq i \leq n$ ), it can not compute  $K'_{i,i+1} = w_i w_{i+1} P$  because without knowing  $w_i$  or  $w_{i+1}$ , the computation of  $K'_{i,i+1}$  is as difficult as to solve the computational Diffie-Hellman problem.

- S4 **Perfect Forward Secrecy:** The scheme also satisfies the perfect forward secrecy. The analysis is the same as above.
- S5 **Anonymity:** In the improved scheme, a user's identifier is transferred encrypted. A user's pseudonym which alters with the protocol is transferred instead of the the user's identifier. The attacker is unable to get the user's identity.
- S6 **Unlinkability:** In the improved scheme, a user's pseudonyms is in use and it changes from session to session. Furthermore, no same information appears in two different executions of the protocol. The outsiders are prevented from tracing the user using the protocol information.

The scheme is proved formally using the automatic tool AVISPA [22]. The authentication between a KGC and a user of the improved scheme is analyzed, stimulating two users in the system. Four states in all are analyzed and the authentication is proven safe. The role specification of the KGC and a user is shown in Figs. 2 and 3, respectively. The result is shown in Fig. 4.

```

role kgc( kc, U1, U2: agent, P, s, ID1.Q1, ID2.Q2 : message, K, K1, K2 : public_key
M, M1, H1, H2, A, A1 : Hash_func, SND, RCV : channel(dy))
played_by kc
def = local
state : nat
Ppub, s1, s2, R1, R2, Rr1, Rr2, Sig1, Sig2 : message

%knowledge(kc) = {inv(k)}
init
state := 0
transition
1. state = 0 ^ RCV(ID1.Q1) ^ RCV(ID2.Q2)=|>
state':=1 ^ Rr1' := new() ^ Rr2' := new()
^ R1' := M(P, Rr1') ^ R2' := M(P, Rr2')
^ s1' := A(Rr1, M1(s, H(ID1.Q1))) ^ s2' := A(Rr2, M1(s, H(ID2.Q2)))
^ Ppub' :=M(P, s)
^ Sig1' := H2(A(R1', M(Ppub', H(ID1.Q1)))) ^ Sig2' := H2(A(R2', M(Ppub', H(ID2.Q2))))
^ SND({{R1', s1', Ppub', Sig1'}_inv(K)}_K1) ^ SND({{R2', S2', Ppub', Sig2'}_inv(K)}_K2)
^witness(kc, U1, u1_kgc_Rr1, R1')
^witness(kc, U1, u1_kgc_s1, s1')
^witness(kc, U1, u1_kgc_ppub, Ppub')
^witness(kc, U2, u2_kgc_Rr2, R2')
^witness(kc, U2, u2_kgc_s2, s2')
^witness(kc, U2, u2_kgc_ppub, Ppub')
end role

```

**Fig. 2.** The role specification of a KGC

## 5.2 Performance Comparison

We compare the security of the improved scheme with available schemes. Sun *et al.*'s protocol [16] is chosen because it is the SOTA research result. Yang *et al.*'s scheme [18], Wan *et al.*'s scheme [17] as well as Kumar *et al.*'s scheme are chosen because these schemes aim to provide anonymity in AGKA. The comparison results are listed in Table 1. In the table, Column 1 to Column 6 correspond to the six security requirements listed in the Sect. 5.1.

```

role user1( kc, U1, U2 : agent, Ppub, K1, K2, Q1, Q2 : public_key, ID1.Q1, ID2.Q2, P :
message, x1, x2 : secret_key, M, M1, H2, A, A1 : hash_func, SND,REV : channel(dy))

played_by U1
def = local
state : nat

T1, s1, R1, Tt1, L1, Ll1, Sig1, SK1, T2, s2, R2, Tt2, L2, Ll2, Sig2 , SK2 : message

%knowledge(U1) = {inv(K1)}
init
state := 0

transition
1. state = 0  $\wedge$  RCV({{R1', s1', Ppub', Sig1'}_inv(Ppub)}_K1)  $\wedge$  RCV({{R2', s2', Ppub',
Sig2'}_inv(Ppub)}_K2)  $\wedge$  Sig1' := H2(A(R1', M(Ppub', H(ID1.Q1))))
 $\wedge$  Sig2' := H2(A(R2', M(Ppub', H(ID2.Q2)))) =|>

% equal(Sig1', M(P, s1))
% equal(Sig2', M(P, s2))
state' := 1
 $\wedge$  request(U1, kc, U1_kgc_r1, R1')
 $\wedge$  request(U1, kc, U1_kgc_s1, s1')
 $\wedge$  request(U1, kc, U1_kgc_ppub, Ppub')

 $\wedge$  request(U2, kc, U2_kgc_r2, R2')
 $\wedge$  request(U2, kc, U2_kgc_s2, S2')
 $\wedge$  request(U2, kc, U2_kgc_ppub, Ppub')

 $\wedge$  Tt1' := new()  $\wedge$  T1' := M(P, Tt1')  $\wedge$  Tt2' := new()  $\wedge$  T2' := M(P, Tt2')
 $\wedge$  SND(U1, T1', {{T1', R1', ID1.Q1}_inv(K1)}_SK)
 $\wedge$  witness(U1, kgc, kgc_u1_Tt1, T1')
 $\wedge$  witness(U1, U2, u2_u1_Tt1, T1')

2. state = 1  $\wedge$  RCV({{kgc, L1'}_inv(Ppub)}_sk1')  $\wedge$  RCV({{kgc, L2'}_inv(Ppub)}_SK2') =|>
state' := 2
 $\wedge$  request(U1, kgc, u1_kgc_Ll1, L1')
 $\wedge$  request(U2, kgc, U2_kgc_Ll2, L2')
SK1' := M(L1', M1(A1(s1', x1';)))
SK2' := M(L2', M1(A1(s2', x2';)))

end role

```

**Fig. 3.** The role specification of a user

From the table, we can see that Sun *et al.*'s protocol doesn't aim to provide anonymity thus it only satisfies the security requirement of secure group key establishment. Sun *et al.*'s scheme doesn't provide unlinkability. Wan *et al.*'s scheme and our scheme satisfy all the six security requirements.

The computational overhead of the improved scheme is compared with the same set of protocols. These protocols are based on either pairing-based cryptosystem or elliptic curve cryptosystem, so we consider the operations including the bilinear pairings, scalar multiplications in pairing-related group  $\mathbb{G}_1$ , hash to  $\mathbb{G}_1$  point, scalar multiplications in elliptic curve group  $G$ , hash to point in  $G$ , and asymmetric operations including signature/verification and enc/decryption. Other cryptographic operations such as hash function, symmetric enc/decryption are omitted because the time for these are neglectable according to the benchmark result in [23]. The notations used in our analysis are listed in Table 2.

```

SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
  TYPED_MODEL
PROTOCOL
  program files/avispa/clagka/./tempdir/clagkasec.if
GOAL
  As Specified
BACKENED
  CL-Atse
STATISTICS
  Analysed : 4 states
  Reachable : 0 states
  Translation : 0.60 seconds
  Computation : 0.00 seconds

```

**Fig. 4.** The role specification of a user**Table 1.** Security comparison

	S1	S2	S3	S4	S5	S6
Sun et al.'s	Yes	Yes	Yes	Yes	No	No
Wan et al.'s	Yes	Yes	Yes	Yes	Yes	Yes
Yang et al.'s	Yes	Yes	Yes	Yes	Yes	No
Kumar et al.'s	No	Yes	Yes	Yes	No	No
Ours	Yes	Yes	Yes	Yes	Yes	Yes

**Table 2.** Notations

Notation	Description
$\tau_{pa}$	Time for bilinear pairing
$\tau_{\mathbb{G}_1}$	Time for scalar multiplication in paring-based group $\mathbb{G}_1$
$\tau_{H_{\mathbb{G}_1}}$	Time to hash a string to a point in paring-based group $\mathbb{G}_1$
$\tau_G$	Time for scalar multiplication in elliptic curve group $G$
$\tau_{H_G}$	Time to hash a string to a point in elliptic curve group $G$

We compare the computation overhead of key agreement phase. The computation cost of a group manager and that of a group member are discussed separately. We also compare the serial cost which is the overall cost of a group manager and a group member, and which can be viewed as the cost to generate a group key [1].

- According to the analysis in [16], in Sun *et al.*'s scheme, the computation cost of a group manager is  $n\tau_{pa} + 5n\tau_{\mathbb{G}_1}$ . The computation cost for a group member is  $\tau_{pa} + 3\tau_{\mathbb{G}_1}$  with precomputation. Thus the serial cost is  $(n+1)\tau_{pa} + (5n+3)\tau_{\mathbb{G}_1}$ .
- In Wan *et al.*'s scheme [17], a group manager should compute one scalar multiplication on  $\mathbb{G}_1$ , a signature and an anonymous encryption to provide the anonymity for the group members. We employ Choon-Cheon's ID-based signature [24] because it is efficient for Wan *et al.*'s scheme. The ID-based anonymous encryption in [25] is adopted. With the two schemes, the overall overhead of a group manager is  $(2n+5)\tau_{\mathbb{G}_1} + n\tau_{H_{\mathbb{G}_1}} + \tau_{pa}$ . A group member should execute 2  $\mathbb{G}_1$  hash, 3  $\mathbb{G}_1$  scalar multiplications, and 2 pairings to compute the group session key. Moreover, a member should also verify the group manager's signature and decrypt the manager's encryption to check if he is the designated group member. A group member's computation cost is  $(2n+3)\tau_{\mathbb{G}_1} + 2\tau_{H_{\mathbb{G}_1}} + 6\tau_{pa}$ . The serial cost is  $(4n+8)\tau_{\mathbb{G}_1} + (n+2)\tau_{H_{\mathbb{G}_1}} + 7\tau_{pa}$ .
- Yang *et al.*'s protocol realizes user anonymity using the pairing-free certificateless cryptography. The scheme requires a group member 6 scalar multiplications on  $G$ . A group manager has to execute  $7n+1$  scalar multiplications on  $G$ . Thus the serial cost is  $(7n+7)\tau_G$ .
- In the key-agreement algorithm of Kumar-Tripathi's scheme, KGC acts as a group manager. A KGC should carry out one  $G$  scalar multiplication for each group member to compute the symmetric session key, a symmetric encryption and a decryption. Since the symmetric enc/decryption is omitted, the overall computation overhead for a group manager is  $n\tau_G$  for  $n$  group members. Each group member should execute 12 scalar multiplications on  $G$ . The serial cost is  $(n+12)\tau_G$ .
- Our improved scheme requires the KGC one scalar multiplication to compute  $T'_i$  and three scalar multiplications for  $K'_i$  in the Step 2, thus the overall cost for a KGC is  $4n\tau_G$ . A group member  $U_i$  has to execute two scalars to compute the symmetric key  $T_i$  and  $K_i$  respectively in the Step 1. In the Step 3,  $U_i$  needs one scalar to compute  $K'_i$  from  $T'_i$ , one scalar multiplication to compute  $W_i$ , and eight scalar multiplications altogether to compute  $K_i^R$  and  $K_i^L$ . Therefore, the overall computation cost for  $U_i$  is  $12\tau_G$ . The serial cost for the establishment of a group key is  $(4n+12)\tau_G$ .

We compare the serial cost of different schemes in Table 3. According to experimental results in [26], the running time of scalar multiplications on  $\mathbb{G}_1$  and bilinear pairings are higher than that of scalar multiplications on  $G$ . A scalar multiplication on  $G$  could be twenty times as fast as a pairing. It can be seen that the proposed scheme has better performance than Sun *et al.*'s scheme because our scheme only needs  $(4n+12)$   $G$  scalar multiplications, and Sun *et al.*'s scheme needs  $5n+3$   $\mathbb{G}_1$  scalar multiplications and  $n+1$  pairings. To compute  $n+1$  pairings in Sun *et al.*'s scheme could be slower than our scheme. Wan *et al.*'s scheme needs  $(4n+8)$   $\mathbb{G}_1$  scalar multiplications and more than 4 pairings, thus the overall computation cost is higher than our scheme. The serial cost of Yang *et al.*'s scheme is  $(7n+7)\tau_G$ ,  $(7n+7)\tau_G > (4n+12)\tau_G$  when  $n > 5/3$ .

**Table 3.** Performance comparison

Scheme	Group manager's cost	Group member's cost	Serial cost
Sun et al.'s	$5n\tau_{G_1} + n\tau_{pa}$	$3\tau_{G_1} + \tau_{pa}$	$(5n + 3)\tau_{G_1} + (n + 1)\tau_{pa}$
Wan et al.'s	$(2n + 5)\tau_{G_1} + n\tau_{H_{G_1}} + \tau_{pa}$	$(2n + 3)\tau_{G_1} + 2\tau_{H_{G_1}} + 6\tau_{pa}$	$(4n + 8)\tau_{G_1} + (n + 2)\tau_{H_{G_1}} + 7\tau_{pa}$
Yang et al.'s	$(7n + 1)\tau_G$	$6\tau_G$	$(7n + 7)\tau_G$
Kumar et al.'s	$n\tau_G$	$12\tau_G$	$(n + 12)\tau_G$
Our scheme	$4n\tau_G$	$12\tau_G$	$(4n + 12)\tau_G$

Therefore, when there is more than one user in the group, our protocol is more efficient than Yang *et al.*'s scheme, and this is always true. Kumar *et al.*'s scheme is the most efficient among all, but it is proven insecure. The improved scheme requires  $3n$  more  $G$  scalar multiplications than Kumar-Tripathi's scheme, and these overhead are used for a KGC to authenticate group members.

According to the analysis above, our scheme achieves strengthened security with improved performance compared with available schemes.

## 6 Conclusion

In this research, we propose attacks to Kumar-Tripathi's anonymous CL-AGKA protocol. We prove that the protocol fails to provide master key secrecy, authenticity and unlinkability. Then the protocol is improved based on computational Diffie-Hellman problem and divisible computational Diffie-Hellman problem. Compared with current schemes, the improved scheme provides strengthened security with the best computational performance.

**Acknowledgments.** The authors would like to thank the Fundamental Research Funds for the Central Universities (JB161508), National Natural Science Foundation of China (No. 61402351), and China 111 Project (B16037) for support.

## References

1. Kim, Y., Perrig, A., Tsudik, G.: Tree-based group key agreement. *ACM Trans. Inform. Syst. Secur.* **7**(1), 60–96 (2004)
2. Boyd, C., Nieto, J.M.G.: Round-optimal contributory conference key agreement. In: Desmedt, Y.G. (ed.) *PKC 2003*. LNCS, vol. 2567, pp. 161–174. Springer, Heidelberg (2003). [https://doi.org/10.1007/3-540-36288-6\\_12](https://doi.org/10.1007/3-540-36288-6_12)
3. 3GPP TS 33.102: 3rd Generation Partnership Project 3GPP, 3G Security, Security Architecture, Technical Specification Group (TSG) SA (2003)
4. Buttner, C., Huss, S.A.: A novel anonymous authenticated key agreement protocol for vehicular ad hoc networks. In: *Proceedings of International Conference on Information Systems Security*, pp. 259–269 (2015)

5. Al-Riyami, S.S., Paterson, K.G.: Certificateless public key cryptography. In: Laih, C.-S. (ed.) ASIACRYPT 2003. LNCS, vol. 2894, pp. 452–473. Springer, Heidelberg (2003). [https://doi.org/10.1007/978-3-540-40061-5\\_29](https://doi.org/10.1007/978-3-540-40061-5_29)
6. Shamir, A.: Identity-based cryptosystems and signature schemes. In: Blakley, G.R., Chaum, D. (eds.) CRYPTO 1984. LNCS, vol. 196, pp. 47–53. Springer, Heidelberg (1985). [https://doi.org/10.1007/3-540-39568-7\\_5](https://doi.org/10.1007/3-540-39568-7_5)
7. Zhang, L., Zhang, F., Wu, Q., Domingoferrer, J.: Simulatable certificateless two-party authenticated key agreement protocol. *Inf. Sci.* **180**(6), 1020–1030 (2010)
8. Lippold, G., Boyd, C., Nieto, J.M.G.: Strongly secure certificateless key agreement. In: Proceedings of 3rd International Conference on Paring Cryptography (Pairing 2009), pp. 206–230 (2009)
9. Yang, G., Tan, C.-H.: Strongly secure certificateless key exchange without pairing. In: Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security, pp. 71–79. ACM (2011)
10. Teng, J., Wu, C.: A provable authenticated certificateless group key agreement with constant rounds. *J. Commun. Netw.* **14**(1), 104–110 (2012)
11. Lu, C., Wu, T., Hsu, C.L.: Certificateless authenticated group key agreement scheme with privacy-preservation for resource-limited mobile devices. *Int. J. Innov. Comput. Inf. Control* **8**(1B), 599–615 (2012)
12. Seo, S.H., Won, J., Sultana, S., Bertino, E.: Effective key management in dynamic wireless sensor networks. *IEEE Trans. Inf. Forensics Secur.* **10**(2), 371–383 (2015)
13. Gu, X., Cao, Z., Wang, Y.: How to get group key efficiently in mobile ad hoc networks. In: Proceedings of Military Communications Conference (MILCOM 2015), pp. 1009–1014 (2015)
14. Heo, S., Kim, Z., Kim, K.: Certificateless authenticated group key agreement protocol for dynamic groups, pp. 464–468 (2007)
15. Lee, E.J., Lee, S.E., Yoo, K.Y.: A certificateless authenticated group key agreement protocol providing forward secrecy. In: Proceedings of International Symposium on Ubiquitous Multimedia Computing, pp. 124–129 (2008)
16. Sun, H., He, B., Chen, C., Wu, T., Lin, C., Wang, H.: A provable authenticated group key agreement protocol for mobile environment. *Inf. Sci.* **321**, 224–237 (2015)
17. Wan, Z., Ren, K., Lou, W., Preneel, B.: Anonymous id-based group key agreement for wireless networks. In: Wireless Communications and Networking Conference, pp. 2615–2620. IEEE (2008)
18. Yang, Y., Zheng, X., Liu, X., Zhong, S., Chang, V.: Cross-domain dynamic anonymous authenticated group key management with symptom-matching for e-health social system. In: Future Generation Computer Systems (2017, in press). <http://www.sciencedirect.com/science/article/pii/S0167739X1730554X>
19. Kumar, A., Tripathi, S.: A pairing free anonymous certificateless group key agreement protocol for dynamic group. *Wirel. Pers. Commun.* **82**(2), 1027–1045 (2015)
20. Xiong, H.: Cost-effective scalable and anonymous certificateless remote authentication protocol. *IEEE Trans. Inf. Forensics Secur.* **9**(12), 2327–2339 (2014)
21. Liu, J., Zhang, Z., Chen, X., Kwak, K.S.: Certificateless remote anonymous authentication schemes for wirelessbody area networks. *IEEE Trans. Parallel Distrib. Syst.* **25**(2), 332–342 (2014)
22. Armando, A., et al.: The AVISPA tool for the automated validation of internet security protocols and applications. In: Etessami, K., Rajamani, S.K. (eds.) CAV 2005. LNCS, vol. 3576, pp. 281–285. Springer, Heidelberg (2005). [https://doi.org/10.1007/11513988\\_27](https://doi.org/10.1007/11513988_27)
23. Dai, W.: “Crypto++ 5.6.5 benchmarks”. <https://www.cryptopp.com/benchmarks.html>

24. Choon, J.C., Hee Cheon, J.: An identity-based signature from gap diffie-hellman groups. In: Desmedt, Y.G. (ed.) PKC 2003. LNCS, vol. 2567, pp. 18–30. Springer, Heidelberg (2003). [https://doi.org/10.1007/3-540-36288-6\\_2](https://doi.org/10.1007/3-540-36288-6_2)
25. Wang, H., Zhang, Y., Xiong, H., Qin, B.: Cryptanalysis and improvements of an anonymous multi-receiver identity-based encryption scheme. IET Inf. Secur. **6**(1), 20–27 (2012)
26. “Miracl library: <https://www.miracl.com/>. Shamus Software Ltd



# A Space Efficient Algorithm for LCIS Problem

Daxin Zhu<sup>1</sup> and Xiaodong Wang<sup>2</sup>(✉)

<sup>1</sup> Quanzhou Normal University, Quanzhou 362000, China

<sup>2</sup> Fujian University of Technology, Fuzhou 350108, China  
wangxd135@139.com

**Abstract.** This paper reformulates the problem of finding a longest common increasing subsequence of the two given input sequences in a very succinct way. An extremely simple linear space algorithm based on the new formula can find a longest common increasing subsequence of sizes  $n$  and  $m$  respectively, in time  $O(nm)$  using additional  $\min\{n, m\} + 1$  space.

**Keywords:** Algorithm · LCIS · Dynamic programming  
Time complexity · Space complexity · Linear space

## 1 Introduction

A sequence  $A = (a_1, \dots, a_n)$  of numbers is an increasing sequence if  $a_1 < a_2 < \dots < a_n$ . Given 2 sequences input sequences  $A = (a_1, \dots, a_n)$  and  $B = (b_1, \dots, b_m)$ , a sequence  $S$  is a common increasing subsequence (CIS) of the 2 sequences if  $S$  is an increasing sequence and  $S$  is a subsequence of both  $A$  and  $B$ . The longest common increasing subsequence (LCIS) of  $A$  and  $B$  is the longest sequence among all CIS of  $A$  and  $B$ .

Yang et al. [11] designed a dynamic programming algorithm that finds an LCIS of two input sequences of size  $n$  and  $m$  in  $O(nm)$  time and space. If  $r$ , the total number of ordered pairs of positions at which the two input sequences match, is relatively small, Chan et al. [6] gave a faster algorithm which runs in  $O(R \log l \log \log n + \text{Sort}(n))$  time, where  $n$  is the length of each sequence and  $R$  is the total number of ordered pairs of positions at which the two sequences match and  $l$  is the length of the LCIS, and  $\text{Sort}(n)$  is the time to sort each input sequence. If the length of the LCIS,  $l$ , is small, Kutz et al. [9] gave a faster algorithm, with an output-dependent expected running time of  $O((m + nl) \log \log \sigma + \text{Sort}(n))$ , where  $\sigma$  is the size of the alphabe. A first linear space algorithm was proposed by Sakai [3, 10].

If the length of the LCIS,  $l$ , is small, Kutz and Brodal [9] gave a faster algorithm which runs in  $O(nl \log n)$  time. If  $r$ , the total number of ordered pairs of positions at which the two input sequences match, is relatively small, Chan et al. [6] gave a faster algorithm which runs in  $O(\min(r \log l, nl + r) \log \log n + n \log n)$  time where  $n$  is the length of each sequence and  $r$  is the total number of ordered pairs of positions at which the two sequences match and  $l$  is the length of

the LCIS. A first linear space algorithm was proposed by Sakai. The space cost of the algorithm of Yang et al. was reduced to linear by a careful application of Hirschbergs divide-and-conquer method [7,8]. The space complexity of the algorithm of Kutz and Brodal [9] was also reduced from  $O(nl)$  to  $O(m)$  by using the same divide-and-conquer method of Hirschberg.

In this paper, we revisit the KBKK algorithm for computing LCIS. Based on a novel property of the LCIS, we find a very simple linear space algorithm but not the Hirschbergs divide-and-conquer method to solve the problem.

## 2 Definitions and Terminologies

In the whole paper we will use  $A = (a_1, \dots, a_n)$  and  $B = (b_1, \dots, b_m)$  to denote the two input sequences of size  $n$  and  $m$  respectively, where each pair of elements in the sequences is comparable.

A data structure called a bounded heap (BH) is used in the KBKK algorithm [1,2,4,5]. It supports the following operations:

- *Insert* ( $H, s, p, d$ ): Insert into the BH  $H$  the key  $s$  with priority  $p$  and associated data  $d$ .
- *DecreasePriority* ( $H, s, p, d$ ): If the BH  $H$  does not yet contain the key  $s$ , perform *Insert*( $H, s, p, d$ ). Otherwise, set this keys priority to  $\min\{p, p'\}$ , where  $p'$  is its previous priority.
- *BoundedMin* ( $H, s$ ): Return the item that has minimum priority among all items in  $H$  with key smaller than  $s$ . If  $H$  does not contain any items with key smaller than  $s$ , return “invalid”.

The above operations of a bounded heap can be supported in  $O(\log \log n)$  amortized time using  $O(n)$  space, where keys are drawn from the set  $\{1, \dots, n\}$ .

The KBKK algorithm begins with a preprocessing step, where it removes from each sequence all elements that do not appear in the other sequence and rename the remaining symbols of the alphabet to  $\{1, 2, \dots, \sigma\}$ . For every remaining element  $s$ , it generates a sorted list  $Occ_s$  that contains the indices of all occurrences of  $s$  in  $B$ . Then, the algorithm identifies common increasing subsequences (CISs) of increasing lengths. In iteration  $i$  it identifies length- $i$  CISs. For each element  $a_j$ , the minimum index  $k$  in  $B$  is identified such that there is a length- $i$  CIS which ends at  $a_j$  and  $b_k$ . The index  $k$  is stored in  $L_i[j]$ .

For each  $1 \leq j \leq n$ ,  $L_1[j]$  can be simply set to the minimum index in the list  $Occ_{a_j}$ , i.e., the earliest occurrence of  $a_j$  in  $B$ . In the  $i$ th iteration, the algorithm checks for each  $a_j$  whether it can extend a length- $(i - 1)$  CIS to a length- $i$  CIS, and if so, identifies the minimum such  $k$ . For this purpose, the algorithm maintains a bounded heap  $H$ . When  $a_j$  is processed, it contains all elements  $a_t, 1 \leq t < j$ , for which  $L_{i-1}[t] < \infty$ . The key of  $a_t$  in  $H$  is  $a_t$  itself and its priority is  $L_{i-1}[t]$ . The algorithm performs the query *BoundedMin* ( $H, a_j$ ) to find the leftmost endpoint of a length- $(i - 1)$  CIS, which contains only elements smaller than  $a_j$ . If  $k$  is this endpoint, then  $L_i[j]$  is set to the first occurrence of  $a_j$  in  $B$  which lies behind  $k$ . This is the leftmost endpoint in  $B$  of a length- $i$  CIS

which ends at  $a_j$ . If  $A$  and  $B$  have a length- $l$  LCIS which ends in  $A$  at index  $j$  and in  $B$  at index  $k$ , then at the end of the algorithm we have:  $i = l$  and  $L_i[j] \leq k$ .

For each iteration of the algorithm,  $O(n)$  operations are performed on the bounded heap  $H$ , each of which takes  $O(\log \log \sigma)$  amortized time. The  $Occ_s$  lists are queried at most  $n$  times. A naive implementation of the  $Occ_s$  lists would require  $\theta(\sigma m)$  time and space, but this complexity can be reduced to  $O(m)$  time, and space by a randomized implementation of the algorithm. Finally, the total expected running time of the algorithm is  $O((m + nl) \log \log \sigma + Sort_\Sigma(m))$ , where  $Sort_\Sigma(m)$  is the time required to sort a length- $m$  input sequence drawn from the alphabet  $\Sigma$ .

To construct the LCIS, the array  $Link$  are used in the algorithm to save the necessary information for backtrack. The array  $Link$  requires  $O(nl)$  space in the worst case. In the KBKK algorithm, the technique developed by Hirschberg [3] for LCS is used to reduce the space complexity from  $O(nl)$  to  $O(m)$ .

For each pair  $(i, j)$ , if  $a_i = b_j$  then  $(i, j)$  is called a match of  $A$  and  $B$ , and  $a_i$  (or  $b_j$ ) is called the match value of  $(i, j)$ . For any two matches  $(i, j)$  and  $(i', j')$ , if  $i' < i$ ,  $j' < j$ , and  $a_{i'} < a_i$ , then  $(i', j')$  is defined to dominate  $(i, j)$ , or  $(i', j')$  be a dominating match of  $(i, j)$ . For every match  $(i, j)$ , define the rank of the match, denoted by  $r(i, j)$ , to be the length of the LCIS of  $A_i = (a_1, \dots, a_i)$  and  $B_j = (b_1, \dots, b_j)$  such that  $a_i$  (and  $b_j$ ) is the last element of the LCIS. It is easy to see that the length of the LCIS of  $A$  and  $B$  is  $l = \max\{r(i, j)\}$ . A critical match in the match set is defined to be a match  $(i, j)$  such that no match  $(i', j')$  of the same set with  $i' \leq i$  and  $j' \leq j$ .

Let  $M^k$  be the set of critical matches with rank  $k$ . In the algorithm of Chan et al., one vEB tree  $T^k$  is used to store the set of critical matches in  $M^k$  for  $1 \leq k \leq l$ , respectively. Then all matches between  $A$  and  $B$  are identified. The matches are arranged in ascending order of their values. The matches  $(i, j)$  with the same value are arranged in descending order of  $j$  and then  $i$ . For each of the matches in this order, the largest rank dominating match can be found by a binary search on the non-empty vEB trees. To determine if there is a dominating match of  $(i, j)$  in  $T^k$ , the operation  $FindLeft(k, (i, j))$  is applied to  $T^k$ , which finds a match  $(i', j')$  in  $T^k$  with the maximum  $j'$  that  $j' < j$ . If a match  $(i', j')$  with  $i' < i$  and  $j' < j$  is returned, then a dominating match of  $(i, j)$  in  $T^k$  is found, and then the rank  $r(i, j)$  can be determined by  $r(i, j) = 1 + r(i', j')$ . Once the rank  $r(i, j)$  is computed, every matches  $(i', j')$  in  $T^{r(i, j)}$  such that  $i \leq i'$  and  $j \leq j'$  will become non-critical. These non-critical matches are removed from  $T^{r(i, j)}$  and the new match  $(i, j)$  becomes a new critical match in  $T^{r(i, j)}$ . At the end of the algorithm,  $l = \max\{r(i, j)\}$ , the length of the LCIS of  $A$  and  $B$  is obtained.

The algorithm computes actually only the length of the LCIS of  $A$  and  $B$ . If an LCIS of  $A$  and  $B$  must be computed, the algorithm may cost more space. In the algorithm, when the rank  $r(i, j)$  is determined by  $r(i, j) = 1 + r(i', j')$ , the match  $(i', j')$  is called a previous match of  $(i, j)$ . If the previous match of each match  $(i, j)$  is recorded in the algorithm, then an LCIS of  $A$  and  $B$  can be

easily constructed from any match in  $T^l$  and the recorded previous matches. To record all previous matches requires extra  $O(R)$  space.

The runtime of the algorithm of Chan et al. consists of three parts.

- (1) It takes  $O(Sort(n) + R)$  time to sort the two sequences  $A$  and  $B$  and identify the matches between  $A$  and  $B$  in ascending order of their values, where  $R$  is the number of matches of  $A$  and  $B$ ;
- (2) It takes  $O(\log l \log \log n)$  time to find the largest rank dominating match of a given match in the binary search because the function  $FindLeft$  is called  $O(\log l)$  times and there are at most  $l$  non-empty vEB trees. Hence it takes  $O(R \log l \log \log n)$  time to find the ranks of all matches;
- (3) Each match can be inserted and deleted at most once from a vEB tree. Each of these operations takes  $O(\log \log n)$  time, and hence it takes  $O(R \log \log n)$  time for all matches. Altogether, it takes  $O(R \log l \log \log n + Sort(n))$  time.

### 3 A Simple Linear Space Algorithm

The space complexity is not given in [6]. Let each vEB tree  $T^k$  use space  $t_k$ ,  $1 \leq k \leq l$ , in the algorithm. If only the length of the LCIS of  $A$  and  $B$  needs to be computed, then it is obvious that the algorithm of Chan et al. requires  $O(\sum_{k=1}^l t_k)$  space. If an LCIS of  $A$  and  $B$  needs to be computed, not just its length, then the algorithm of Chan et al. will require  $O(R + \sum_{k=1}^l t_k)$  space. It would be quadric since  $R$  can be as large as  $O(n^2)$  in the worst case. In this note, we will prove first that  $O(\sum_{k=1}^l t_k) = O(n)$ . Then, a simple linear space algorithm to compute the LCIS of  $A$  and  $B$  based on the algorithm of Chan et al. is presented.

The vEB trees  $T^k$ ,  $1 \leq k \leq l$ , constructed by the algorithm of Chan et al. has a very nice property as stated in the following lemma.

**Lemma 1.** *At any time of the algorithm, for each match  $(i, j) \in T^k$ ,  $1 < i \leq n$ ,  $1 < j \leq n$ ,  $1 < k \leq l$ , there must be a match  $(i', j') \in T^{k-1}$  such that  $(i', j')$  dominates  $(i, j)$ .*

*Proof.* Let

$$\begin{cases} k = \min_{1 < t \leq j} \{t \mid L_i[t] < \infty, a_t \leq a_j\} \\ r = L_i[k] \end{cases} \tag{1}$$

It follows from the above that  $1 \leq r < k \leq j$ ,  $L_{i-1}[r] < \infty$  and  $a_r < a_k \leq a_j$ . If  $L_i[r] < \infty$ , it follows from (1) that  $k \leq r$ . A contradiction. Therefore,  $L_i[r] = \infty$ .

The proof is complete.

Based on Lemma 1, the KBKK algorithm can be improved to a very simple linear space algorithm as follows.

---

**Algorithm 1.** LCIS

---

**Input:**  $A, B$   
**Output:** LCIS of  $A$  and  $B$   
 $i \leftarrow 1$ ;  
**for**  $j=1$  **to**  $n$  **do**  
   $L[j] \leftarrow \text{MinimumKey}(\text{Occ}_{a_j})$   
**end**  
**while**  $i < n$  **and**  $L[j] < \infty$  **for some**  $j$  **do**  
   $H \leftarrow \emptyset$ ;  $i \leftarrow i + 1$ ;  
  **for**  $j=1$  **to**  $n$  **do**  
     $t \leftarrow L[j]$ ;  $L[j] \leftarrow \infty$ ;  
     $(j', k') \leftarrow \text{BoundedMin}(H, a_j)$ ;  
    **if**  $(j', k') \neq \text{invalid}$  **then**  
       $L[j] \leftarrow \min\{k : k \in \text{Occ}_{a_j} \wedge k > k'\}$ ;  
      **if**  $L[j] < \infty$  **then**  $d[j] \leftarrow i$   
    **end**  
    **if**  $t < \infty$  **then**  $\text{DecreasePriority}(H, a_j, t, (j, t))$ ;  
  **end**  
**end**

---

In the above algorithm, the array  $L$  is used to store the values of  $L_i$  in the KBKK algorithm, since in the main loop of the algorithm only  $L_{i-1}$  and  $L_i$  are used.  $t$  is used to hold the value of  $L_{i-1}[j]$  to be used for the next  $j$ .  $H$  is a bounded heap. When  $a_j$  is processed, it contains all elements  $a_t, 1 \leq t < j$ , for which  $L[t] < \infty$ . The key of  $a_t$  in  $H$  is  $a_t$  itself and its priority is  $L[t]$ . The array  $d$  is used to record the current length of the CIS which ends in  $A$  at index  $j$ . The array  $d$  can be used to build an LCIS of  $A$  and  $B$  as follows.

---

**Algorithm 2.** Build LCIS

---

**Input:**  $d$   
**Output:** An LCIS of  $A$  and  $B$   
 $i \leftarrow \max_{1 \leq j \leq n} d[j]$ ;  $t \leftarrow \infty$ ;  
**for**  $j=n$  **downto**  $1$  **do**  
  **if**  $d[j] = i$  **and**  $a_j < t$  **then** **Print**  $a_j$ ;  $t \leftarrow a_j$ ;  $i \leftarrow i - 1$ ;  
**end**

---

The algorithm searches for an LCIS in a reversed order. The variable  $i$  is the current length of the LCIS to come up, and  $t$  is the current element of the LCIS at position  $i$ . If  $i > 1$  and  $t = a_j$ , we can always find an index  $1 \leq k < j$  such that  $d[k] = i - 1$  and  $a_k < a_j$  by Lemma 1. Therefore, the algorithm can always stop when  $i = 1$ , and outputs an LCIS of  $A$  and  $B$  in a reverse order. The time complexity of the algorithm does not change. The space required to backtrack an LCIS is clearly  $O(n)$ .

The table  $L$  constructed by the KBKK algorithm has a very nice property as stated in the following lemma.

**Lemma 2.** *For each pair  $(i, j), 1 < i \leq l, 1 < j \leq n$ , if  $L_i[j] < \infty$ , then there must be an index  $r$  such that*

$$\begin{cases} 1 \leq r < j, \\ a_r < a_j, \\ L_{i-1}[r] < \infty \\ L_i[r] = \infty. \end{cases} \quad (2)$$

*Proof.* It follows from the above that  $1 \leq r < k \leq j, L_{i-1}[r] < \infty$  and  $a_r < a_k \leq a_j$ . If  $L_i[r] < \infty$ , it follows from (1) that  $k \leq r$ . A contradiction. Therefore,  $L_i[r] = \infty$ .

The proof is complete.

Based on the formula (2), we can reduce the space cost of the algorithm *LCIS* to  $\min\{n, m\} + 1$ . When computing a particular row of the dynamic programming table, no rows before the previous row are required. Only two rows have to be kept in memory at a time, and thus we need only  $\min\{n, m\} + 1$  entries to compute the table. A space efficient algorithm to compute the length of LCIS of  $X$  and  $Y$  can be described as follows.

---

**Algorithm 3.** Linear Space

---

**Input:**  $X, Y$

**Output:** The length of LCIS of  $X$  and  $Y$

**for**  $i = 1$  **to**  $n$  **do**

$L(0) \leftarrow 0;$

**for**  $j = 1$  **to**  $m$  **do**

**if**  $x_i > y_j$  **and**  $L(j) > L(0)$  **then**  $L(0) \leftarrow L(j);$

**if**  $x_i = y_j$  **and**  $L(j) < L(0) + 1$  **then**  $L(j) \leftarrow L(0) + 1;$

**end**

**end**

$L(0) \leftarrow \max_{1 \leq j \leq m} L(j);$

**return**  $L(0)$

---

In the above algorithm, the array  $L$  is used to store the values of  $L_i$  in the KBKK algorithm, since in the main loop of the algorithm only  $L_{i-1}$  and  $L_i$  are used.  $t$  is used to hold the value of  $L_{i-1}[j]$  to be used for the next  $j$ .  $H$  is a bounded heap. When  $a_j$  is processed, it contains all elements  $a_t, 1 \leq t < j$ , for which  $L[t] < \infty$ . The key of  $a_t$  in  $H$  is  $a_t$  itself and its priority is  $L[t]$ . The array  $d$  is used to record the current length of the CIS which ends in  $A$  at index  $j$ . The array  $d$  can be used to build an LCIS of  $A$  and  $B$  as follows.

---

**Algorithm 4.**  $h(i_0, i_1, j_0, j_1, l, u, L, I)$ 


---

```

for  $i = i_1$  downto  $i_0$  do
  if  $x_i > l$  and  $x_i < u$  then
     $L(0) \leftarrow 0;$ 
    for  $j = j_1$  downto  $j_0$  do
      if  $x_i < y_j$  and  $L(j) > L(0)$  then  $L(0) \leftarrow L(j);$ 
      if  $x_i = y_j$  and  $L(j) < L(0) + 1$  then  $L(j) \leftarrow L(0) + 1; I(j) \leftarrow i;$ 
    end
  end
end

```

---

The algorithm searches for an LCIS in a reversed order. The variable  $i$  is the current length of the LCIS to come up, and  $t$  is the current element of the LCIS at position  $i$ . If  $i > 1$  and  $t = a_j$ , we can always find an index  $1 \leq k < j$  such that  $d[k] = i - 1$  and  $a_k < a_j$  by Lemma 1. Therefore, the algorithm can always stop when  $i = 1$ , and outputs an LCIS of  $A$  and  $B$  in a reverse order. The time complexity of the algorithm does not change. The space required to backtrack an LCIS is clearly  $O(n)$ .

Finally, our main result can be completed in the following theorem.

**Theorem 1.** *Let  $X = x_1x_2 \cdots x_n$  and  $Y = y_1y_2 \cdots y_m$  be two input sequences over an alphabet  $\Sigma$  of size  $n$  and  $m$  respectively. A longest common increasing subsequences of  $X$  and  $Y$  can be computed in time  $O(nm)$  using additional  $\min\{n, m\} + 1$  space.*

**Example.** For the given input sequences of  $X = (3, 5, 1, 2, 7, 5, 7)$  and  $Y = (3, 5, 2, 1, 5, 7)$ ,  $f(i, j)$ , the length of an LCIS in  $LCIS(X_i, Y_j)$  is listed below.

$$f = \begin{pmatrix} 1, 0, 0, 0, 0, 0 \\ 1, 2, 0, 0, 2, 0 \\ 1, 2, 0, 1, 2, 0 \\ 1, 2, 1, 1, 2, 0 \\ 1, 2, 1, 1, 2, 3 \\ 1, 2, 1, 1, 2, 3 \\ 1, 2, 1, 1, 2, 3 \end{pmatrix}$$

It follows from the above that the length of any LCIS of  $X$  and  $Y$  is 3. The LCIS (1, 5, 7) of  $X$  and  $Y$  can be generated by the algorithm.

## 4 Concluding Remarks

We have found a simple method to build an LCIS of the KBKK algorithm. This makes to find an LCIS is as easy as to find the length of LCIS. The time complexity does not change. The question whether the time complexity can be reduced to  $O(n^{2-\varepsilon})$  is still open.

**Acknowledgment.** This work was supported in part by the Quanzhou Foundation of Science and Technology under Grant No.2013Z38, Fujian Provincial Key Laboratory of Data-Intensive Computing and Fujian University Laboratory of Intelligent Computing and Information Processing.

## References

1. Abboud, A., Backurs, A., Williams, V.V.: Quadratic-time hardness of LCS and other sequence similarity measures. In: Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2015), pp. 59–78 (2015)
2. Abboud, A., Hansen, T.D., Williams, V.V., Williams, R.: Simulating branching programs with edit distance and friends or: a polylog shaved is a lower bound made. In: Proceedings of the 48th Annual ACM Symposium on Symposium on Theory of Computing (STOC 2016), pp. 375–388 (2016)
3. Backurs, A., Indyk, P.: Edit distance cannot be computed in strongly subquadratic time (unless SETH is false). In: Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC 2015), pp. 51–58 (2015)
4. Backurs, A., Indyk, P.: Which regular expression patterns are hard to match? In: Proceedings of the 57th Annual Symposium on Foundations of Computer Science, (FOCS 2016), pp. 457–466 (2016)
5. Backurs, A., Tzamos, C.: Improving viterbi is hard: better runtimes imply faster clique algorithms. In: Proceedings of the 34th International Conference on Machine Learning (ICML 2017), pp. 311–321 (2017)
6. Chan, W., Zhang, Y., Fung, S.P.Y., Ye, D., Zhu, H.: Efficient algorithms for finding a longest common increasing subsequence. *J. Comb. Optim.* **13**, 277–288 (2007)
7. Cygan, M., Mucha, M., Wegrzycki, K., Włodarczyk, M.: On problems equivalent to  $(\min, +)$ -convolution. In: Proceedings of the 44th International Colloquium on Automata, Languages, and Programming (ICALP 2017), pp. 22:1–15 (2017)
8. Hirschberg, D.S.: A linear space algorithm for computing maximal common subsequences. *Commun. ACM* **18**(6), 341–343 (1975)
9. Kutz, M., Brodal, G.S., Kaligosi, K., Katriel, I.: Faster algorithms for computing longest common increasing subsequences. *J. Discrete Algorithms* **9**, 314–325 (2011)
10. Sakai, Y.: A linear space algorithm for computing a longest common increasing subsequence. *Inf. Process. Lett.* **99**, 203–207 (2006)
11. Yang, I.H., Huang, C.P., Chao, K.M.: A fast algorithm for computing a longest common increasing subsequence. *Inf. Process. Lett.* **93**(5), 249–253 (2005)



# Improving the Efficiency of Dynamic Programming in Big Data Computing

Xiaodong Wang<sup>1</sup> and Daxin Zhu<sup>2</sup>(✉)

<sup>1</sup> Fujian University of Technology, Fuzhou 350108, China

<sup>2</sup> Quanzhou Normal University, Quanzhou 362000, China  
dex@qztc.edu.cn

**Abstract.** In this paper we present the extremely simple algorithms to improve the performance of dynamic programming, one of the fundamental techniques for solving optimization problems, in the environment of data-intensive computing. These algorithms are applied to several NP hard combinatorial optimization problems. The presented algorithms decrease the time and space complexity of dynamic programming algorithms by exploiting word parallelism. The computational experiments demonstrate that the achieved results are not only of theoretical interest, but also that the techniques developed may actually lead to considerably faster algorithms.

**Keywords:** Algorithm · Big data · Dynamic programming  
Time complexity · Space complexity · Computational experiments

## 1 Introduction

An optimization problem exhibits the property of optimal substructure if an optimal solution to the problem consists of a sequence of optimal solutions to subproblems. Such problems may be solved through dynamic programming as this framework takes advantage of overlapping subproblems to decrease the computational effort.

In this paper, we consider the following 4 *NP*-hard problems.

(1) **Subset-Sum Problem:**

Given a set  $S = \{a_1, a_2, \dots, a_n\}$  of  $n$  nonnegative integers and an additional integer  $m$ , the subset-sum problem asks to choose a subset  $\bar{S}$  of  $S$  such that  $\sum_{\bar{S}} a_i$  is maximized without exceeding  $m$ . The integers  $a_i \leq m$  are frequently denoted the weights and  $m$  the capacity.

(2) **Bounded Subset-Sum Problem:**

The bounded subset-sum problem is a generalization of the subset-sum problem where each weight may be chosen a bounded number of times.

Given a set  $S = \{(a_1, d_1), (a_2, d_2), \dots, (a_n, d_n)\}$  of nonnegative integer pairs, and an additional integer  $m$ , the problem asks to maximize the sum  $\sum_S a_i x_i$  without exceeding  $m$ , respecting that  $x_i \in \{0, 1, \dots, d_i\}$ . The integers  $a_i$  denoted the weights,  $d_i$  denoted the bounds and  $m$  denoted the capacity.

**(3) Unbounded Subset-Sum Problem:**

Given a set  $S = \{a_1, a_2, \dots, a_n\}$  of  $n$  nonnegative integers and an additional integer  $m$ , the unbounded subset-sum problem asks to maximize the sum  $\sum_S a_i x_i$  without exceeding  $m$ , where  $x_i \in \{0, 1, \dots\}$ .

**(4) Two-Partition Problem:**

The two-partition problem can be formulated as follows: Two sets of integers  $S = \{a_1, a_2, \dots, a_n\}$  and  $T = \{a'_1, a'_2, \dots, a'_n\}$  are given together with an  $m$ . Find two subsets  $\bar{S}$  of  $S$  and  $\bar{T}$  of  $T$  such that  $\sum_{\bar{S}} a_i = \sum_{\bar{T}} a'_i$  is maximized without exceeding  $m$ .

The 4 problems described above are *NP*-hard, but they can be solved through dynamic programming in time  $O(nm)$  through a straightforward recursion by Bellman [1].

A problem exhibits the property of optimal substructure can be solved through dynamic programming as this framework takes advantage of overlapping subproblems to decrease the computational effort [4].

Applying dynamic programming to *NP*-hard problems may lead to algorithms with pseudo polynomial running time. The dynamic programming algorithms generally have the additional benefit that we do not only obtain a single solution but a whole table of optimal sub-solutions corresponding to different values of the constraints. Solving the problem for all values of the constraints will change the problems considered to polynomial problems with respect to the input and output size.

Algorithms exploiting word-parallelism have been studied in the last years within the field of algorithms dealing with polynomial problems.

In the computation model of word RAM [2, 5, 10], the largest coefficient in the input and output data is assumed to be  $U$  and the word size of the computer to be  $w$ .

When deriving the time and space complexity of the presented algorithms we assume that  $w = \Theta(\log U)$ .

In a word RAM, the operations like binary and, binary or, and bitwise shift with  $w$  bits can be implemented in constant time on words of size  $w$ . None of the developed algorithms make use of multiplication apart from the process of indexing multidimensional tables. The indexing can however be implemented by a single shift operation per index if the domain size of the index is extended to the nearest larger power of two. For an excellent survey on sorting and searching algorithms on the word RAM see [3].

The algorithms in the computation model of word RAM have improved the time complexity of several problems and given a new insight into the design of algorithms. Unfortunately the improved time complexities seldom lead to practical improvements of the performance, as they are beaten by algorithms with theoretically worse time complexity but where the constants in the complexity term are much smaller.

The present paper exploits word parallelism for the 4 *NP*-hard problems described above by using the bitset class. As the computational effort for solving these problems is huge, word parallelism is of great importance both from a theoretical and practical aspect.

The organization of the paper is as follows.

In the following 3 sections we describe our new algorithm design paradigm. In Sect. 2 we give an extremely simple algorithm for the 4 *NP*-hard problems described above with the bitset class, where each bit indicates whether a subset-sum can be achieved with a sum which equals the bit position. In the new algorithm we exploit word parallelism to decrease the running time of the dynamic programming algorithm to  $O(nm/\log m)$ , where  $n$  is the number of integers given and  $m$  is the sum capacity.

In Sect. 3 we give a computational study of the presented algorithm which demonstrates that the achieved results are not only of theoretical interest, but also that the techniques developed may actually lead to faster algorithms.

Some concluding remarks are in Sect. 4.

## 2 The Algorithms Using Bitset Class

For the subset-sum problem, a straightforward dynamic programming algorithm can be designed as follows.

Let  $t_{i,j}$  for  $i = 0, \dots, n, j = 0, \dots, m$  be a solution to the subset-sum problem defined on items  $S_i = \{a_1, \dots, a_i\}$  with  $m = j$ .

To initialize the recursion we set  $t_{0,j} = 0$  for  $j = 0, \dots, m$ .

Subsequent values of  $t$  can be computed recursively as

$$t_{i,j} = \max\{t_{i-1,j}, t_{i-1,j-a_i} + a_i\} \quad (1)$$

for  $i = 1, \dots, n, j = 0, \dots, m$ .

The optimal value of the problem is given by  $t_{n,m}$ .

Since the table  $t_{i,j}$  has size  $O(nm)$  and each entry can be computed in constant time, the time and space complexity of the dynamic programming algorithm for the subset-sum problem is  $O(nm)$ .

Notice that the dynamic programming algorithm for the subset-sum problem not only solves the problem for the given capacity  $m$  but for all capacities  $0, \dots, m$ . This variant of the problem appears in the industry where the capacity is not known exactly in advance but where the choice depends on the solutions achievable [12].

Other approaches for the subset-sum problem include branch-and-bound [7, 9] and balanced dynamic programming [8]. The first approaches run in exponential time while the second has the time and space complexity  $O(n \max_{a_i \in S} a_i)$ . Although these approaches are frequently able to solve the subset-sum problem faster than dynamic programming algorithm, they only return the solution value for a single capacity value.

The dynamic programming algorithm for subset-sum can be adapted for exploiting word parallelism.

Let  $b_{i,j}$  denote the indicator of subset-sums. That is  $b_{i,j} = 1$  if and only if there exists a subset of the weights  $S_i = \{a_1, a_2, \dots, a_i\}$  which sums to  $j$ .

In this way,  $b_{0,j} = 0$  for  $j = 0, \dots, m$  and the subsequent values of  $b$  can be computed recursively as

$$b_{i,j} = b_{i-1,j} \quad \text{or} \quad b_{i-1,j-a_i} \tag{2}$$

for  $i = 1, \dots, n, j = 0, \dots, m$ .

According to the formula 2, we can design an extremely simple algorithm for the subset-sum problem with the bitset class as follows.

---

**Algorithm 1.** Subet-Sum( $a$ )

---

```

b[0] ← 1
for  $i = 1$  to  $n$  do
     $b \leftarrow b$  or  $b \ll a_i$ 
end for
return  $b$ 
    
```

---

In the algorithm stated above, the parameter  $a$  is a vector represents the input set  $S$ . The result returned by the algorithm Subet-Sum is a bitset  $b$  representing the indicator of the subset-sums.

**Theorem 1.** *The optimal values of the subset-sum problem can be computed in  $O(nm/\log m)$  time and  $O(m/\log m)$  space using algorithm Subet-Sum ( $a, b$ ).*

**Proof.** Since only one bit is needed for representing each entry in  $b_{i,j}$  we can use a bit set to store the indicator of subset-sums. This bit set  $b$  takes up  $O(m/\log m)$  space.

To derive the bit set for  $b_{i,*}$  from the current bitset for  $b_{i-1,*}$  we simply copy current bitset  $b$  to a new bitset  $b$  **or**  $b \ll a_i$  which is the bitset  $b$  shifted left by  $a_i$  bits. Then we merge the two bitset  $b$  and  $b \ll a_i$  through a bitset operation **or**.

The whole operation for the “for” loop of the algorithm can be done in  $O(m/\log m)$  time as the bitset operation **or** of two words can be done in constant time and each word can be shifted in constant time.

The optimal value  $z$  is found as the largest value  $j \leq m$  for which  $b[j] = 1$ . This can easily be done in  $O(m/\log m)$  time by a linear search to find the rightmost bit. The whole operation takes  $O(nm/\log m)$  time. This proves the time complexity.

If the optimal value of the subset-sum problem is  $z$ , the corresponding subset  $S'$  of  $S$  satisfies  $\sum_{S'} a_i = z$  can be computed by the following algorithm Construct( $a, b, z$ ), where  $a$  is a vector representing the input set  $S$  and  $b$  a bitset computed by the algorithm Subet-Sum( $a$ ).

---

**Algorithm 2.** Construct( $a, b, z$ )

---

```

for  $i = n$  down to  $0$  do
   $b \leftarrow b \gg a_i$ 
  if  $b[z] = 0$  then
    output  $a_i$ 
     $z \leftarrow z - a_i$ 
  end if
end for

```

---

The whole operation for the “for” loop of the algorithm Construct( $a, b, z$ ) can be done in  $O(m/\log m)$  time as the bitset operation or of two words can be done in constant time and each word can be shifted in constant time. The time complexity of the algorithm Construct( $a, b, z$ ) is therefore  $O(nm/\log m)$ .

We now consider the problem (2), the bounded subset-sum problem.

Since every integer number can be uniquely represented as a partial sum of the sequence of powers of 2. Each of the powers of 2 appears at most once in such a sum, i.e. its selection is a binary decision. Therefore we can model the variable  $x_i$  in the bounded subset-sum problem by binary variables each of them equal to a power of 2 times a single copy of his item.

More formally, using binary coding we can convert the bounded subset-sum problem to an ordinary subset-sum problem with  $\sum_{i \in S} \log d_i$  items.

We introduce  $n_i = \lfloor \log d_i \rfloor$  artificial items  $a_i, 2a_i, \dots, 2^{n_i}a_i$  for every integer  $a_i$  in the set  $S$  such that the artificial items are multiples of  $a_i$  by a power of 2.

The optimal solution of the resulting instance of the ordinary subset-sum problem is equivalent to the optimal solution of the original bounded subset-sum problem since every possible value of  $x_i \in \{0, 1, \dots, d_i\}$  can be represented by a sum of binary variables described above.

Since  $d_i \leq m$  the number of items is  $O(n \log m)$ .

The bounded subset-sum problem can then be solved by the following algorithm Bounded-Subet-Sum( $a, d$ ).

---

**Algorithm 3.** Bounded-Subet-Sum( $a, d$ )

---

```

 $b[0] \leftarrow 1$ 
for  $i = 1$  to  $n$  do
  while  $d_i \neq 0$  and  $a_i \neq 0$  do
     $b \leftarrow b \text{ or } b \ll a_i$ 
     $a_i = \lfloor a_i/2 \rfloor$ 
     $d_i = \lfloor d_i/2 \rfloor$ 
  end while
end for
return  $b$ 

```

---

The total running time of the algorithm Bounded-Subet-Sum( $a, d$ ) is obviously  $O(n \log m \cdot m/\log m) = O(nm)$ . The space complexity of the algorithm is

$O(m/\log m)$ . This improves the complexity of the dynamic programming algorithm which has time complexity  $O(m \sum_{i=1}^n d_i)$  and space complexity  $O(nm)$ .

The unbounded subset-sum problem can be solved as the bounded subset-sum problem by imposing an upper bound  $\lfloor m/a_i \rfloor$  on the number of times the integer  $a_i$  can be chosen.

---

**Algorithm 4.** Unbounded-Subet-Sum( $a, d$ )

---

```

b[0] ← 1
for  $i = 1$  to  $n$  do
   $d = \lfloor m/a_i \rfloor$ 
  while  $d \neq 0$  and  $a_i \neq 0$  do
     $b \leftarrow b \text{ or } b \ll a_i$ 
     $a_i = \lfloor a_i/2 \rfloor$ 
     $d = \lfloor d/2 \rfloor$ 
  end while
end for
return  $b$ 

```

---

The time and space complexities of the algorithm Unbounded-Subet-Sum( $a, d$ ) are the same as the algorithm Bounded-Subet-Sum( $a, d$ ).

To solve the problem (4) the two-partition problem, we first use the algorithm Subet-Sum for the input set  $S$  and  $T$ . This gives us two bitsets representing the solutions of the subset-sum problem for the input set  $S$  and  $T$  respectively. Then, use binary and to add these bitsets obtaining a new bitset  $b$ . In this bitset, if  $b[j] = 1$  then it is possible to obtain  $\sum_S a_i = \sum_T a'_i$ .

The algorithm can be described as follows.

---

**Algorithm 5.** Two-Partition( $c, d$ )

---

```

 $b \leftarrow$  Subet-Sum( $c$ )
 $b \leftarrow b \ \&$  Subet-Sum( $d$ )
return  $b$ 

```

---

For the optimal solution, we simply use a linear search starting from  $m$  and running downward to find the largest value of  $j$  for which  $b[j] = 1$ .

The time and space complexities of the algorithm Two-Partition( $c, d$ ) are obviously  $O((n_S + n_T)m/\log m)$  and  $O(m/\log m)$ .

### 3 Computational Experiments

In this section, we give some computational experiments on the performance of our new algorithm for the subset-sum problem exploiting word parallelism. We

have restricted the experiments to the variant of subset-sum problems where the output of the algorithm should be a complete list of optimal solution values for all capacities up to  $m$ .

To test the performance of the algorithms, we have chosen five types of data instances from Martello and Toth [7]:

P3  $a_i$  uniformly random distributed in  $[1, 10^3]$  and  $\lfloor b = n10^3/4 \rfloor$

P6  $a_i$  uniformly random distributed in  $[1, 10^6]$  and  $\lfloor b = n10^6/4 \rfloor$

Evenodd  $a_i$  even, uniformly random distributed in  $[1, 10^3]$  and  $b = 2\lfloor n10^3/8 \rfloor + 1$

Avis  $a_i = n(n+1) + i$  and  $b = n(n+1)\lfloor (n-1)/2 \rfloor + n(n-1)/2$

Todd  $a_i = 2^{\lfloor \log_2 n \rfloor + n + 1} + 2^{\lfloor \log_2 n \rfloor + i} + 1$   
and  $b = \lfloor \frac{1}{2} \sum_{i=1}^n a_i \rfloor$

It is known that every branch-and-bound algorithm enumerates an exponentially growing number of nodes when solving Evenodd problems [6]. Any recursive algorithm will perform poorly for the Avis problems [3]. Any algorithm which uses upper bounding tests will have to enumerate an exponential number of states for the Todd problems [3].

We have used the bitset class of STL [11] to test our algorithm. A drawback of the bitset class of STL is the size of the bitset must be fixed before compilation. To compile the program using the bitset class of STL, it is needed to pass in a fixed integer value that specifies the number of bits that a bitset can hold. This fixed size constraint can help make bitset operations fast at the expense of programming flexibility.

The solution to the dynamic size declaration might be found by using the boost dynamic-bitset [12, 13] to store the bitsets. This will give us more natural access to operations on the bitset as a whole.

Another way to use the bitset class dynamically is to write our own bitset class using a bit vector. It is also not a difficult task.

Our computational experiments were carried out on a personal computer with Pentium (R) Dual Core CPU 2.10 GHz and 2.0 Gb RAM. The word size of the processor is  $w = 32$ .

As expected, the bitset class exploiting word parallelism leads to a considerably more efficient algorithm which in general is at least one order of magnitude faster than the dynamic programming algorithm. This speedup may seem surprising as the word size is  $w = 32$ , but due to the decreased space complexity caching gets improved, which results in additional speedup. For large problems, which do not fit within the cache, the speedup decreases.

## 4 Concluding Remarks

This paper has introduced the extremely simple algorithms for 4 *NP*-hard subset-sum like problems with the bitset class. The presented algorithms decrease the time and space complexity of dynamic programming algorithms by exploiting word parallelism.

The computational experiments in Sect. 3 demonstrate that the achieved results are not only of theoretical interest, but also that the techniques developed may actually lead to considerably faster algorithms.

As dynamic programming algorithms in general are very space consuming, it is possible to solve larger problems by exploiting word parallelism. Reducing the space complexity also improves the caching of the processor, thus leading to an additional decrease in the observed computational time.

Throughout the paper we assume that  $w = \Theta(\log U)$ , i.e. that the word size is of the same magnitude as the coefficients in the input and output data. In practice one will normally use the largest possible word size supported by the processor, thus obtaining an additional decrease in computational time. Modern processors support 128 bit integer arithmetics, and thus word parallelism may be of significant importance.

**Acknowledgment.** This work was supported in part by the Quanzhou Foundation of Science and Technology under Grant No. 2013Z38, Fujian Provincial Key Laboratory of Data-Intensive Computing and Fujian University Laboratory of Intelligent Computing and Information Processing.

## References

1. Baker, J., Bond, C., Corbett, J.C., et al.: Megastore: providing scalable, highly available storage for interactive services. In: 5th Biennial Conference on Innovative Data Systems Research, CIDR 2011 (2011)
2. Diehl, M.: Database Replicatio with Mysql. Linux Journal, May 2010
3. Dodis, Y., Patrascu, M., Thorup, M.: Changing base without losing space. In: Proceeding of 42nd ACM Symposium on Theory of Computing (STOC) (2010)
4. Frenkel, E., Nikolaev, A., Ushakov, A.: Knapsack problems in products of groups. *J. Symbolic Comput.* **74**, 96–108 (2016)
5. Goerigk, M., Gupta, M., Ide, J., Schobel, A., Sen, S.: The robust knapsack problem with queries. *Comput. Oper. Res.* **55**, 12–22 (2015)
6. Hans, K., Pferschy, U., Pisinger, D.: Knapsack Problems. Springer Verlag, Berlin (2005). <https://doi.org/10.1007/978-3-540-24777-7>
7. He, Y., Zhang, X., Li, W., Li, X., Wu, W., Gao, S.: Algorithms for randomized time-varying knapsack problems. *J. Comb. Optim.* **31**, 95–117 (2016)
8. Kong, X., Gao, L., Ouyang, H., Li, S.: Solving large-scale multidimensional knapsack problems with a new binary harmony search algorithm. *Comput. Oper. Res.* **63**, 7–22 (2015)
9. Krishnamoorthy, B.: Thinner is not always better: cascade knapsack problems. *Oper. Res. Lett.* **45**, 77–83 (2017)
10. Schulze, B., Paquete, L., Klamroth, K., Figueira, J.: Bi-dimensional knapsack problems with one soft constraint. *Comput. Oper. Res.* **78**, 15–26 (2017)



11. Ssulami, A.M., Mathkour, H.: Faster string matching based on hashing and bit-parallelism. *Inf. Process. Lett.* **123**, 51–55 (2017)
12. Swamy, C.: Improved approximation algorithms for matroid and knapsack median problems and applications. *ACM Trans. Algorithms* **49**, 1–49 (2016)
13. Viola, E.: Bit-probe lower bounds for succinct data structures. In: *Proceeding of 41st ACM Symposium on Theory of Computing (STOC)*, pp. 475–482 (2009)

# Traceable and Complete Fine-Grained Revocable Multi-authority Attribute-Based Encryption Scheme in Social Network

Yanmei Li, Fang Qi, and Zhe Tang<sup>(✉)</sup>

School of Information Science and Engineering, Central South University,  
Changsha 410083, China  
{liyanmei\_, csqifang, tz}@csu.edu.cn

**Abstract.** Nowadays the data and user information involved in the social network with the nature of high complexity. More and more service providers and users will share the data in the cloud servers. To keep the shared data confidential against untrusted third-party service providers, settle single point failure and solve performance bottlenecks of authorized center as well as secret key abuse of malicious users who disclose their own private key, we propose a multi-authority attribute-based encryption scheme which supports traceability and fine-grained revocation mechanism in social network. The scheme is based on traditional attribute based encryption, which can realize distributed access control and support complete fine-grained revocation mechanism. The security is proved in the standard model, which effectively solves the above problems.

**Keywords:** Attribute-based encryption · Revocable · Traceable  
Fine-grained · Social network

## 1 Introduction

In recent years, the attribute-based encryption-ABE [1] mechanism has become a hotspot of cryptography. This scheme associates the user's private key and ciphertext with different attribute sets, respectively. When and only when the user's private key can satisfy the access policy in the ciphertext, the ciphertext can be decrypted, which can flexibly display the access policy and reduce the network bandwidth and computation cost of network node. Therefore, fine-grained access control can be widely used. In ABE scheme, since the user's private key is generated from attribute authority, the private key may be leaked by attribute authority or users and it is hard to judge who should be responsible for the leakage. Thus the problem of private key abuse is prominent. Attribute revocation is a useful way to solve above problem and also is a matter of research. According to the way of attribute revocation, there are two types: direct revocation and indirect revocation. Peng and Zhang [6] pointed out that, in indirect revocation,

only the users who are not revoked can update the secret key. The first proposed ABE scheme only supports threshold access control policy. To express a more flexible access control policy, researchers proposed two kinds ABE schemes: one is key-policy attribute-based encryption (KP-ABE) [2, 3], in which user's private key associates with the access policy; the other is ciphertext-policy attribute-based encryption (CP-ABE) [4, 5], in which the access policy is embedded in the ciphertext. Considering that the CP-ABE can resist collusion attack, guarantee the confidentiality and realize flexible access control. In this paper, we propose a traceable and complete fine-grained revocable multi-authority attribute-based encryption scheme. The contributions are as follows:

- Proposed scheme can reduce the risk of single point failure and bottleneck.
- The ciphertext can contain multiple revocation lists and the complete fine-grained control can be realized.
- We proposed the traceable algorithm to locate the malicious user, which can guarantee the security.

## 2 Proposed Scheme

In this section, we will introduce the traceable and complete fine-grained revocable multi-authority attribute-based encryption scheme in detail. The scheme contains the following eight algorithms:

**GlobalInit.** Suppose  $G, G_T$  are cyclic group with order  $N = P_1 P_2 P_3$ , where  $P_1, P_2, P_3$  are distinct big prime numbers,  $e$  is the bilinear mapping  $e : G \times G = G_T$ . Suppose  $G_{P_1}$  is the subgroup of  $G$  with order  $P_1$ .  $g$  is the generator of  $G_{P_1}$ .  $G_{P_3}$  is the subgroup of  $G$  with order  $P_3$ .  $Y$  is the generator of  $G_{P_3}$ . We can randomly select a subgroup  $h$  is  $G_{P_1}$  and two parameters  $a, \alpha$  from  $Z_N$ . Input  $1^\tau$  where  $\tau$  is the security parameter, this algorithm outputs the system public key:

$$GPAR = (N, g, h, Y, e(g, g)^a, g^a) \quad (1)$$

**CASetup.** This algorithm runs in  $d$ th authority  $CA_d$  input the system public key  $GPAR$  and the index  $d$  of corresponding authority  $CA_d$ .  $CA_d$  randomly chooses  $a_d, \alpha_d \in Z_N$  and runs the algorithm. Finally, the algorithm returns  $CA'_d$ 's system public key and system master key as follows:

$$CAPK_d = e(g, g)^{\alpha_d} \quad (2)$$

$$CAMSK_d = (a_d, \alpha_d) \quad (3)$$

$CA_d$  publishes the public key and keeps the private key, the revoke list  $T_d$  is initialized as an empty list.

**AASetup.** On input the system public key  $GPAR$ , the attribute domain  $U_k$  managed by attribute authority  $AA_k$  and the index  $k$ , suppose  $att \in U_k$ , randomly select  $S_{att} \in_R Z_N$  and suppose  $H_{att} = g^{S_{att}}$ . This algorithm outputs the public key:

$$AAPK_k = (H_{att}) \quad (4)$$

of  $AA_k$  private key:

$$AASK_k = (S_{att}) \quad (5)$$

and publishes the public key and keeps the private key.

**Encrypt.** Input the plaintext  $M$ , access structure  $T = (A, \rho)$ , system public key  $GPAR$ , the public key  $CAPK_d$  of  $CA_d$ , public key  $AAPK_k$  of  $AA_k$ . Finally the algorithm outputs the ciphertext  $CT$ . The access structure  $T$  is presented by a LSSS matrix  $(A, \rho)$ , where  $A$  is  $l \times n$  matrix.  $\rho$  will map every line  $A_x$  of matrix  $A$  into an attribute  $\rho(x)$ , where  $x \in (1, 2, \dots, n)$ .  $\rho$  is not allowed to map the different line into a same attribute. Then, to build up a  $n$ -dimension vector  $\vec{v} = (s, v_1, v_2, \dots, v_{n-1})$ , randomly select  $n - 1$  numbers  $v_1, v_2, \dots, v_{n-1}$  from  $Z_N$  and a secret  $s$ . Suppose  $\lambda_x = A_x \cdot \vec{v}$ .  $R(x)$  is the revoke list of users and  $S(x) = U - R(x)$  (suppose  $S(x) \neq \emptyset$ ). If  $S(x) \neq U$  and  $R(x) \neq \emptyset$ , randomly select  $\gamma_x \in Z_N$  and compute  $C_{x,1}, C_{x,2}$ , where  $C_{x,2}$  is the information of attribute revocation. The ciphertext  $CT$  is:

$$CT = \left\{ \begin{array}{l} C_0 = M \cdot e(g, g)^{s \cdot \alpha_d}, C_1 = g^{\alpha_d}, \\ C_2 = g^s, C_3 = g^{\lambda_x}, \\ C_{x,0} = H_{att}^{\lambda_x}, C_{x,1} = g^{\gamma_x}, \\ C_{x,2} = (g^{\gamma_x} \cdot \prod_{j \in S(x)} g_{n+1-j} \cdot g^d)^{\lambda_x}, \\ C_{x,3} = (\prod_{j \in R(x)} g_{n+1-j})^{\lambda_x} \cdot g^{\alpha_d} \end{array} \right\} x \in \{1, 2, \dots, l\} \quad (6)$$

**CAKeyGen.** By this algorithm, the user registers his/her identity information  $gid$  with the authorization center  $CA_d$  to obtain the their own secret key. When  $CA_d$  received the users identity  $gid$ , it randomly selects a number  $r_{gid,d} \in Z_N$  and  $r_{gid} = \prod_{d=1}^D r_{gid,d}$ , inputs system public key  $GPAR$  and the private key  $CAMSK_d$  of  $CA_d$ . This algorithm returns the users public key:

$$aucpk_{gid,d} = \{g^{\alpha_d \cdot g^{a \cdot r_{gid,d}}}\} \quad (7)$$

and users private key:

$$auck_{gid,d} = \{d, gid, L_{gid,d} = g^{r_{gid,d}}\} d \in \{1, 2, \dots, D\} \quad (8)$$

**AAKeyGen.** This algorithm is used to generate the users attribute private key. When attribute authority receives the private key generation request from attribute  $att (att \in U_k)$ , input the system public key  $\{aucpk_{gid,d}\}_{d \in D}$ . Attribute authority  $AA_k$  computes  $CA_d$ 's user's attribute private key:

$$auask_{att,gid,d} = (L_{gid,d})^{S_{att}} = g^{r_{gid,d} S_{att}} = (H_{att}^{gid,d}) \quad (9)$$

then combined with the user attribute private key of all authorization centers  $CA_d$ , we can get user's private key is:

$$auask_{att,gid} = \prod_{d=1}^D auask_{att,gid,d} = (L_{gid})^{S_{att}} \quad (10)$$

It is clear that the users identity is embedded in the private key.

**Decrypt.** If the visitor can satisfy the access structure  $T = (A_{l \times n}, \rho)$ , then there exists recovery parameter  $\omega_x \in Z_N$  which can satisfy  $\sum_{\rho(x) \in A} \omega_x \cdot A_x = (1, 0, \dots, 0)$ . If the visitor cannot satisfy the access structure  $T = (A_{l \times n}, \rho)$ , the algorithm return  $\perp$ , which means visitor cannot decrypt the ciphertext.  $D_x$  can be get from:

$$\frac{e(g^\alpha g^{gid} uask_{gid,d}, C_3) e(C_3, C_{x,3}) e\left(g, \left(\prod_{j \in S(x)} g_{n+1-j}\right)^{\lambda_x}\right) \cdot e(C_3, C_{x,1})}{e\left(L_{gid,d}, (H_{att})^{\lambda_s}\right) e(g, C_{x,2}) e\left(g, \left(\prod_{j \in R(x)} g_{n+1-j}\right)^{\lambda_x \gamma_x}\right)} \quad (11)$$

$$= e(g, g)^{(gid + \alpha_c) \lambda_x}$$

Suppose  $\omega_x$  is the recovery parameter of the  $x$ th line of matrix  $A$ . According to LSSS recovery algorithm, the plaintext  $M$  can be recovered:

$$C_0 \cdot \left(\prod_{x \in \ell} D_x^{\omega_x}\right)^{-1} \cdot e(g^{gid}, g^s) = M \quad (12)$$

**Trace.** This algorithm is run by authority, because the attribute and identity is bond to the secret key. The algorithm will generate the pirate decoder, and then the identity of malicious user can be traced according to  $T$ .

### 3 Security Analysis

In this section, we will analysis the security performance. First, we will build semi-function ciphertext and semi-function secret key. Semi-function ciphertext randomly select two numbers  $c, t \in Z_N$ , suppose  $R_x$  is the subgroup of  $G_{P_2}$ ,  $g_2$  is the generator of  $G_{P_2}$ . Randomly select a vector  $\vec{u} = (c, u_2, u_3, \dots, u_k)$  from  $Z_N^k$ . The semi-function ciphertext is CT. Semi-function secret key: there are two kinds semi-function secret key randomly select a number  $n \in Z_N$ . The first kind of semi-function secret key is:

$$\begin{aligned} auck'_{gid,d} &= \{d, gid, L_{gid,d} = g^{r_{gid,d}} g_2^{n + \vec{A}_x \cdot \vec{u}}\} \\ uask_{att,gid} &= \prod_{d=1}^D auask_{att,gid,d} = (L - gid^{S_{att}}) g_2^{n + \vec{A}_x \cdot \vec{u}} \end{aligned} \quad (13)$$

The second kind of semi-function secret key is:

$$\begin{aligned} auck'_{gid,d} &= \{d, gid, L_{gid,d} = g^{r_{gid,d}} g_2^n\} \\ uask_{att,gid} &= \prod_{d=1}^D auask_{att,gid,d} = (L - gid^{S_{att}}) g_2^{n + \vec{A}_x \cdot \vec{u}} \end{aligned} \quad (14)$$

For any  $1 \leq k \leq q$  we define a serial security game, where  $q$  is the maximum number that adversary queries ciphertext:

- $\text{Game}_{real}$ : The game is a real attack game defined earlier in this article.
- $\text{Game}_0$ : This game is like  $\text{Game}_{real}$ , the difference is that challenger returns semi-function ciphertext, so the game equals to  $\text{Game}_{0,2}$ .
- $\text{Game}_{final}$ : In the challenge phase of the security game, the challenger returns a semi-function ciphertext with random message to the adversary.

**Lemma 1.** *If there is a polynomial time algorithm adversary  $\Gamma$  that  $\text{Game}_{real}$  and  $\text{Game}_0$  can be distinguished by non-ignorable advantage  $\varepsilon$ , then we can construct a polynomial time algorithm  $\delta$  that can overcome the assumption 1 by advantage  $\varepsilon$ .*

*Proof.* The polynomial time algorithm receives a hypothetical 1 with condition  $(g, Y, T)$ .

Init: The adversary outputs an access structure  $T^* = (A_{l \times n}^*, \rho)$  that is challenged, for any  $gid \in GID$ , arbitrarily lists the specified user revocation list  $R_{\rho(x)}^*$ , and the algorithm  $\delta$  runs  $\Gamma$ .

Setup: Suppose  $\omega^* = \{\rho(x)\}$ , generate the system public key:

$$GPAR = (N, g, h, Y, e(g, g)^\alpha, g^a) \quad (15)$$

System private key  $CAMSK_d = (a_d, \alpha_d)$  and  $AASMK_k = (S_{att})$ .

Phase 1: When adversary asks about the private key rival, algorithm can use the system private key and secret key generation algorithm, then returns normal private key to adversary.

Challenge: Adversary submits two plain messages  $M_0, M_1$  with the same length-randomly select a vector  $\vec{v}' = (1, v'_1, v'_2, \dots, v'_n)$ , then choose a message  $M_\theta$  from  $M_0, M_1$ , the ciphertext is:

$$CT' = \left\{ \begin{array}{l} C_0' = M_\theta \cdot e(g, g)^{s \cdot \alpha_d}, C_1' = g^{a_d} g_2^c, \\ C_2' = g^s g_2^t, C_3' = g^{A_x \cdot \vec{v}'} g_2^{ct}, \\ C'_{x,0} = H_{att}^{A_x \cdot \vec{v}'}, C_{x,1}' = g^{\gamma_x}, \\ C_{x,2}' = \left( g^{\gamma_x} \prod_{j \in S(x)} g_{n+1-j} \cdot g^d \right)^{A_x \cdot \vec{v}'}, \\ C_{x,3}' = \left( \prod_{j \in R(x)} g_{n+1-j} \right)^{A_x \cdot \vec{v}'} \cdot g^{\alpha_d} \end{array} \right\}, x \in \{1, 2, \dots, l\} \quad (16)$$

Phase 2: Repeat Phase1s operation. Guess: Adversary outputs the guess  $\theta'$  of ciphertext  $M_\theta$ .

- If  $T = g^s \in G_{P_1}$  the query ciphertext is a normal ciphertext, the  $\text{Game}_{real}$  game is running:

- If  $T = g^s X^c \in G_{P_1 P_2}$ , then  $\vec{u} = c \cdot \vec{v}^t$ , the query ciphertext is a semi-function ciphertext, and the  $Game_0$  game is running.

Therefore, the algorithm  $\delta$  can overcome the hypothesis 1 with the advantage  $\varepsilon$  according to the adversary's output.

## 4 Conclusion

In this paper we proved that the scheme is secure in view of the problems of single authorization center and coarse-grained control in social network. This scheme not only solves the single node failure and performance bottlenecks of the authorization center, but also solves the problem of private key leakage by malicious users and achieves the completely fine-grained revocation. In fact, there are other aspects of revocation, such as the multi-authority CP-ABE [7], cyclical update of the user's private key [8] and other cancellation programs. Future direction of our research is how to generate a more secure and efficient solution based on existing works.

**Acknowledgments.** This work is supported by the National Natural Science Foundation of China under Grants No. 61632009 and the Science and Technology Project of Changsha under Grant No. kq1701089, and Fundamental Research Funds of Central South University under Grant No. 2017zzts713.

## References

1. Bethencourt, J., Sahai, A., Waters, B.: Ciphertext-policy attribute-based encryption. In: IEEE Symposium on Security and Privacy, pp. 321–334. IEEE Computer Society (2007)
2. Rahulamathavan, Y., Veluru, S., Han, J., et al.: User collusion avoidance scheme for privacy-preserving decentralized key-policy attribute-based encryption. *IEEE Trans. Comput.* **65**(9), 2939–2946 (2016)
3. Attrapadung, N., Libert, B., de Panafieu, E.: Expressive key-policy attribute-based encryption with constant-size ciphertexts. In: Catalano, D., Fazio, N., Gennaro, R., Nicolosi, A. (eds.) PKC 2011. LNCS, vol. 6571, pp. 90–108. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-19379-8\\_6](https://doi.org/10.1007/978-3-642-19379-8_6)
4. Wang, H.P., Zhao, J.J.: Ciphertext-policy attribute-based encryption with anonymous access structure. *Comput. Sci.* 2016
5. Ning, J., Dong, X., Cao, Z., et al.: White-box traceable ciphertext-policy attribute-based encryption supporting flexible attributes. *IEEE Trans. Inf. Forensics Secur.* **10**(6), 1274–1288 (2015)
6. Peng, K., Zhang, X.: Adaptively security CP-ABE scheme supporting attribute revocation. *Comput. Eng.* **41**(4), 151–155 (2015)
7. Li, Q., Xiong, J., Xiong, J., et al.: Provably secure unbounded multi-authority ciphertext-policy attribute-based encryption. *Secur. Commun. Netw.* **8**(18), 4098–4109 (2015)
8. Phan, D.H., Trinh, V.C.: Identity-based trace and revoke schemes. In: Boyen, X., Chen, X. (eds.) ProvSec 2011. LNCS, vol. 6980, pp. 204–221. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-24316-5\\_15](https://doi.org/10.1007/978-3-642-24316-5_15)

# The All Seeing Eye: Web to App Intercommunication for Session Fingerprinting in Android

Efthimios Alepis and Constantinos Patsakis(✉)

Department of Informatics, University of Piraeus,  
80, Karaoli & Dimitriou, 18534 Piraeus, Greece  
{talepis,kpatsak}@unipi.gr

**Abstract.** The vast adoption of mobile devices in our everyday lives, apart from facilitating us through their various enhanced capabilities, has also raised serious privacy concerns. While mobile devices are equipped with numerous sensors which offer context-awareness to their installed apps, they can be also exploited to reveal sensitive information when correlated with other data or sources. Companies have introduced a plethora of privacy invasive methods to harvest user's personal data for profiling and monetizing purposes. Nonetheless, up to now, these methods were constrained by the environment they operate, e.g. browser vs mobile app, and since only a handful of businesses could have access to both of these environments, the conceivable risks can be calculated and the involved enterprises can be somehow monitored and regulated. This work introduces some novel user deanonymisation approaches for device fingerprinting in Android. Having Android AOSP as our baseline, we prove that web pages, by using several inherent mechanisms, can cooperate with installed mobile apps to identify which sessions operate in specific devices and consequently to further expose users' privacy.

**Keywords:** Android · Privacy · Deanonymization · App collusion

## 1 Introduction

The unprecedented growth of mobile usage has radically transformed our daily lives. Besides the great advances in our communications, mobile devices have changed the way we create, process and consume information, as they realise pervasive and ubiquitous computing. Among others, one of the most significant emerged changes is how we value information. The fact that people are constantly and effortlessly connected to the Internet via devices which empower people's unobstructed communication, information flow and entertainment, in many occasions results in disregarding or underestimating the value of the information they consume and offer to third-parties.

As far as information offering is concerned, the value of the provided information to third-parties is most of the cases considerably high, something that is not

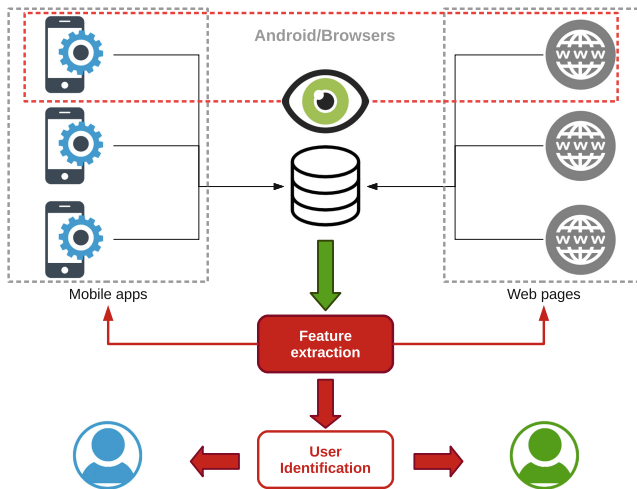


always understood by the users. For instance, one might share his location with an app or a web page neglecting the fact that this single piece of information also consists of a very sensitive piece of data which can be exploited for various purposes. Indicative uses for such location sharing could be the recommendation of other users in proximity for communication purposes, or even for sharing a ride. Aggregating location data from numerous users can provide real-time traffic analytics or insight about resource requirements in a smart city. Apparently, this information can stimulate businesses' prosperity by enabling the implementation of further customer-centered services. Therefore most companies are striving to extract from users as much information as possible.

While service personalisation can be considered as a noble cause, companies tend to exploit data further for profiling and targeted advertising, tactics that can expose users to many privacy hazards. This trend is highlighted by the fact that many companies are providing APIs which harvest user data to create fine grained user profiles, containing a lot of sensitive user information. Such practices have also led to the introduction of methods such as browser and device fingerprinting. Nonetheless, thus far mobile apps and web pages are considered as two diverse ecosystems as they refer to two discrete software environments with radical differences in their information flow and data usage. This distinction works for the benefit of users' privacy, since it allows some parts of their activities to remain isolated and hence private. For instance, it prevents an app from knowing which web pages a user visits, or a web page from knowing which apps a user is using and when. On the contrary, enabling access between these two environments could allow for a web page to communicate with an installed app to recover further personal and sensitive information from local files or sensor measurements, and hence to further reveal one's interests.

The goal of this work is to illustrate that there are currently several means to realise user identification in Android, regardless of the environment a software module is operating on. Despite the privacy hesitations of people towards the well-known tech giants or independent browsers, we provide some concrete examples proving that an "All Seeing Eye", a software entity able to monitor users' actions across both the web and the application environments, can be easily created. Such an entity, in the form of an online database equipped with some additional services can correlate information from web pages and mobile apps in order to identify individuals. After a thorough investigation in the related scientific literature and to the best of our knowledge, the authors of this paper have concluded that this problem has been so far partially studied, as current literature is focused on methods which examine each software ecosystem independently and not both of them as a whole. In fact, the proposed methods in this work can be considered as an extension of device fingerprinting as they do not solely depend upon unique characteristics of device components or hardware identifiers. We name these methods as "session fingerprinting" since their goal is to reveal whether web-browsing and software sessions operate simultaneously in a device and identify the user.

The generic concept of this work, in a simplified form, is illustrated in Fig. 1. Each side of this figure is dedicated to the two software “ecosystems”, namely web pages and mobile apps. Obviously, there is a crosscut from the OS, namely Android, since it manages calls from both ecosystems in a mobile device, as well as from the browsers which by their definition as applications belong in both ecosystems. The “All Seeing Eye” acts as a Command and Control, C&C, server which collects information from web pages and apps, correlates it and transmits “commands” and the corresponding information to both sides. The commands may range from “retrieve a list of installed apps” and “scan local storage for files containing X”, to “display ad Y” or “application Z send webpage data W”. Therefore, the “All Seeing Eye”, as the orchestrator of all performed actions by apps and web pages, can ultimately reveal user identities.



**Fig. 1.** Basic concept

While similar attempts have been made in the past, it is rather important to note that methods trying to escape the browser’s environment without user’s consent are considered to be malware and usually exploit browser’s vulnerability. Especially in the case of Android, passing a single bit of information from a benign browser to an app is rather difficult, given that it has not only to bypass the browser sandbox but additional obstacles due to Android’s security model which will be discussed later on.

## 2 Session Fingerprinting

When someone browses the Internet, his session is considered anonymous unless he has logged in a web page. While this anonymity is very convenient for ensuring

user’s privacy, companies strive to find ways to bypass it and to profile users. Frequently, as the benign goal behind these actions is regarded the adaptation and/or personalisation of a web page according to the corresponding user profile, which translates to better usability and increased content quality, which in turn may increase both views and viewers. In most of the cases, this personalisation targets the advertising industry, since by deanonymising an individual a business is able to display ads tailored to users’ preferences and therefore to increase both business’ and service providers’ profits.

One of the most widely used methods in achieving this is browser fingerprinting, which tries to deanonymise users by exploiting noticeable differences in the usage of different browsers such as the underlying OS, user agent, browser version, monitor size, or even installed fonts and plugins [10, 15, 16]. More advanced methods go a step further by exploiting device specific variations to identify individual devices. For smartphones, it has been shown that sensors, such as accelerometers, or speakers and microphones may have unique characteristics which differ, not only across models, but also across the devices within they operate due to calibration errors and frequency distortions [6, 9, 20].

While these methods have been proven efficient in many cases, they are usually subject to errors and software updates which could render a previous fingerprint useless. For instance, a browser update may change the user agent or the fonts, making impossible its linking to the previous fingerprint. However, what a company actually needs is to be able to correlate information with other affiliated parties in order to determine whether the user has been simultaneously operating another *session*. As a typical example can be regarded the parallel usage of a web page and a mobile app. Note that while the latter means that the app is running in the background, it is a typical situation in almost all mobile OSes. In this regard, both parties should try to create a unique ID for each session and also communicate it with each other in order to deanonymise the user. We name the methods for extracting these IDs as *session fingerprinting*.

Apparently, these methods depend on the existence of cooperating apps in the mobile device. We argue that this is a weak assumption as the considered adversary in this work is mainly an ad network. Due to the prevalence of the freemium model in Android, most applications are free and, most of the times, they come with at least one preinstalled ad component. However, our requirements do not imply escalated privileges, hence the resulting applications are easier to be accepted by the users.

Although both browsers and the Android OS have such privileges, namely are able to deanonymise the users and have data about them coming from multiple sources, the level of trust a user has to both of them is the ultimate one. User choose them because they trusts that they will act honestly and they will protect them from threats and, above all, they will not stalk them. Moreover, it is important to highlight that despite the OSes restrictions, the different existing ad frameworks may indeed perform user profiling, and in many cases, a quite intense one, but still they cannot escape the browser or the app ecosystem within they operate. Nonetheless, they are continuously acting more rogue,

despite their sandbox environment [12]. Stevens et al. [19] found that some ads would use undocumented permissions, such as read/write to calendar or access location and camera. Grace et al. [13] found that around half of them would probe the corresponding apps to determine whether they could abuse them for harvesting sensitive user information. More recently, it has been documented that ad networks are using commercial ultrasonic tracking technologies like SilverPush, Lisnr and Shopkick to determine either users' location, or the commercials they are watching [2]. Notwithstanding their invasiveness, to the best of our knowledge, none of them has been able to pass information from a browser to an app within the Android system. Instead, this kind of communication, whenever reported, was strictly among apps that used the same ad network.

### 3 Problem Setting

In principle, the information that can be collected from a web page is derived solely via the launched browser and is strictly limited to the browser's environment. Any attempt to access resources beyond the browser sandbox is considered as a security violation and therefore is characterized as malicious. To this end, browsers allow very limited exposure of user data to a web page. For instance, a web page cannot read from or write to the storage of a mobile device unless this action is user initiated. To overcome these restrictions, adversaries may resort to browser extensions [14] which provide even more capabilities. On top of that, nowadays, due to the growth of mobile devices, several standards, like HTML5, have been introduced to allow browsers to access additional resources, such as location, camera, or microphone, upon direct and explicit user consent. As a result, web pages have a growing set of capabilities, yet quite limited in comparison to apps.

On the other hand, Android apps reside on a different environment. Contrary to web pages, mobile apps “live” on the operating system and thus have more direct access to the hardware. Again, their access is limited according to the granted privileges from the users plus their scope is more fixed as they fulfill specific user needs. Due to this restriction, apps cannot determine which web pages a user visits. A critical distinction between Android apps and web pages is that apps always pass through an “installation” process. This step represents a user acknowledgment regarding the specific resources an app is allowed to use inside the environment it is executed. Notably, since Android Marshmallow users may grant and revoke permissions to specific resources, like camera, microphone, location etc., which are called “dangerous” and may hinder security and privacy issues. On the contrary, this is not the case for web pages where users do not have preconditions for visiting them. In many instances, users would like to be able to use “one-time apps” to accomplish specific tasks like using a retailer's app when browsing his web page. To address this need, Google recently introduced *instant apps*, which do not require installation, and have more permissions and/or capabilities than common web pages. However, instant apps, like web pages, are also restricted from accessing hardware identifiers to prevent user profiling.

**Table 1.** Capabilities of apps and web pages.

	Native apps	Instant apps	Web pages
Access device identifiers	✓		
Device external storage	✓		
Push notifications	✓		
List of installed apps	✓		
Access body sensors	✓		
Direct communication with installed apps	✓		
Receive broadcasts from OS or 3rd party apps	✓		
Run on the background	✓		
Change device settings	✓		
Access user calendar	✓	✓	
Access motion sensors	✓	✓	
Access contacts	✓	✓	
Access phone calls	✓	✓	
Access sensors	✓	✓	
Access environmental sensors	✓	✓	
Access location	✓	✓	✓
Access microphone	✓	✓	✓
Access position sensors	✓	✓	✓
Access camera	✓	✓	✓
Access internal storage (own storage)	✓	✓	✓
High precision timestamps	✓	✓	✓

Key differences in the number of capabilities of Android apps, both native and instant, and web pages are illustrated in Table 1. As expected from the earlier discussion, it can be easily noticed that installed applications have far more access to device resources than web pages, since users install apps granting themselves the corresponding permissions. On the contrary, due to the nature of the Web, users may visit a large number of different web pages on a daily basis, without knowing their quality, intentions, source or content. Hence, both Android and browsers make significant efforts, e.g. running in a sandbox environment, towards protecting users from malicious web page behaviour. Unquestionably, if an app had been able to communicate with a web page without restrictions, the entire underlying security infrastructure would have been rendered useless. In fact, even the most widely used apps in Android are not able to communicate directly with their web page. For instance, in the case of three well-known and widely used apps, Facebook, Instagram and Twitter, they do not transfer information to their corresponding web page or any other cooperating web page

when a user has logged in the app. Instead, a “Connect with Facebook/Twitter” button usually appears, requiring further user interaction and most importantly being realized by users. Evidently, had these apps been able to transfer this kind of information to the browser, they would have done it already long time ago, not only for facilitating users, but for increasing further the amount of collected user data and the quality of provided services.

Creating a mechanism being able to transmit an identifier from the browser to a cooperating installed app in a user’s device, or vice versa, would allowed for the installed app to identify the individual who visited the cooperating web page and subsequently, that web page would be instantly granted access to the same resources as those of the installed app. Further analyzing this, after both parties, namely web pages and apps have identified themselves lying in the same user’s device, they are able to create a covert channel. In the least sinister scenario, a web page cooperating with an app would be able to access a user’s contacts, SMS messages or even storage and microphone, without obtaining user’s consent, and would displayed ads perfectly tailored to the user’s profile, albeit violating his privacy. However, in a true malicious scenario, user data would be harvested by web pages and personalized exploits would be pushed to users’ devices to further exploit their personal data while they surf in the WWW.

In most of the Android cases documented by researchers, information is leaked from one app to another through a covert channel [8,11]. Although Rushanan et al. in [18] achieve a goal similar to ours, their study concerns only the desktop environment. Their approach consists in exploiting the Web Workers API in order to increase the CPU and memory utilization. By monitoring both CPU and memory usage, they manage to pass messages from a web page to an app in a desktop computer. However, this attack scenario is not possible in an Android device. For devices up to Marshmallow, while apps could monitor the `/proc/` folder and extract some information about memory usage, the recovered information is far from being considered fine-grained and does not include CPU usage. With the introduction of Nougat, apps are allowed to only access the contents of their own `/proc/PID` folder (<https://developer.android.com/about/versions/nougat/android-7.0-changes.html>), so this method does not work any more for AOSP. The only other alternative for an app to have this kind of access is to request the system-level permission `PACKAGE_USAGE_STATS` (<https://developer.android.com/reference/android/app/usage/UsageStatsManager.html>). The fact that their attack does not apply for passing messages in Android is also proved by the authors’ statement that in Android they managed just to launch a resource depletion attack against the browsers. Moreover, the aforementioned restrictions in Nougat prevent apps from accessing `/proc/net` which could otherwise reveal the domain names but not the full URL a user has visited.

Notably, developers in many occasions, despite Google’s recommendations (<https://developer.android.com/training/articles/user-data-ids.html>), use `ANDROID_ID` as a unique identifier. To restrict this, Google required for apps to request the dangerous permission `READ_PHONE_STATE` (<https://developer.android.com/reference/android/Manifest.permission.html>). Clearly, since this

ID is unique, installed apps may identify instances and correlate users and behaviours. Since such actions violate user privacy, even though they are performed locally only among installed apps, in the latest preview of Android O, Google decided to block this behaviour so that each app receives a different `ANDROID_ID`. More precisely, in Android O for each combination of application package name, signature, user, and device we end up with a different `ANDROID_ID` (<https://developer.android.com/preview/behavior-changes.html>). To further support users controlling their unique identifiers, Google has recently announced the new changes coming in Android O [5], regarding device identifiers. In this regard, Android O is limiting the use of device-scoped identifiers that are not resettable and is also updating the way that applications request account information and providing more user-facing control. The latter signifies that Google is not only aware of such deanonymization issues, but is also constantly working on refining its platform to mitigate these threats and restrict unauthorised and unregulated app to app communication, let alone web to app communication.

## 4 Intercommunication Between Apps and Web Pages

In the following subsections, we provide a set of concrete examples as Proofs of Concept, which showcase how apps and web pages can mutually create and consequently transmit unique IDs that allow them to link their usage and to communicate sensitive attributes to each other, realizing what the authors of this paper have introduced as “session fingerprinting”. The authors of this paper have responsibly disclosed the mentioned security issues described in the next subsections, regarding unauthorized communications between apps and web pages.

### 4.1 Location

Location awareness has undoubtedly increased the potential of many applications since it allows them to adapt accordingly and render their information based on location specific criteria, drastically improving e.g. user recommendations. Obviously, location is a sensitive piece of information as it can disclose many private attributes, ranging from work and residence location, to entertainment preferences and political/religious beliefs if correlated with other sources of information. Therefore, mobile OSes allow applications to access location data only if the user grants a corresponding permission. In Android this permission is provided either as *Fine* or as *Coarse* location.

Similarly, beyond the support for media in HTML5, the standard enables web pages to access user location. Since this information is sensitive, the browser specifically requests for user permission to be granted, even though this kind of information can be used for other purposes as well. Once the browser gets access to user’s location, the response contains apart from the longitude and the latitude, the accuracy and the timestamp [17] as well. Moreover, depending on the implementation, it may also return other values, such as heading and speed.

Interestingly, in our research we have come up with proofs that this information can be correlated with location data information from an Android app. More precisely, while an Android app could monitor a user's location and is able to correlate the coordinates with the ones that are received from a web page, one could argue that since these requests to location data are not made simultaneously, from the browser and the app, the actual identity of the user is not disclosed. This argument can be clearly supported either because other nearby users may be also implicated or because the web page gets this information only once. However, practically this is not the case. For reasons such as decreasing battery consumption, since the usage of the GPS is rather greedy, Android app developers may choose to use the "last known location" feature through `getLastLocation` of `LocationServices` which fetches the location from its cache [1]. Then, based on the accompanied timestamp the developer can determine whether he needs to request a new reading or not. Yet, what seems quite interesting in this case is that our findings reveal that by accessing the device's last known coarse location from an app, we may end up having data about the precise last location request made by a web page. It should be emphasized that "coarse" location is the one that should be used here, since "fine" location is accessed exclusively by android apps and not web pages and hence is not suitable for our method.

Apparently, if a mobile app monitors the last known location and its corresponding timestamp determined by the `Network Location Provider` and communicates this information to the All Seeing Eye, the latter is able to determine whether it coincides with user's coordinates and timestamp received from a web page. Beyond a doubt, this combination of data is quite "unique". Once a correlation is found, the All Seeing Eye can create a covert channel that will serve both the app and the web page to exchange data regarding this session.

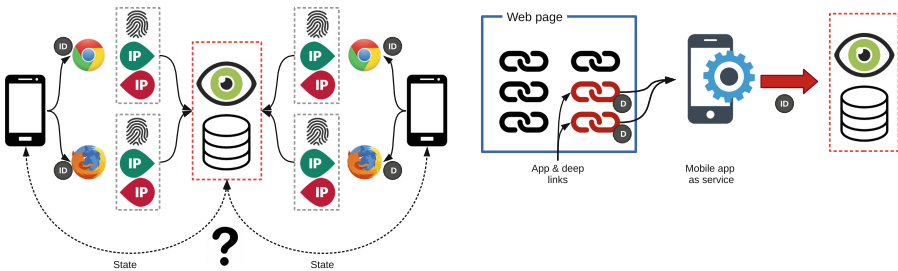
## 4.2 Browser Fingerprinting

In the previous example a dangerous, yet very commonly used permission, namely location, was used to identify a user. Nevertheless, one could achieve the same result without such permissions. A more stealth method is to utilize browser fingerprinting. To this end, we assume that the victim has installed an application which does not request any dangerous permission. According to the Android permission model, such applications are allowed to list all the installed applications in a device, hence the adversary also has knowledge of all the installed browser applications. This way, the app can subsequently open all the available browsers through intents and point them to a desired URL in order to obtain a fingerprint from them. Note that as of Nougat, an application cannot determine which the foreground application is, a piece of information that would have been very valuable for the adversary, however, the authors have already notified Google of a new method to achieve this in all versions prior to Nougat (Android Issue no 23504, triaged). Each time a browser is fingerprinted by the app, a random nonce is created and is sent in the web page request allowing the adversary to determine to whom its fingerprint belongs. This kind of attack utilizes malicious intents, while for "covering traces" purposes the



cooperating malicious web pages could redirect to a commonly used web page (e.g. a search engine), after accomplishing the ID exchanging job. A scenario where no intents are needed also exists, where a malicious app may use its native webview component in order to accomplish the aforementioned task.

Since mobile devices have less “unique” characteristics compared to personal computers, an extension to browser fingerprints is to additionally use even more mobile device characteristics, further conforming to the “session fingerprinting” proposed term. Indeed, both a mobile app and a web page can obtain knowledge about the internal and the external IP of the mobile device. For the former one could potentially use WebRTC [4] which is known to leak several pieces of private information [3]. Therefore, when a user visits a web page, the web page queries the “All Seeing Eye” to determine if someone with the specific browser fingerprint, public and local IPs has a cooperating app running at this timeframe. In general, the chances of this query returning more than one result are slim. Nonetheless, in a corporate environment where many people might have mobile devices of the same model, some instances may exist. In such environments one could have two identical devices with the same internal IP, if e.g. two users with the same smartphone model use the corporate WiFi on different floors or departments. To further reduce the query results, the “All Seeing Eye” could request the state of each device. In this case, additional information that can be cross-checked between apps and web pages include, but are not limited to, the following: battery information (both state and charging level), interval since device’s last noticeable movement (this could be determined e.g. via accelerometers), interval since last proximity (via proximity sensor), light measurements, positioning (e.g. facing up or down), or even some connection statistics such as `downloadMax` (one of the new features of HTML5 through the Network Information API [7]). From this information one can easily determine which user is using a smartphone at a specific timeframe and essentially to eliminate the possibilities of having false positives. The process is illustrated in Fig. 2a.



(a) Identification through browser fingerprinting.

(b) Identification through app and deep links.

**Fig. 2.** User identification methods.

In this figure, we may notice that both web pages inside browsers and also web pages inside applications' webview components are able to collect unique fingerprints, namely internal and external IPs, accompanied by information regarding the devices' state and communicate them to the "All Seeing Eye" which will then be responsible for finding exact matches between them.

### 4.3 App and Deep Links

Most users install plenty of apps on their devices, even though many of them might have some overlapping functionality. To facilitate user interaction between websites and Android applications the Web Intents framework was introduced, allowing a developer to specify how a hyperlink is handled on the user device, e.g. open the phone to dial up an already prepared number, or use Skype for a specific contact. However, Android supports further features through *App* and *Deep links*. The concept behind both of these types of web links is to open specific apps depending on the link. As an example, Facebook and Twitter apps are triggered when the user taps on a link referring to content of the corresponding site.

Interestingly, this kind of functionality is automatically activated when a user installs an application which provides such features, without requesting any user approval. Moreover, Android activities may run on the foreground in a "hidden" mode, either by using transparent themes or by utilizing floating zero-sized activities. Practically, this creates a hidden communication channel between web pages and apps that can be used to identify users as illustrated in Fig. 2b. We assume that an app is installed in a user's device having at least one "browsable" (declared inside Apps Manifest file) activity in order to enable "cooperation" with web pages. On the other side, web pages embed some special "intent" hyperlinks which also have the ability to carry a random ID, different for every interaction. Once a user taps in one of these links, the ID is bundled and transferred to the app, using the "getExtras" method inside the "Intent" Android class. As a final step, once again the data is communicated to the All Seeing Eye, deanonymising the user. This kind of "interaction" can be seen as a directed web-to-app communication, where an app has the ability to be reached from web pages. At the same time one or more web pages may utilize this "functionality" by transmitting an ID which is essential for the entire described scenario. Next subsection describes how this local channel communication can be achieved through the opposite direction of interaction.

### 4.4 Direct App to Web Communication

Following the same logic as in the previous subsections, an app is also able to directly reach a web page through device's installed browser. Once again, Android Intents are deployed in order to launch an installed browser and to pass a URL. The cooperating app is able to inject one or more string values inside a URL as parameters, which in our case is a simple, randomly generated ID and correspondingly fire a malicious intent towards a browser. As a next

step, the loaded web page extracts the ID from the URL’s “location search” property and thus a local covert channel between the app and the launched web page is created. Of course, the web page is again able to communicate the ID to the All Seeing Eye making the user’s profile available to others as well. As already discussed, this kind of method “leaves some traces” namely it opens a web browser, however the corresponding malicious web page can hide its traces with a simple redirection.

A quite significant detail in the described process is that the reached by the app web page should initially use a client side programming language in order to retrieve the transmitted ID. This is essential in order to create the local covert channel between the app and the web page, since an app ID directly transmitted to a web server would lose the ability to “find its way back” to the corresponding device’s web page.

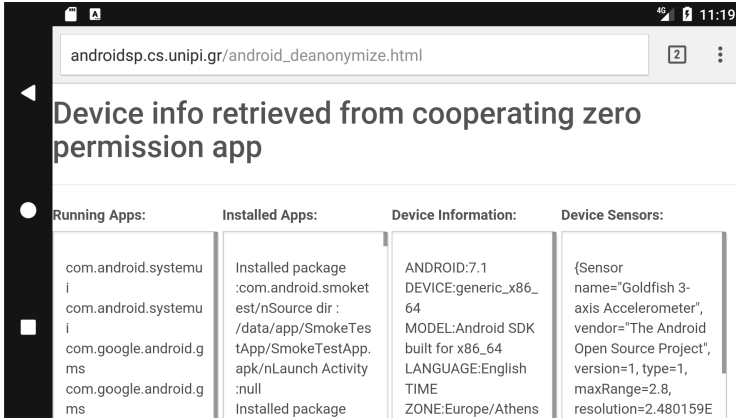
## 5 Experimental Results and Statistics

In order to provide an estimation of the potential exposure of users to these threats, the authors of this paper used data by utilizing “Tacyt” (<https://tacyt.elevenpaths.com>). Tacyt is an innovative cyber intelligence tool that facilitates research in Android mobile apps environments with big data technology. The aim of Tacyt is to enable quick detection, discovery and analysis of these threats to reduce their potential impact on organizations. Enabling app data mining and detection, enables research and analysis of the collected information from Google Play and other markets. Due to the implementation of Tacyt, the responses are per app version, and not per app, nonetheless they provide a very good overview of apps dating back to at least three years ago.

Table 2 presents the results from the performed queries. In our first query we tried to identify how many apps provide a deep or app link. This information is declared in the manifest of each application and is clearly marked with the `android.intent.category.BROWSABLE` tag in the XML. The next two rows involve app versions which required the `ACCESS_COARSE_LOCATION` and Internet permission which could be potentially used to deanonymize users. Similarly, using PublicWWW, a source code search engine for web pages, we found more than 96,000 web pages to use geolocation features in their code for locating users.

**Table 2.** Results from Tacyt.

	Google play		Other markets	
	Available	Unavailable	Available	Unavailable
App & Deep links	432,204	87,483	71,686	34
ACCESS_COARSE_LOCATION	996,326	215,335	154,749	29
INTERNET	4,044,922	1,046,310	533,492	149
Total versions in market	4,207,542	1,095,398	576,204	155



**Fig. 3.** Device info as obtained from a web page through a zero permission app installed in a device running Android 7.1.1.

Finally, we have implemented a proof of concept app that is able to cooperate with web pages without requesting any dangerous permissions. The app is available at [androidsp.cs.unipi.gr/android\\_deanonymize.html](http://androidsp.cs.unipi.gr/android_deanonymize.html). Once installed, the app recovers a lot of information from the device, such as installed and running apps, device info as well as measurements from many sensors which do not require any dangerous permissions. Eventually, as illustrated in Fig. 3, when the user visits the web page through his mobile browser he can verify that the web page has recovered all this information without requesting his consent. It is worth to be noted, that apart from providing access to sensors that a web page would not normally have, the measurements for these sensors are also listed in a fine grained mode, e.g. access to accelerometer detailed measurements which could be used for fingerprinting [9].

## 6 Conclusions

The ever increasing use of mobile devices exposes user privacy in numerous ways. Despite the fact that mobile OSes take several measures to protect their users, attackers seem to always be one step ahead. Nonetheless, most would agree that the state of the art countermeasures guarantee an independence between the browser and the mobile apps in a way that they cannot exchange information. Taking Android as our reference platform, we introduce new methods that exploit various inherent mechanisms to practically guarantee absolute identification with limited resource usage. Moreover, the proposed methods extend the notion of device fingerprinting to what we call session fingerprinting. Our techniques can be performed without accessing unique device characteristics or using dangerous permissions. In this regard, our techniques imply a bigger threat, as the covert channel that is created between the web pages and the apps cannot be traced easily.

Due to the fact that all the aforementioned mechanisms are inherent in Android, one cannot rule out the possibility of these mechanisms already been exploited, enabling unauthorised and unregulated cooperation between the two ecosystems. Clearly, this would greatly expose users' privacy, bypassing the permission model of the most widely used mobile platform. Addressing such issues is a rather challenging task, because, apart from changing the native Android mechanisms, it is also required for the OS to determine the context of some calls, either to prohibit access to resources, or to obfuscate the underlying information, since the calls seem legitimate.

**Acknowledgments.** This work was supported by the European Commission under the Horizon 2020 Programme (H2020), as part of the *OPERANDO* project (Grant Agreement no. 653704) and is based upon work from COST Action *CRYPTACUS*, supported by COST (European Cooperation in Science and Technology). The authors would like to thank *ElevenPaths* for their valuable feedback and providing them access to Tacyt.

## References

1. Android developers: getting the last known location (2017). <https://developer.android.com/training/location/retrieve-current.html>
2. Arp, D., Quiring, E., Wressnegger, C., Rieck, K.: Privacy threats through ultrasonic side channels on mobile devices. In: 2nd IEEE European Symposium on Security and Privacy (EuroS&P) (2017)
3. Beltran, V., Bertin, E., Crespi, N.: User identity for webrtc services: a matter of trust. *IEEE Internet Comput.* **18**(6), 18–25 (2014)
4. Bergkvist, A., Burnett, D.C., Jennings, C., Narayanan, A., Aboba, B.: WebRTC 1.0: real-time communication between browsers (2016). <https://www.w3.org/TR/webrtc/>
5. Blog, A.D.: Changes to device identifiers in Android O (2017). <https://android-developers.googleblog.com/2017/04/changes-to-device-identifiers-in.html>
6. Bojinov, H., Michalevsky, Y., Nakibly, G., Boneh, D.: Mobile device identification via sensor fingerprinting. arXiv preprint [arXiv:1408.1416](https://arxiv.org/abs/1408.1416) (2014)
7. Cáceres, M., Jiménez Moreno, F., Grigorik, I.: Network information API (2017). <http://wicg.github.io/netinfo/>
8. Chandra, S., Lin, Z., Kundu, A., Khan, L.: Towards a systematic study of the covert channel attacks in smartphones. In: Tian, J., Jing, J., Srivatsa, M. (eds.) *SecureComm 2014*. LNCS, vol. 152, pp. 427–435. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-23829-6\\_29](https://doi.org/10.1007/978-3-319-23829-6_29)
9. Dey, S., Roy, N., Xu, W., Choudhury, R.R., Nelakuditi, S.: Accelprint: imperfections of accelerometers make smartphones trackable. In: *Proceedings of the Network and Distributed System Security Symposium (NDSS)* (2014)
10. Eckersley, P.: How unique is your web browser? In: Atallah, M.J., Hopper, N.J. (eds.) *PETS 2010*. LNCS, vol. 6205, pp. 1–18. Springer, Heidelberg (2010). [https://doi.org/10.1007/978-3-642-14527-8\\_1](https://doi.org/10.1007/978-3-642-14527-8_1)
11. Gasiór, W., Yang, L.: Exploring covert channel in Android platform. In: 2012 International Conference on Cyber Security (CyberSecurity), pp. 173–177. IEEE (2012)

12. Goodin, D.: Beware of ads that use inaudible sound to link your Phone, TV, Tablet, and PC (2015). <http://arstechnica.com/tech-policy/2015/11/beware-of-ads-that-use-inaudible-sound-to-link-your-phone-tv-tablet-and-pc/>
13. Grace, M.C., Zhou, W., Jiang, X., Sadeghi, A.R.: Unsafe exposure analysis of mobile in-app advertisements. In: Proceedings of the Fifth ACM Conference on Security and Privacy in Wireless and Mobile Networks, WISEC 2012, pp. 101–112. ACM (2012)
14. Kapravelos, A., Grier, C., Chachra, N., Kruegel, C., Vigna, G., Paxson, V.: Hulk: eliciting malicious behavior in browser extensions. In: USENIX Security, pp. 641–654 (2014)
15. Mowery, K., Bogenreif, D., Yilek, S., Shacham, H.: Fingerprinting information in Javascript implementations. In: Proceedings of W2SP, vol. 2, pp. 180–193 (2011)
16. Mowery, K., Shacham, H.: Pixel perfect: fingerprinting canvas in HTML5, pp. 1–12 (2012)
17. Popescu, A.: Geolocation API Specification, 2nd edn. (2016). <https://www.w3.org/TR/geolocation-API/>
18. Rushanan, M., Russell, D., Rubin, A.D.: MalloryWorker: stealthy computation and covert channels using web workers. In: Barthe, G., Markatos, E., Samarati, P. (eds.) STM 2016. LNCS, vol. 9871, pp. 196–211. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46598-2\\_14](https://doi.org/10.1007/978-3-319-46598-2_14)
19. Stevens, R., Gibler, C., Crussell, J., Erickson, J., Chen, H.: Investigating user privacy in Android ad libraries. In: Proceedings of the 2012 Workshop on Mobile Security Technologies (MoST) (2012)
20. Zhou, Z., Diao, W., Liu, X., Zhang, K.: Acoustic fingerprinting revisited: generate stable device id stealthily with inaudible sound. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 429–440. ACM (2014)

# An Efficient Hierarchical Identity-Based Encryption Scheme for the Key Escrow

Yuanlong Li, Fang Qi, and Zhe Tang<sup>(✉)</sup>

School of Information Science and Engineering, Central South University,  
Changsha 410083, China

{liyuanlong, csqifang, tz}@csu.edu.cn

**Abstract.** Hierarchical Identity-Based Encryption (HIBE) is a generalization of Identity-Based Encryption (IBE) that mirrors an organizational hierarchy, however, the residual key escrow problem has still remained. This paper introduces a new scheme, based on the security notion of anonymous ciphertext indistinguishability against key generation center (ACI-KGC) security proposed by Chow, to remove the inherent key escrow problem. The scheme employs multiple Identity Certification Authorities (ICAs) which can ensure that the Private Key Generators (PKGs) distribute private key without knowing the list of user's identity information, so as to prevent malicious PKGs from decrypting message on behalf of user or maliciously make user's private key public. Security analysis shows that the scheme can solve the key escrow with keeping the high-efficiency and security of HIBE system. In contrast to Chen's T-HIBE and Chow's scheme, to generating the user's private key, our scheme does not require the multiple KPAs or the complex interactive protocol which take too much extra computation costs. And comparing with GS-HIBE, our scheme focuses on solving the key escrow problem with outstanding performance and security of HIBE system.

**Keywords:** Hierarchical identity-based encryption · Key escrow  
ACI-KGC · Private Key Generator · Identity certification authority

## 1 Introduction

### 1.1 Hierarchical Identity-Based Encryption

HIBE was first introduced by Horwitz and Lynn with a 2-level HIDE scheme [1]. The first efficient construction for HIBE was proposed by Gentry and Silverberg where security is based on the Bilinear Diffe-Hellman (BDH) assumption in the random oracle model [2]. A subsequent construction was proposed by Boneh and Boyen based on BDH where the security without random oracle under a weaker selective-ID secure [3, 4]. Subsequently, a series of variants have been presented to achieve stronger notions of security in standard model since then [5, 6].

In HIBE system, identities are vectors. The identity of lower-levels Private Key Generators (PKGs) or users denoted by a vector  $\{ID_1, ID_2, \dots, ID_t\}$  and

a postfix identity of  $t$  is referred to the layer of the domain PKG. There are four algorithms about how to generate private keys of lower-levels PKGs or users, and towards the HIBE, the private key of an identity  $\{ID_1, ID_2, \dots, ID_t\}$  at level  $t$  can be generated by its parent domain PKG with identity  $(ID - tuple_t)$ . Below, we present procedure of the algorithmic description in brief:

**Setup:** The root PKG takes a security parameter  $\lambda$  as input to expose the system public parameter and the master public key.

**KenGen:** The domain PKG utilizes the identity information of recipients combining with its master private key to generate the private keys of recipients at lower-levels.

**Encrypt:** The sender utilizes the identity information of recipient as the public key to produce the ciphertext to recipient.

**Decrypt:** The recipient can decrypt the ciphertext through the private key generated by the upper-layers.

## 1.2 Overview of Our New HIBE Scheme

HIBE is a natural tool for protecting communication confidentiality and privacy, but the PKG knows all the private keys it generated and thus can easily decrypt messages on behalf of users, or maliciously make users' private keys public, that is, key escrow problem [7, 8]. Due to this problem, we present a new and secure HIBE scheme, which efficiently solves the key escrow problem and give the more secure way to distribute the private keys for entities:

1. This paper presents this method based on the scheme of Chow's security notion but without the complicated interactive key generation protocol [9, 10].
2. We combine Identity Certification Authorities (ICAs) and hierarchical PKGs to generate the private keys of all the users. Through the user's identity certification, ICA can simplify the function and authority of PKGs.
3. With ICA and PKG generating the private key of user, we apply this procedure into the whole HIBE system. And then, we can analysis the security of our scheme through the game between challenger and adversary.

## 1.3 Organization

This paper is organized as follows. Section 2 describes the notions of bilinear map and some computational complexity. In Sect. 3, we present our modified HIBE scheme which is used in removing key escrow problem in detail. Section 4 analyzes the security of our scheme. In Sect. 5, we make a comparison among GS-HIBE, Chen's T-HIBE [5, 6] and our new EF-HIBE scheme. Finally, we summarize the main result obtained and conclude the paper.

## 2 Preliminaries

Below, we review the background knowledge of Hierarchical Identity-Based Encryption and describe the implication of bilinear pairing. Additionally, we also review the definition of the complexity assumption on which the security of the scheme we presented is based.



## 2.1 Bilinear Groups Pairing

The bilinear groups were first introduced by Boneh et al. [4]. Let  $G_1$  and  $G_2$  be two cyclic groups of prime order  $p$  where  $g$  is a generator of  $G_1$  and  $e : G_1 \times G_1 \rightarrow G_2$  is a bilinear groups, which should have the following properties.

**Bilinear:**  $\forall g_1, g_2 \in G_1, \forall a, b \in Z_p^*$ , We have  $e(ag_1, bg_2) = e(g_1, g_2)^{ab}$ ;

**Non-degeneracy:**  $e(g_1, g_2) \neq 1$ ;

**Computable:** There is an efficient algorithm to compute  $e(g_1, g_2)$  for any  $g_1, g_2 \in G_1$ .

## 2.2 Complexity Assumptions

Below we will present the complexity assumption and the main idea about Bilinear Diffie-Hellman (BDH), which has been used to construct the first efficient HIBE scheme of GS-HIBE [2]. Let  $G_1, G_2$  be two cyclic groups of prime order  $p$ ,  $e : G_1 \times G_1 \rightarrow G_2$  be an admissible bilinear group and let  $a, b, c \in_u Z_p^*$  be chosen at homogeneous randomly and  $g$  is a generator  $G_1$ . The BDH assumption has been introduced by BF-IBE which is not probabilistic polynomial-time algorithm  $\mathcal{B}$  can compute  $e(g, g)^{abc}$  from the tuple  $(g, g^a, g^b, g^c)$  with more than a negligible advantage  $\varepsilon$ . The advantage of  $\varepsilon$  is

$$\left| \Pr \left[ B(g, g^a, g^b, g^c) = e(g, g)^{abc} \right] \right| \geq \varepsilon \quad (1)$$

The probability is much greater than both the random choice of  $a, b, c$  in  $Z_p^*$  and the generator  $g$  in  $G_1$ .

**Definition 1.** The BDH assumption holds in  $G$  if no  $t$ -time algorithm has advantage at least  $\varepsilon$  in solving the BDH problem in  $G$ .

## 2.3 ACI-KGC Security Definition

The security notion of anonymous ciphertext indistinguishability against the key generation center (ACI-KGC) was first proposed by Chow [10, 11], this secure model provided an “embedded identity encryption” random oracle, which use anonymity against a malicious KGC to prevent user privacy from referring to the secret keys of users, and then avoiding the inherent key escrow problem. If the KGC cannot distinguish the intended recipient of the ciphertext from the identity space in polynomial time, it cannot decrypt the ciphertext by enumerating and generating all identity-based secret keys. Thus the notion of ACI-KGC depends on the number of random bits of identity, which protected the recipient identity of encryption message is anonymous. Based on that, we can ensure that KGC cannot realize whose the secret keys belong to, and then it cannot forge the signatures of recipients. Consequently, the scheme of Chow’s is proved to be safe.

Specific to the scheme, the definition of that can be represented for a game between the adversary  $\mathcal{A}$  and challenge  $\mathcal{C}$ . The meaning of the game is that the

adversary  $\mathcal{A}$  has to distinguish the two messages from challenged ciphertext in polynomial-time. If not, the IBE system is ACI-KGC secure. Below, we review the game  $Exp_{IBE,\mathcal{A}}^{ACI-KGC}(\lambda)$  as follows.

**Setup:**  $(\lambda) \rightarrow (param)$  The challenger  $\mathcal{C}$  runs the initialized algorithm Setup to obtain the public parameters  $params$  and a description about plaintext space  $M$ .

**KeyGen:**  $(gen, param) \rightarrow (mpk, msk)$  The adversary  $\mathcal{A}$  runs the KeyGen algorithm to generate the master public key  $mpk$  and master secret key  $msk$ . It gives  $mpk$  to the challenger and saves the  $msk$  to itself.

**Query:**  $A_{(mpk, ID^*)}^{enco(m_i)} \rightarrow (m_0^*, m_1^*)$   $\mathcal{A}$  makes the  $q_E$  times embedded-identity encryption oracle queries with message  $m_i$  as input (for  $i \in \{1, 2, \dots, q_E\}$ ), the algorithm  $enco(mpk, ID^*)^{(m_i)}$  returns  $Enc(mpk, ID^*, m_i)$  as the ciphertext based on a specific identity  $ID^*$ . Then  $\mathcal{A}$  finds two plaintext with equal length  $m_0^*, m_1^*$ , if  $m_0^*, m_1^* \notin M$  or  $|m_0^*| \neq |m_1^*|$  then returns 0.

**Challenger:**  $\mathcal{A}$  sends  $m_0^*, m_1^*$  to the challenger  $\mathcal{C}$ ,  $\mathcal{C}$  selects a random  $b \in \{0, 1\}$  and returns the challenger ciphertext  $C = Enc(mpk, ID^*, m_b^*)$  to  $\mathcal{A}$ .

**Guess:**  $\mathcal{A}$  outputs a bit  $b'$ , if  $b = b'$   $\mathcal{B}$  returns 1 ( $\mathcal{A}$  winning the game), else returns 0. Where the advantage of  $\mathcal{A}$  is defined as:

$$\left| \Pr \left[ Exp_{IBE,\mathcal{A}}^{ACI-KGC}(\lambda) = 1 \right] - \frac{1}{2} \right| \quad (2)$$

**Definition 2.** The IBE scheme is  $(t, q, \varepsilon)$  ACI-KGC secure if there is no such an adversary of polynomial time complexity algorithm to make at most  $q_E$  times embedded-identity encryption oracle queries, in other words, the adversary has lost the game.

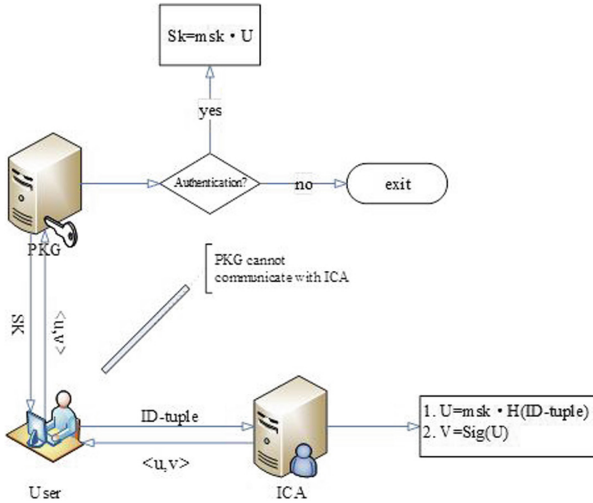
### 3 Construction of Escrow-Free HIBE

This section first presents the foundation of our HIBE scheme and describes the scheme about how to remove key escrow problem from HIBE.

#### 3.1 Removing Key Escrow Problem from HIBE

The essence of the key escrow problem is that the PKG owned the master secret key which has generated and known the private keys of users. In order to restrict the authority on-PKG, we will take over the ICA to weaken the function of PKG. As a trusted third party, ICA is responsible for issuing some kinds of certificates, based on that, we can ensure the privacy of users' private keys. ICA has the function to generate the special identity certificates which can conceal the privacy information of users. Accordingly, PKG has the authority to verify the users' identity and then generate the private keys which will be sent to recipients through the secure key distributing channel.

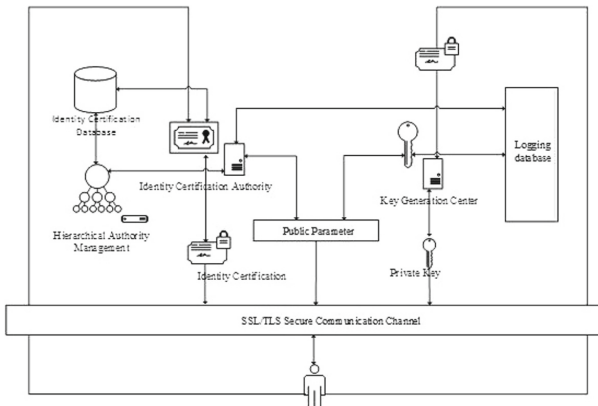
In order to reduce the costs of storing and processing in HIBE, we only setup one ICA for each layer, as well improving the security of the HIBE, the private



**Fig. 1.** The procedure of single mechanism.

keys of users can be generated by the domain PKG from upper-layers. First, the single mechanism of our scheme is given in Fig. 1 in which the procedure of the identity certification and private key generation is clearly depicted.

The user sends the identity to ICA which can generate the identity certificate of user and then return back to the recipient. With the intermediary of recipient’s identity certificate, PKG can generate the private key to the recipient without knowing the list of recipient’s identity information, which corresponds to the recipient and sends the private key to the recipient through the secure key distributing channel. In Fig. 2, we give the general framework of our scheme. Based on that, we can easily get the procedure of the identity certification and the private key generation.



**Fig. 2.** Our scheme architecture

### 3.2 Our New Escrow-Free HIBE Scheme

Our construction is based on the GS-HIBE which is the first efficient model for HIBE, and it was proposed by Gentry and Silverberg where the security is based on the Bilinear Diffie-Hellman (BDH) assumption in the random oracle model. First, we list some proper variables for a simple description in our EF-HIBE scheme in Table 1. Then, we give the algorithmic description of our scheme briefly as followings.

**Table 1.** Proper nouns for a single description

Proper nouns	Description
$PKG$	Private key generator
$G_1, G_2$	$G_1$ is the addition cyclic groups $G_2$ is the multiplicative cyclic groups
$p$	The prime order of $G_1, G_2$
$P_0$	The generator of $G_1$
$ID(id - tuple_t)$	The identity of <i>users</i> or <i>PKGs</i> at $t$ level
$SK$	Private key of <i>users</i>
$h_0, h_t^i$	Public keys of root <i>PKG</i> , and domain <i>PKGs</i>
$s_t^i$	The secret information of the domain node $i$ belongs to $t$ level
$s_0, s_t^i$	$s_0$ is the master key of root <i>PKG</i> $s_t^i$ is the random number as the master keys of the domain node $i$ belongs to $t$ level

**Setup:** The initialize algorithm of system can take over the security parameter  $\lambda$  as inputs to obtain the system parameters. We divide the algorithm into two parts, the PKG setup and the ICA setup. The PKGs setup includes root PKG setup and domain PKGs setup. On account of only one ICA in each layer, which does not refer to the parent-child or the domain ICAs, the setup of each ICA is as the same as others.

**Root PKG Setup:**

1. In brief, we can generate the cyclic group  $G_1, G_2$  and the bilinear pairing  $e : G_1 \times G_1 \rightarrow G_2$  through running the group generator function  $IG$ ;
2. The root PKG picks a random number  $s_0 \in Z_p^*$  as its master secret key where  $p$  is the prime order of the cyclic group. Based on the random number, we can calculate the master public key  $h_0 = s_0 \cdot P_0$  where  $P_0$  is the generator of  $G_1$ .
3. Choose four unidirectional hash functions  $H_1 = \{0, 1\}^* \rightarrow G_1, H_2 = \{0, 1\}^n \rightarrow G_2, H_3 = \{0, 1\}^n \times \{0, 1\}^n \in Z_p^*, H_4 = \{0, 1\}^n \rightarrow \{0, 1\}^n$ ;

4. The definition of message is  $M = \{0, 1\}^n$  and the ciphertext space is  $C = \{0, 1\}^n \times \{0, 1\}^n$ ;
5. Expose the system parameters which are  $\{G_1, G_2, e, H_1, H_2, H_3, H_4, p, P_0, h_0\}$ .

**Domain PKG Setup:** The master secret keys of domain PKGs are generated by the PKGs of parent node in the upper-layers, the same as the private keys of users in the lower-layers. We generate the master secret keys of domain PKGs by the parent node PKGs through the unique identity of domain PKGs, which generate the private keys of users in the lower-layers.

1. We calculate  $Q_t = H_1(ID - tuple_t)$ , where  $ID - tuple_t$  is the identity of PKG in the  $t$  level;
2. The domain PKG picks a random number  $s_t^i \in Z_p^*$  as the secret value, which represents the  $i$ th element in  $t$  level;
3. Calculate the master secret key of domain PKG through the hash value  $S_t^i = S_{t-1}^j + s_{t-1}^j \cdot Q_t$ ,  $S_{t-1}^j$  is the upper layer PKG which means the parent node of the domain PKG;
4. The master public key of domain PKG is  $h_t^i = s_t^i \cdot P_0$ .

**ICA Setup:** The root ICA and domain ICA has the same authority and parameters, which are responsible for generating identity certificates to the users at the lower-layers.

1. The ICA picks a random number  $s_{t/ICA} \in Z_p^*$  as the master private key;
2. Then calculate the master public key of  $t$  level  $P_{t/ICA} = s_{t/ICA} \cdot h_t^i$ .

**Extract:** The user applies for the private key to the PKG at the upper-level with his identity certificate.

1. The user first authenticates its identity to ICA, ICA computes the hash value of user's identity  $Q_t = H_1(ID - tuple_t)$ ;
2. Compute the intermediate value  $U_t = s_{t/ICA} \cdot Q_t$ ;
3. Run the signature algorithm  $V_t = sig(U_t)$  where the  $sig(\cdot)$  is a normal signature algorithm;
4. User sends the identity certificates  $(U_t, V_t)$  to the parent-node PKG to authenticate and generate the private key of recipient. First, the parent-node PKG verifies the  $U \stackrel{?}{=} veri(V_t)$ , if success, the parent-node generates the user private key  $SK_t = S_{t-1}^i + s_{t-1}^i \cdot U_t$ , else quit.

**Encrypt:** To encrypt message by utilizing the recipient's identity as the public key, the sender dose the following.

1. Compute the hash value of user's identity  $Q_t = H_1(ID - tuple_t)$ ;
2. The sender picks random bits  $\sigma = \{0, 1\}^n$ ;
3. The sender computes  $r = H_3(\sigma, M)$ ;

4. Compute the elements of ciphertext, then set the ciphertext to be  $C = \{rP_0, rQ_2, \dots, rQ_t, V, W\}$ 
  - (a) Compute  $V = \sigma \oplus H_2 \left( g_{ID-tuple_t}^r \right)$ ,  $g_{id-tuple_t} = e(P_0/ICA, Q_1)$ ;
  - (b) Compute  $W = M \oplus H_4(\sigma)$ .

**Decrypt:** When the recipient receives the ciphertext, then decrypts the ciphertext by utilizing its private key. We give the calculation steps as follows.

1. The recipient computes
 
$$\begin{aligned}
 & V \oplus H_2 \left( \frac{e(rP_0, SK_t)}{\prod_{t=2}^n e(P_t/ICA, rQ_t)} \right) \\
 &= V \oplus H_2 \left( \frac{e \left( rP_0, \sum_{t=1}^n s_{t-1/PKG}^i U_t \right)}{\prod_{t=2}^n e(s_{t-1/ICA} \cdot s_{t-1/PKG} \cdot P_0, rQ_t)} \right) \\
 &= V \oplus H_2 \left( \frac{e(rP_0, s_0/PKG \cdot s_0/ICA \cdot Q_1) \cdot \prod_{t=2}^n e(rP_0, s_{t-1/PKG}^i \cdot s_{t-1/ICA} \cdot Q_t)}{\prod_{t=2}^n e(s_{t-1/ICA} \cdot s_{t-1/PKG} \cdot P_0, rQ_t)} \right) \\
 &= V \oplus H_2 \left( \frac{e(rP_0, s_0/PKG \cdot s_0/ICA \cdot Q_1) \cdot \prod_{t=2}^n e(rP_0, s_{t-1/PKG}^i \cdot s_{t-1/ICA} \cdot Q_t)}{\prod_{t=2}^n e(s_{t-1/ICA} \cdot s_{t-1/PKG} \cdot Q_t, rP_0)} \right) \\
 &= V \oplus H_2 \left( e(rP_0, s_0/PKG \cdot s_0/ICA \cdot Q_1) \right) \\
 &= V \oplus H_2 \left( e(P_0/ICA, Q_1)^r \right) \\
 &= \sigma;
 \end{aligned}$$
2. And then, the recipient computes

$$W \oplus H_4(\sigma) = M \oplus H_4(\sigma) \oplus H_4(\sigma) = M;$$

3. Last, compute

$$r = H_3(\sigma, M);$$

From the complete process of our computing, we can assess the situation of  $r$ . If it is the same as the calculated results, we can reach the conclusion of security of the scheme, or else, the scheme is not secure.

## 4 Security Analysis

We give the security analysis about our new Escrow-Free HIBE scheme based on the ACI-KGC security. As described above, we give the details of our new EF-HIBE, which generates the private keys of users in the framework without knowing the users' identity information. The security of new EF-HIBE relies on the sequence of security games. We design the indistinguishable games and give the detailed statements to prove the security of our new EF-HIBE as follows, which maintains the ACI-KGC security.

**Theorem 1.** Breaking through the mission of new EF-HIBE can be contributed to the ability to solving the BDH problems. If the BDH is hard, the proposed new scheme is similar to GS-HIBE which can be proved to be ACI-KGC secure.

**Proof.** We prove the theorem by an indistinguishable game between challenger and adversary. Suppose that there exists an adversary  $\mathcal{A}$  that breaks the ACI-KGC security of our new EF-HIBE. We construct a Probabilistic Polynomial Time (PPT) algorithm  $\mathcal{B}$  to solve the BDH problems.

**Setup:**

1.  $\mathcal{B}$  exposes the public system parameters  
 $params = \{G_1, G_2, e, P_0, p, h_0, H_1, H_2, H_3, H_4\}$  where  $P$  is the generator of  $G_1$  and the prime order of  $\{G_1, G_2\}$  is  $p$ ,  $\{H_1, H_2, H_3, H_4\}$  are four hash functions,  $\mathcal{B}$  returns the  $params$  to  $\mathcal{A}$ .
2.  $\mathcal{A}$  Randomly picks  $s_i \in Z_p^*$ , and computes the public master keys of domain PKGs  $P_{PKG_i} = s_i \cdot P_0$ , it returns  $P_{PKG_i}$  to  $\mathcal{B}$ .

**Phase 1:** In the new EF-HIBE system, the function of ICA encrypts and conceals the identity information of users. Moreover, in order to increase the security of our scheme, we give more power to the adversary  $\mathcal{A}$  which can make the  $q_E - times$  embedded-identity encryption oracle queries with message  $m_i (i \in \{1, 2, \dots, q_E\})$  as adaptive input.  $\mathcal{B}$  responds these queries with the selected identity  $ID_i^*$  which leaks to the  $\mathcal{A}$ .

**Challenge:**

1. After the adversary  $\mathcal{A}$  selecting to end the phase 1,  $\mathcal{A}$  inputs two messages  $m_1, m_2$  with the equal length as the object to be challenged;
2.  $\mathcal{B}$  randomly generates a bit  $b \in \{0, 1\}$ , the random number  $\sigma$ , and then computes the correlative coefficient  $r = H_3 \{\sigma, m\}$ ;
3. Based on these elements,  $\mathcal{B}$  encrypts the  $m_b^*$  with the identity public key  $ID^*$ , and responds the ciphertext

$$C = \{rP_0, rQ_2, \dots, rQ_t, \sigma \oplus H_2(e(P_{ICA}, Q_1)^r), m \oplus H_4(\sigma)\} \text{ to } \mathcal{A}.$$

**Guess:**  $\mathcal{A}$  outputs a bit  $b'$ , if  $b = b'$   $\mathcal{B}$  outputs 1, else outputs 0.

Where the advantage of  $\mathcal{A}$  is defined as  $|\Pr [b = b'] - \frac{1}{2}|$ .

**Probability Analysis:** In the above indistinguishable game, under the basis of what the adversary  $\mathcal{A}$  knows the public master keys of root PKG and ICA,  $\mathcal{B}$  can solve the BDH problem with the non-negligible advantage  $\varepsilon$ . As a result,  $\mathcal{A}$  can win the above challenge game on the condition of encrypting and hiding the identity information of users by ICA with a non-negligible probability.

$$\left| \Pr \left[ B(g, g^a, g^b, g^c) = e(g, g)^{abc} \right] \right| > \varepsilon \quad (3)$$

where  $g$  is the generator of  $G_1$ .

In the new scheme above, there are no efficient solutions to solve the BDH problem. Additionally, the solution of users' identity information relies on the

intractability of the Elliptic Curve Discrete Logarithm (EDCL) problem, and the identity information of users is encrypted and hidden by ICA with its master public key, the hardness of computing the identity information of users by PKG is the same with solving the ECDL problem. Based on above that, our scheme can be proved secure.

## 5 Performance Analysis

In the above sections we have described our scheme in detail and give the security analysis. Researches on HIBE system almost concentrate on the improvement of security and performance. In this section, we analyze the performance of our scheme in communication and computation costs, and divide the analysis into two parts, including the single mechanism and the whole scheme.

### 5.1 The Comparison of Single Mechanism

In this part, we simplify our whole scheme into a single mechanism, and make comparison of the performance among Chow's scheme, the single mechanism of Chen's T-HIBE scheme [6] and ours. The simple descriptions of the computation notations are listed in Table 2.

**Table 2.** Description of notation

Notation	Description
$E$	The operation time of bilinear map $e$
$SM$	The operation time of scalar-multiplication based on group
$PM$	The operation time of point-multiplication based on group
$EA$	The operation time of exponentiation based on group
$Sig$	The operation time of signature based on group
$Veri$	The operation time of verification based on group
$H$	The operation time of hash function
$t$	The level of the recipient in HIBE

The mission of identity authentication and key generation in Chow's scheme is handed over to the ICA and KGC. Chow verified the legitimacy of users' identities through the identity certificates issued by ICA. After verification, KGC utilized the anonymous key generation protocol [9] consisted of four polynomial-time algorithms to generate the private keys for users, but which costs much more extra communication and computation costs than ours. In the Chen's mechanism, PKG and KPAs combined the blind factor and private key number factor which were generated by users to generate the private key elements, then the final private keys was generated by users themselves. Especially, our single mechanism is the most efficient among the three schemes according to the Table 3 [12].



**Table 3.** The comparison of three single mechanism

	Chow's	Chen's	Ours
Setup	PKG: randomly pick $msk$ and generate $mpk$ ICA: randomly pick $msk$	PKG: generate $mpk$ $n * KPA$ : randomly pick $msk$ and generate $mpk$ User: compute the blind factor $X_1$ and $X_2$	PKG: generate $mpk$ ICA: randomly pick $msk$ and generate $mpk$
Extract	ICA: $U = sig(ID)$  PKG: $ID = veri(U)$ Carry out the interactive protocol to generate	PKG: blind $SK$ elements $SK_1$ and $Sig(ID)$ $n * KPA_i$ : $veri(ID)$ and blind $SK$ elements $SK_1$  User: Combing $SK_1$ and $SK_2$ to generate the final $SK$	ICA: $U = msk \cdot H(ID)$ $V = Sig(U)$  PKG: $U = veri(V)$ and generate $SK$
Encrypt	$enc(M, ID)$	$enc(M, ID)$	$enc(M, ID)$
Decrypt	$dec(C, SK)$	$dec(C, SK)$	$dec(C, SK)$

### 5.2 The Comparison of Whole Scheme

We make a comparison of performance among GS-HIBE, Chen's T-HIBE and ours, then show that as follows in Table 4. The major computing cost is evaluated by the operation time of the multiplication and exponentiation based on the group, especially the bilinear pairing.

**Table 4.** The comparison of total costs in the three scheme

Cost	Scheme		
	GS-HIBE	T-HIBE	Our-EF-HIBE
$Sig$	1	1	1
$Veri$	1	$n + 1$	1
$SM$	$2t$	$t + 8$	$t + 1$
$PM$	$t$	$2t + n$	$2t - 2$
$EA$	1	3	1
$E$	$t + 1$	$t + n + 5$	$2t$
$H$	$2(t + 1)$	$t + n + 5$	$2t + 6$
$EF - Problem$	N	Y	Y

Our scheme utilizes ICA to generate the private keys to restrict the power of PKG, which efficiently solves the key escrow problem in HIBE system. The numbers in Table 4 indicates the extra communication and computation times. Compared with the GS-HIBE, even though we significantly increase extra computation costs in the key generation procedure, our scheme solves the inherent difficulty of key escrow problem. Additionally, from the comparison, we can easily find that our scheme is more efficient than Chen's T-HIBE.

## 6 Conclusion

In this paper, we proposed a new scheme based on anonymous ciphertext indistinguishability against key generation center (ACI-KGC) security. Additionally, our new scheme is similar to GS-HIBE which the main idea of our scheme was how to solve the inherent key escrow problem in the HIBE system. The third trust part ICA utilized the identity certificates to conceal the private identity information of users to restrict the power of PKG, which prevented the malicious PKG from getting the private keys to decrypt the ciphertext easily. From the comparison, our new EF-HIBE system was more efficient than others and we believed that our scheme could be more benefit for the communication devices which have limited computation ability. Moreover, our future work will concentrate on applying the escrow-free HIBE we proposed to practice.

**Acknowledgments.** This work is supported by the National Natural Science Foundation of China under Grants No. 61632009 and the Science and Technology Project of Changsha under Grant No. kq1701089, and Fundamental Research Funds of Central South University under Grant No. 2017zzts711, and the Guang-dong Provincial Natural Science Foundation under Grant 2017A030308006 and High-Level Talents Program of Higher Education in Guangdong Province under Grant 2016ZJ01.

## References

1. Horwitz, J., Lynn, B.: Toward hierarchical identity-based encryption. In: Knudsen, L.R. (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 466–481. Springer, Heidelberg (2002). [https://doi.org/10.1007/3-540-46035-7\\_31](https://doi.org/10.1007/3-540-46035-7_31)
2. Gentry, C., Silverberg, A.: Hierarchical ID-based cryptography. In: Zheng, Y. (ed.) ASIACRYPT 2002. LNCS, vol. 2501, pp. 548–566. Springer, Heidelberg (2002). [https://doi.org/10.1007/3-540-36178-2\\_34](https://doi.org/10.1007/3-540-36178-2_34)
3. Boneh, D., Boyen, X.: Efficient selective-ID secure identity-based encryption without random Oracles. In: Cachin, C., Camenisch, J.L. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 223–238. Springer, Heidelberg (2004). [https://doi.org/10.1007/978-3-540-24676-3\\_14](https://doi.org/10.1007/978-3-540-24676-3_14)
4. Boneh, D., Boyen, X.: Secure identity based encryption without random Oracles. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 443–459. Springer, Heidelberg (2004). [https://doi.org/10.1007/978-3-540-28628-8\\_27](https://doi.org/10.1007/978-3-540-28628-8_27)
5. Chen, L., Harrison, K., Soldera, D., Smart, N.P.: Applications of multiple trust authorities in pairing based cryptosystems. In: Davida, G., Frankel, Y., Rees, O. (eds.) InfraSec 2002. LNCS, vol. 2437, pp. 260–275. Springer, Heidelberg (2002). [https://doi.org/10.1007/3-540-45831-X\\_18](https://doi.org/10.1007/3-540-45831-X_18)

6. Chen, P., et al.: T-HIBE: a trustworthy HIBE scheme for the OSN privacy protection. In: IEEE SocialSec 2015, Liverpool, UK, pp. 72–79, October 2015. <https://doi.org/10.1109/SocialSec2015.11>
7. Gentry, C.: Certificate-based encryption and the certificate revocation problem. In: Biham, E. (ed.) EUROCRYPT 2003. LNCS, vol. 2656, pp. 272–293. Springer, Heidelberg (2003). [https://doi.org/10.1007/3-540-39200-9\\_17](https://doi.org/10.1007/3-540-39200-9_17)
8. Goyal, V.: Reducing trust in the PKG in identity based cryptosystems. In: Menezes, A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 430–447. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-74143-5\\_24](https://doi.org/10.1007/978-3-540-74143-5_24)
9. Boldyreva, A.: Threshold signatures, multisignatures and blind signatures based on the gap-Diffie-Hellman-group signature scheme. In: Desmedt, Y.G. (ed.) PKC 2003. LNCS, vol. 2567, pp. 31–46. Springer, Heidelberg (2003). [https://doi.org/10.1007/3-540-36288-6\\_3](https://doi.org/10.1007/3-540-36288-6_3)
10. Chow, S.S.M.: Removing escrow from identity-based encryption. In: Jarecki, S., Tsudik, G. (eds.) PKC 2009. LNCS, vol. 5443, pp. 256–276. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-00468-1\\_15](https://doi.org/10.1007/978-3-642-00468-1_15)
11. Sui, A.F., et al.: Separable and anonymous identity-based key issuing. In: ICPADS 2005, Fuduoka, Japan, pp. 275–279, July 2005. <https://doi.org/10.1109/ICPADS.2005.263>
12. Chase, M.: Efficient non-interactive zero-knowledge proofs for privacy applications. Ph.D dissertation, Brown University Providence (2008)

# An Improved Pre-copy Transmission Algorithm in Mobile Cloud Computing

Xianfei Huang<sup>1,2(✉)</sup>, Nao Wang<sup>1,2</sup>, and Gaocai Wang<sup>1,2</sup>

<sup>1</sup> School of Computer and Electronic Information, Guangxi University,  
Nanning 530004, Guangxi, People's Republic of China  
huangl023oblivion@outlook.com, 748227@qq.com,  
wanggcgx@163.com

<sup>2</sup> Guangxi Colleges and University Key Laboratory of Parallel and Distributed  
Computing Technology, Nanning 530004, China

**Abstract.** With the rapidly development of mobile cloud computing, in order to provide a better experience for mobile users, the computing and storage functions of mobile terminals are migrated to the cloud. In the application research of virtual machine migration, the transmission rate of wireless channel will be changed with the changes of network condition. This paper focuses on the pre-copy transmission algorithm in the process of virtual machine migration in mobile cloud computing. A pre-copy transmission algorithm based on optimal stopping theory (PTAOST) was proposed, the algorithm constructs an optimal stopping problem based on an optimal transmission rate in the process of virtual machine migration. And we solve the model by using the optimal stopping theory, and get an optimal transfer rate. So the total amount of data and migration time is reduced during the process of virtual machine migration. In simulations, the PTAOST is compared with the transmission algorithms offered by some literatures, and we obtained some comparable results about the total migration data and total migration time to different transmission algorithms. The simulation results show that the proposed algorithm has less the total amount of migration data and total time of migration and improves the performance of migration.

**Keywords:** Mobile cloud computing · Optimal stopping theory  
Virtual machine migration · Optimal transmission rate · Pre-copy

## 1 Introduction

With the rapid development of mobile networks, mobile applications have been given more new functions. Therefore, a higher-performance must be provided by the computing, storage capacity and battery capacity in mobile terminal (MT). In order to meet the requirements of performance, a new application model was created by introducing cloud computing into mobile network. The new application is called mobile cloud computing (MCC). The main function of MCC is migration. In general, MCC can be defined as that MT gets the usage and deliver of necessary infrastructure, platforms and software resources from the cloud in wireless networks [1]. There are three typical migration systems in MCC migration: MAUI, Clonecloud and Cloudlet.

MAUI was old system, and two different computation offloading systems (Cloudlet and Clonecloud) were proposed by researchers for new hardware technology. Cloudlet uses dynamic virtual machine synthesis technology, the applications of virtual machine in the mobile terminals are migrated to the neighbor cloud to achieve the support of calculation. Amendola et al. pointed out that the online migration of virtual machines could provide the support for large numbers of computation and storage in the future of 5G networks [2]. Manaka et al. studied the online migration of virtual machines in order to make end-to-end services having low latency in the data center [3]. In a wireless network, however, it is not enough to satisfy the performance requirements of the delay, and we also need to consider the performance of the transport strategy used in the migration process at wireless network [4, 5]. In addition, to optimize these performances, we must have knowledge about the transport strategies in the virtual machine migration process. There is a typical virtual machine transmission strategies, which is pre-copy migration strategy. In the MCC environment, the characteristics of the transmission rate were needed to take full account at the pre-copy migration strategy in the migration process of wireless network. Since the pre-copy migration strategy is affected by the transmission rate and the dirty page generation rate, when the quality of the wireless network channel isn't good, the transmission rate will reduce the virtual machine migration performance, so channel detection during the virtual machine migration transfer should be considered as an important condition.

An improved pre-copy algorithm in the MCC environment is proposed by this paper. Considering the feature that the capacity and quality of the wireless channel will change with time, the optimal stopping theory is introduced into pre-copy algorithm to obtain the optimal transmission rate of the virtual machine application in the migration process, and further to optimize the total time and total data in the migration performance of the virtual machine.

## 2 Related Work

In recent years, the pre-copy algorithm has been optimized from different aspects as far as the existence problems in the traditional pre-copy algorithm. Two categories can be divided via analyzing the characteristics of memory in the pre-copy algorithm: one is to compress the memory page or to shorten the dirty page of the memory page achieving the minimum amount of data per round [6, 7]. For example, in paper [6], in order to solve the problem of memory page in the pre-copy algorithm, the Markov model was used to predict the generation of the working set, so that the total time and downtime of the migration can be effectively reduced by transmitting dirty pages which were low modification probability. In reference [7], the memory compression algorithm, Lempel-Ziv, was used to compress the memory that needs to be transferred, so that the data which needs to be transmitted was reduced. The other category was optimized by analyzing the effect of dirty page generation on resources (not including memory pages), and using probabilistic and statistical methods [8, 9]. For example, in the literature [8], the generation of dirty pages will have a greater impact on the network retransmission, and the impact will increase the cost of the network. A policy was proposed to optimize the memory, by predicting a modification time of next round.

In the literature [9], the parameter values are given via analyzing the key parameters in the migration (including the dirty page generation rate). The migration strategy of the virtual machine is determined by the model guidance method. Although these algorithms have a certain effect on reducing the total time of migration and the total amount of data transmitted, the researchers did not take into account the impact of the channel quality on the migration process of the virtual network.

The optimal stopping theory has offered a better solution for the optimization of wireless networks. For example, in paper [10], the energy consumption of the transmission process is modeled, then the optimal transmission rate threshold was calculated by using the optimal stopping theory, so the model could effectively reduce the energy consumption. So the optimal stopping theory was introduced into the virtual machine migration process in the wireless network. The optimal stopping theory is a study of when-to-stop, and the stop has the most advantage for decision makers.

In summary, the migration performance of the pre-copy algorithm is focused on. However, in the mobile network, the effective migration performance optimization algorithm also needs to meet the requirements of low total data volume and total time at the same time. In this paper, the migration performance of the pre-copy algorithm is proposed in the mobile network. Considering a change of the transmission rate with the quality of the wireless channel, and the improvement of the performance of the pre-copy algorithm by the change, this paper will propose a pre-copy transmission algorithm based on the optimal stopping theory.

### 3 Problem Description and Model Establishment

#### 3.1 Problem Description

The pre-copy phase is the part that we want to improve and optimize. Assuming the detection time is  $T$ , the number of rounds detected in this round is  $n$ , the number of dirty pages which are transmitted in the previous round is  $D_{i-1}$ , the generation rate of dirty pages is  $C$ , and the wireless channel gain with some varing tailed is  $g$  (usually Rayleigh or Rician are used to simulate changes in the wireless channel). If the rate obtained by detection is  $R_n$ , according to the Shannon formula.

$$R = W \log_2(1 + g * S_{NR}) \quad (1)$$

Then, the channel transmission rate  $R_n$  is generated based on the channel bandwidth  $W$ , the channel gain  $g$ , and the signal-to-noise ratio  $S_{NR}$ . If the next migration is carried out immediately after the end of a migration transfer, the transmission time for each round is  $D_{i-1}/R_s$ , and  $R_s$  is the rate at the time of transmission of the sending device. If the initial time is not ideal, we need to detect the channel, then the detection of the time is  $(D_{i-1} + CnT)/R_n + nT$ . If the initial transmission time  $D_{i-1}/R_s$  is greater than the detected transmission time  $(D_{i-1} + CnT)/R_n + nT$ , then we can transfer the dirty pages after  $n$  times of detection, thus transmission rate can be achieved to reduce the transmission time.

### 3.2 Model Establishment

Because the optimal stopping problem is limited in our study, the maximum number of detections  $M$  should be given. The relationship between the new dirty page rate  $C$  generated by the XEN-based virtual machine and the working set  $Work$  is given by paper [9], there is  $Work = \gamma CN_i T$ , the  $\gamma$  is given in reference [9]. In order to calculate how much income time we can get, we have:

$$\delta = \frac{D_{i-1}}{R_s} - \frac{D_{i-1} + (1 - \gamma)CN_i T}{R_{N_i}} - N_i T \quad (2)$$

Assuming that the maximum transmission rate of the channel is  $R_{max}$ , when  $R_N > R_{max}$  and  $\delta = 0$ , Eq. (2) is equivalent to:

$$R_{max} = \frac{D + (1 - \gamma)CMT}{D/R_s - MT} \quad (3)$$

Assuming that the  $R_N$  is subject to the Rayleigh distribution, then finding optimal transfer rate  $R_N$  can be converted into a special secretary problem with the specific conditions in the optimal stopping theory by Eq. (2). First, we need to give a random variable  $y_k (1 \leq k \leq M)$  to the absolute rank of the  $R_N$  according to the secretary problem. If we want to make the  $k^{\text{th}}$  detection of the  $R_N$  which is the optimal value at  $x_j$ , the  $k^{\text{th}}$  detection  $R_N$  at  $x_j$  probability is:

$$P(y_k = 1 | x_k = j) = \frac{P(y_k = 1, x_k = j)}{P(x_k = j)} = \begin{cases} \frac{k}{M} & j = 1 \\ 0 & j \neq 1 \end{cases} \quad (4)$$

Defining the random variable  $z_k$  for the  $k^{\text{th}}$  detection channel's reward, since  $R_N$  follows the Rayleigh distribution, assuming that the probability density of  $R_N$  is  $f_R(r)$  and the cumulative probability distribution is  $F_R(r)$ . According to Eq. (2), we have:

$$z_k = \int_{R_{th,k}}^{R_{max}} f_R(r) dr \quad (5)$$

Then, the definition of the  $k^{\text{th}}$  detection of the conditions obtained for the reward is  $w_k(j) = E\{z_k = 1 | x_k = j\}$ , combining with Eqs. (5) and (4), we have.

$$w_k(j) = \begin{cases} \frac{k}{M} \int_{R_{th,k}}^{R_{max}} f_R(r) dr & j = 1 \\ 0 & j \neq 1 \end{cases} \quad (6)$$

For  $r = 1, \dots, M$ , there is a stop rule  $\xi(r)$ , and the  $r - 1$  number of detected rate time is smaller than the  $r$  detected rate time. If we don't send data by last  $M - 1$  transmission rate, but by the  $M$  time, and then there is such a stop rule [11]:

$$P(\xi(r) = k) = \frac{r-1}{k(k-1)} \quad (7)$$

According to Eq. (7), the expected reward  $\varphi(r; M)$  which is corresponded to (6) is:

$$\begin{aligned} \varphi(r; M) &= E\{z_{\xi(r)}\} = \sum_{k=r}^M w_{(k)}(1)P(\xi(r) = k) \\ &= \frac{r-1}{M} \sum_{k=r}^M \frac{\int_{R_{M,k}}^{R_{\max}} f_R(r) dr}{k-1} \end{aligned} \quad (8)$$

It is now necessary to find out whether there is a  $r^*$  which let the expected reward  $\varphi(r; M)$  having the maximum value, so that the  $r^*$  is the optimal result of the special secretary problem, which bring with specific conditions. We have the following deduction:

**Deduction 1:** There is a  $r^*(1 \leq r^* \leq M)$  which made the  $\varphi(r; M)$  having a maximum value, according to the expected reward  $\varphi(r; M)$  given by Eq. (2). The optimal stop rule is:

$$r^* = \min \left\{ r \geq 1 \mid \lambda(r; M) = P_{\delta_r} - \sum_{k=r+1}^M \frac{P_{\delta_k}}{k-1} \geq 0 \right\} \quad (9)$$

**Proof:**  $\varphi(r; M)$  is a single-peak function for  $r$ , then there must be  $\varphi(r; M) - \varphi(r+1; M) \geq 0$ .

$$\begin{aligned} \varphi(r; M) - \varphi(r+1; M) &= \frac{1}{N} \left( P_{\delta_r} - \sum_{k=r+1}^M \frac{P_{\delta_k}}{k-1} \right) \geq 0 \\ \Leftrightarrow P_{\delta_r} - \sum_{k=r+1}^M \frac{P_{\delta_k}}{k-1} &\geq 0 \end{aligned}$$

## 4 Pre-copy Transmission Algorithm Based on Optimal Stopping Theory

This section provides a detailed introduction to the pre-copy transmission algorithm based on the optimal stop theory. It is first necessary to obtain the maximum stop time of the transmission algorithm according to the formula (3). Secondly, we need to obtain the optimal stopping time according to the formula (9), and then obtain the optimal transmission rate by the optimal stopping time. After obtaining the optimal



transmission rate, we enter the pre-copy transmission phase. This is an iterative process. The first step, the time of the last round of transmission was calculated; the second step, according to the last round of the transmission time and dirty page generation rate, this round's data size of dirty pages was calculated; the third step, if the transfer data at this round was less than the threshold or greater than the previous round, enter the shutdown state. Finally, all the data of rounds should be transmitted, or return back to the first step. The detailed algorithm is described as follows:

---

**Algorithm 1: Pre-copy transmission algorithm based on optimal stopping theory (PTAOST)**

---

**Input:**  $Vmen, Vthd, RS, C$  /\*The virtual machine data size, the last threshold, the initial transfer rate, the dirty page generation rate\*/

**Output:**  $Vmig, Tmig$  /\*Migrate the total amount of data, migrate the total time\*/

1. **Begin**

2.  $V(1) \leftarrow Vmen$  /\*Initialize the initial transfer data size\*/

3.  $M \leftarrow \min\{N > 1 | R_N \geq R_{max}\}$  /\* Maximum stop time\*/

4.  $r^* \leftarrow \min\{1 \leq r \leq M | \lambda(r; M) \geq 0\}$  /\*Optimal number of detections\*/

5.  $S^* \leftarrow \max\{S_i | 1 \leq i \leq r^*\}$  /\*The maximum rate within the optimal number of detections\*/

6.  $R \leftarrow \text{First}\{S_N | r^* \leq N, S_N \geq S^*\}$  /\*Optimal transmission rate\*/

7. **For**  $1 \leq i \leq \max\_step$  **do** /\*Pre-copy transmission algorithm\*/

8.  $T(i) \leftarrow V(i)/R$

9.  $\gamma \leftarrow a + \omega T(i) + \chi C$

10.  $Work(i+1) \leftarrow \gamma * T(i) * C$

11.  $V(i+1) \leftarrow T(i) * C - Work(i+1)$

12. **if**  $V(i+1) \leq Vthd$  **or**  $V(i+1) > V(i)$

13.  $V(i+1) \leftarrow T(i) * C$  /\* Last transfer data size\*/

14.  $T(i+1) \leftarrow V(i+1)/R$

15.  $Tdown \leftarrow T(i+1) + Tresume$

16. **break**

17. **end if**

18. **end for**

19. **End**

20.  $Vmig \leftarrow \sum_{i=1}^{\max\_step} V(i)$

21.  $Tmig \leftarrow \sum_{i=1}^{\max\_step} T(i)$

22. **Return**  $Vmig, Tmig$

---

According to the analysis of algorithm, the maximum number of steps 3 (the maximum stop time) is  $M$ . Step 4 to step 6 is a stage which is used to find the optimal rate, the maximum number of times which find the optimal rate does not exceed the maximum stop time, so the maximum number of times is also  $M$ . Step 7 to 18 is the pre-copy migration phases, since there is only one time for loop, the maximum number of steps for this loop is  $\max\_step$ , and the 20 and 21 step, the maximum number of times does not exceed  $\max\_step$ . In summary, the time complexity of PTAOST is  $O(n)$ .

## 5 Simulation Results and Analysis

In simulations, the wireless channel model of the experimental environment is assumed to be Rayleigh channel fading model [12]. According to the channel gain probability density given in reference [13] and the Shannon formula (1), we can get the probability distribution formula of the transmission rate as follows:

$$F_R(r) = 1 - \exp\left(-\frac{(2^{r/W} - 1)^2}{2\sigma^2 S_{NR}^2}\right) \quad (10)$$

The parameters in simulations are shown in Table 1.

**Table 1.** Simulation of experimental parameter values

Parameters	Description	Values
$W$	Bandwidth [MHz]	1
$S_{NR}$	Signal-to-noise ratio [dB]	$10^{-1}$
$\sigma$	Mean variance of channel gain	1
$g$	Channel gain	0–4

### 5.1 Performance Comparison and Analysis for Different Algorithms

In order to examine the values of pre-copy transmission algorithm in the virtual machine migration process, Matlab tools is used to simulate our transmission algorithm and two other transmission algorithms. The other transmission algorithms are as follows.

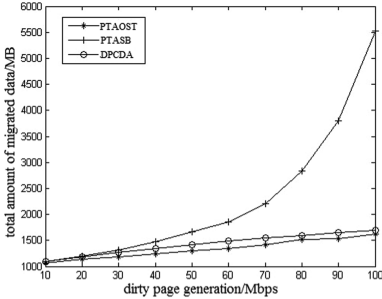
- (1) Pre-copy Transmission Algorithm based on the Sooner the Better (PTASB) [11].
- (2) Dirty Page Custom Delivery Algorithm (DPCDA) [9]. According to the data presented in paper [9], the transmission rate was set to  $R = C + 100$ .

In simulations, the transmission rate  $R$  under different transmission strategies is the key parameter of the average migration data  $V_{mig}$  and the average time of migration  $T_{mig}$ .

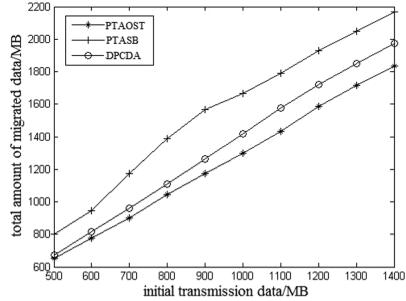
### 5.2 Average Amount of Migrated Data

In simulations, in the conditions of 1G initial transmission data size and the different dirty page generation rate  $C$ , the source host uses the transmission algorithm which was given in Sect. 4 for 10,000 times. Then an average value of the total migration data of the simulation experiment was observed with 10,000 times, the average value  $V_{mig}$  is shown in Fig. 1. The relationship between the total migration data and the dirty page generation rate was described in Fig. 1 during  $T = 0.1$ , and under three different transmission algorithms. It can be seen: (1) The result can be seen from comparing with the PTAOST when  $C = 20\text{--}60$  Mbps, and the total amount of data transmission is more than the DPCDA 100 MB. (2) It can be seen from Fig. 1, whatever the dirty page generation rate is, PTAOST is always better than DPCDA.

Figure 2 show the simulated value  $V_{mig}$  of total migrated data at  $C = 50$  Mbps, and  $V_{mig}$  was under the different initial transfer data and  $T$ . It can be seen that: When the initial transmission of data is too large, PTAOST performance is better than the other two algorithms (DPCDA and PTASB).



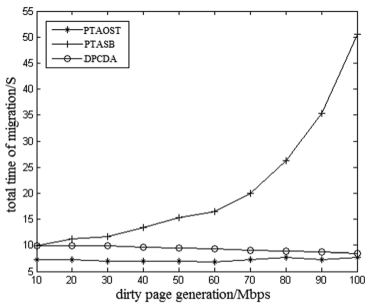
**Fig. 1.** The relationship between the total amount of migrated data and the rate of dirty page generation ( $T = 0.1$ )



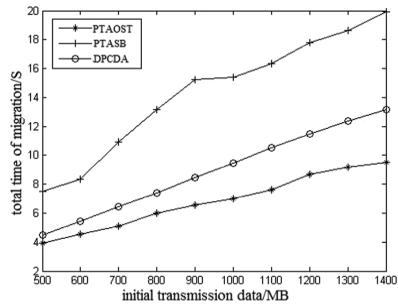
**Fig. 2.** The relationship between the total amount of migrated data and the amount of initial transmission data ( $T = 0.1$ )

### 5.3 Average Migration Time

Figure 3 show the average migration time  $T_{mig}$  of different dirty page generation rates and different detection times  $T$ . As can be seen from Fig. 3: (1) PTASB has a larger change. (2) The PTAOST performs better when the value is small. (3) the total migration time of the PTAOST changes smaller than that of the DPCDA with the increase of the generation rate of dirty pages.



**Fig. 3.** The relationship between the total time of migration and the rate of dirty page generation ( $T = 0.1$ )



**Fig. 4.** The relationship between the total time of migration and the amount of initial transmission data ( $T = 0.1$ )

Figure 4 show the relationship between the total time of migration and the amount of initial transmission data under  $C = 50$  Mbps. the relationship is an incremental relationship. (1) Both of the PTAOST and DPCDA are better than the PTASB and have

a smaller total time of migration. (2) When the initial transmission of data is too large, the total time of migration reduced by PTAOST was more than the other two algorithms.

In summary, the total migration data and the total time of migration can be reduced by pre-copy transmission algorithm, which was based on the optimal stopping theory. The performance improvement for the algorithm in virtual machine migration, and better experience given to users are an achievement of the algorithm proposed by this paper.

## 6 Conclusions

With the rapid growth of mobile networks, improving the performance of mobile terminals, especially the performance of the virtual machine migration process in mobile cloud computing, has become an inevitable requirement. With the performance problem of transmission algorithm in mobile cloud computing, the study on performance optimization of transmission in mobile cloud computing was proposed in this paper. Considering the characteristics of the transmission rate in the quality of the wireless channel, a pre-copy transmission algorithm of the optimal stopping theory is also proposed. In order to migrate the virtual machine applications efficiently, the quality of the wireless channel should be periodically detected by the source host. For finding the optimal transmission rate, the probability distribution of the dirty page production rate and the transmission rate of the wireless channel were used to construct the optimal stopping problem, and then the proof of the optimal stopping problem have been given. Before the virtual machine migration begins, the optimal transfer rate was obtained by the source host in detecting the wireless channel for the migration of virtual machine applications. Then the total migration data and the total time of migration can be reduced by the optimal rate. The experimental results show that the proposed algorithm has less total data and total time, and the performance of virtual machine migration can be effectively improved by our method.

**Acknowledgments.** This research is supported in part by the National Natural Science Foundation of China under Grant No. 61562006, in part by the Natural Science Foundation of Guangxi Province under Grant No. 2016GXNSFBA380181 and in part by the Key Laboratory of Guangxi University.

## References

1. Asrani, P.: Mobile cloud computing. *Int. J. Eng. Adv. Technol.* **2**(4), 606–609 (2013)
2. Amendola, D., Cordeschi, N., Baccarelli, E.: Bandwidth management VMs live migration in wireless fog computing for 5G Networks. In: 2016 5th IEEE International Conference on Cloud Networking, pp. 21–26 (2016)
3. Manaka, Y., Hasegawa, K., Koizumi, Y., et al.: On live migration and routing integration for delay-sensitive cloud services in wireless mesh networks. In: 2015 IEEE International Conference on Communications (ICC), pp. 454–459 (2015)

4. Wen-Li, Z., Bing, G., Yan, S., et al.: Computation offloading on intelligent mobile terminal. *Chin. J. Comput.* **39**(5), 1021–1038 (2016)
5. Yong, C., Jian, S., Cong-Cong, M., et al.: Mobile cloud computing research progress and trends. *Chin. J. Comput.* **40**(2), 273–295 (2017)
6. Sun, G., Gu, J., Hu, J., et al.: Improvement of live memory migration mechanism for virtual machine based on pre-copy. *Comput. Eng.* **37**(13), 36–39 (2011)
7. Mingsong, S., Wenwen, R.: Improvement on dynamic migration technology of virtual machine based on Xen. In: *Proceedings of the 2013 8th International Forum on Strategic Technology*, pp. 124–127. IEEE Press, Washington (2013)
8. Ting-Wei, C., Pu, Z., Zhong-Qing, Z.: Memory optimization algorithm of migration based on Xen virtual machine. *Comput. Sci.* **40**(9), 64–66 (2013)
9. Haikun, L., Hai, J., Chengzhong, X., et al.: Performance and energy modeling for live migration of virtual machines. In: *Proceedings Of the 20th International Symposium on High Performance Distributed Computing (HPDC 2011)*, San Jose, pp. 171–182 (2011)
10. Ying, P., Gao-Cai, W., Shu-Qiang, H., et al.: An energy consumption optimization strategy for data transmission based on optimal stopping theory in mobile networks. *Chin. J. Comput.* **39**(6), 1162–1175 (2016)
11. Anagnostopoulos, C., Hadjiefthymiades, S.: Delay-tolerant delivery of quality information in ad hoc networks. *J. Parallel Distrib. Comput.* **71**(7), 974–987 (2011)
12. Simon, M.K., Alouini, M.S.: *Digital Communications Over Fading Channels*. Wiley, Hoboken (2005)
13. Chen, H., Baras, J.S.: Distributed opportunistic scheduling for wireless Ad-Hoc networks with block-fading mode. *IEEE J. Sel. Areas Commun.* **31**(11), 2324–2337 (2013)

# Motivation of DDOS Attack-Aware Link Assignment between Switches to SDN Controllers

Sameer Ali<sup>1</sup>, Saw Chin Tan<sup>1</sup>(✉), Lee Ching Kwang<sup>2</sup>,  
Zulfadzli Yusoff<sup>2</sup>, Reazul Haque<sup>1</sup>, Ir. Rizaludin Kaspin<sup>3</sup>,  
and Salvatore Renato Ziri<sup>3</sup>

<sup>1</sup> Faculty of Computing and Informatics, Multimedia University,  
63100 Cyberjaya, Malaysia  
sameer.ali@szabist.edu.pk, sctan1@mmu.edu.my,  
reazul.mmu@gmail.com

<sup>2</sup> Faculty of Engineering, Multimedia University, 63100 Cyberjaya, Malaysia  
{cklee, zulfadzli}@mmu.edu.my

<sup>3</sup> Telekom Malaysia Research and Development, TM Innovation Centre,  
Lingkaran Technokrat, Cyberjaya, Malaysia  
{Rizaludin, salvatore}@tmrnd.com.my

**Abstract.** In this paper, we are going to investigate the attack-aware link assignment for the connectivity of switches to SDN controllers under DDoS attacks. The infected or broken link of switch to controller will create disconnection in network and interruption in services. As a result, such as reliability and availability will be impaired. Therefore, on the basis of the analysis, we are going to highlight the importance of the attack-aware link-assignment for SDN switches to controllers.

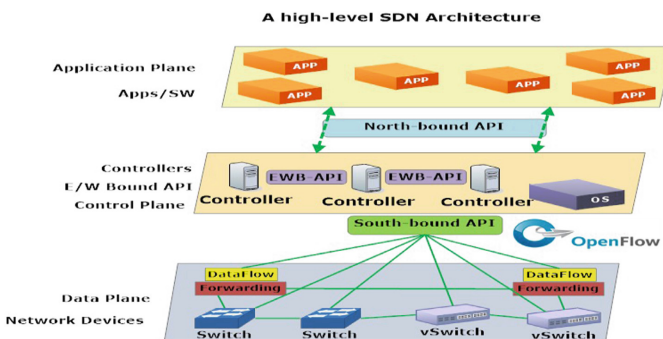
**Keywords:** Software-defined networks · DDoS attack  
Attack-aware link assignment · Propagation model

## 1 Introduction

Software-defined networking (SDN) is a collection of numerous network technologies which are agile and flexible as the virtualized servers and storage devices are concerned of the current data center networks. The architecture of software-defined network allows the network engineers and administrators to rapidly implement policies and procedures as per network requirement [1]. SDN networks are dynamic as compared to the traditional networks which are static and inflexible. In traditional networks, the control plane and data plane are connected and operated together. However, in SDN the control plane and data plane are separated which gives the flexibility of programming and centralize management of networks. Controllers perform as the core part of the SDN network. SDN Controllers offer a central view of the complete network. These controllers enable network engineers to make decision and give command to the control plane such that switches and routers and can be operated more effectively. SDN

uses southbound Application Programming Interface (API) to transfer the data to the switches and routers [2]. SDN uses northbound API to communicate between the requests and business applications to the control plane. This enables network administrators to programmatically manage the flow of traffic and organize different services directly. Alternatively, there are many other networks such as social media, smart phones, cloud clustering. These modern technologies are persistent to reduce the legacy networks to their limits. Computer systems and other storage devices have been benefited from unbelievable novelties in virtualization and automation. Network administrators may turn up new computer systems and storage device examples in instantly, and only can be detained up for weeks by inflexible and often time taking physical network operations.

SDN is one of the predominant networking paradigms that seeks to simplify network control logic from the underlying hardware and introduces real-time network programmability enabling innovation. SDN has the impending to transform traditional data centers by providing an elastic way to control over the network so it can operate more like the virtualized forms of computers and storage devices [2, 3, 4]. Figure 1 illustrates a high-level SDN architecture with its three layers i.e. data plane, control plane and application plane. In this figure switches and virtual switches are located in the data plane. All these are connected with the control plane via southbound API which provides communication interface through OpenFlow protocol. The control plane contains controller software where network operating system resides. Application layer is the tier three in SDN which contains end user business applications and security tools.



**Fig. 1.** Software-defined networking a high level architecture.

A Distributed Denial of Service (DDoS) attack is an effort to make an online service inaccessible. DDoS attack may target a wide range of significant network resources, from financial organizations to newscast websites and services [3, 4, 5]. The victim of a DDoS attack can be of both the end-user targeted systems and all network systems compromised by malicious traffic. In a DDoS attack, the arriving traffic

overflows the victim system from many different sources like potentially more than thousands. This effectively makes it difficult to stop the attack simply by delaying only an IP address. It is actually difficult to differentiate authentic user traffic from attacking traffic when dealing with so numerous points of sources [5, 2].

DDoS attacks can be of either vertically or horizontally [18]. The DDoS attacker may attack using a huge number of malicious traffic. The propagation of DDoS attacks are as illustrated in Figs. 2 and 3.

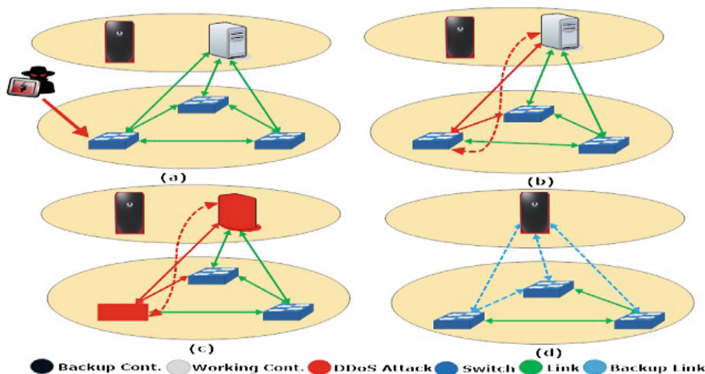


Fig. 2. Illustrates the vertical DDoS attack in SDN [18].

In case of a vertical attack in SDN network the infected switch may be harmful for overall network component. Shown in (Fig. 2a) where the attacker is attacking on the switch directly. When the infected switch communicates with the controller in (Fig. 2b) then the controller also become infected. Subsequently, this infection continues spreading in (Fig. 2c) where the controller becomes disabled. In Fig. 2(d) shows the backup links and backup controller with the restored network operations. In (Fig. 3a) which illustrates that DDoS attacker has attacked on a switch where other switches are also connected horizontally to the first one therefore due to this attack the infection gets spread horizontally to other neighbors. In (Fig. 3b) it is evident that infection is passing to the controller. This spreading to (Fig. 3c) which results the controller become disabled. In (Fig. 3d) shows that the backup controller restored the network services through provided backup controller and links.

The link assignment in SDN is the establishment of a link between switches to SDN controllers in both directions. DDoS attacks are vulnerable which result of network unavailability and interruption in services. We are investigating in attack-aware link assignment to highlight its importance in SDN.

Section 2 describes the related work of DDoS attacks under SDN. Section 3 describes the investigation and analysis of attack-aware link-assignment and finally Sect. 4 concludes with the suggestions for future work.



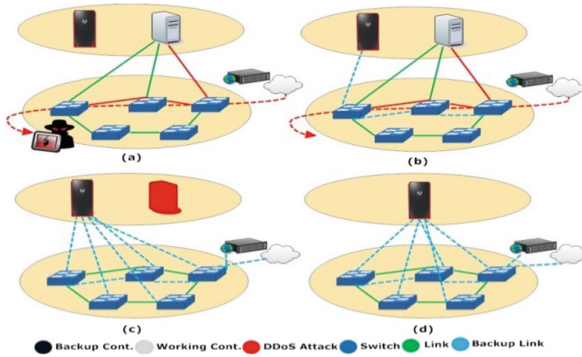


Fig. 3. Illustrates the horizontal DDoS attack in SDN [18].

## 2 Related Work

Comprehensive study of the literature review on software-defined networks and DDoS attacks in SDN is discussed here. The concept of DDoS attacks are getting more dominant especially in cloud computing areas. Currently, software-defined based networks have attracted many researchers in networking environment. The authors in paper [6] discuss about the new challenges and characteristics in cloud computing about the DDoS attacks by providing a broad assessment against mechanism of defense to DDoS attacks in SDN networks. Article [7] describes that the DDoS attacks are the reason to ruin the network resources by flooding the harmful packets through the compromised systems. Attacker aims is to make the network traffic malicious. Paper [8] highlights the network security problems and proposed a method against DDoS attacks in SDN that whatever the data packet is being sent to controller by the switch should be thoroughly examined by the controller using a new specific entry with hard timeout and idle timeout of which the values can be smaller as compare to normal flow entries. The author's paper [9] has developed the methods to detect the DDoS attacks on software-defined network's capability of low monitoring. Their simulation shows that the proposed methods instantly can locate the victims of DDoS attacks with using the numbers of constrained of flow monitoring rules. In paper [10] authors discuss the new features and developments in cloud computing environment about the DDoS attack. Their research suggested that the DDoS attacks in the area of cloud computing can be reduced by the use of SDN. In [11] the authors reported on security impacts, specifically the impacts of DDoS attacks and its defense mechanism in a large enterprise network. Their perspective is that the SDN technology is very useful for enterprise networks to defend from DDoS attack if the mechanism of defense is designed adequately.

The authors in [12] developed several kinds of controller placement algorithms for SDN. These algorithms are also useful to maximize SDN reliability because any network failure may cause the disconnection between the control plane and data planes. Their algorithms were evaluated by using actual practical topologies. They define the path of control as the route set that is used for communication between switches and the

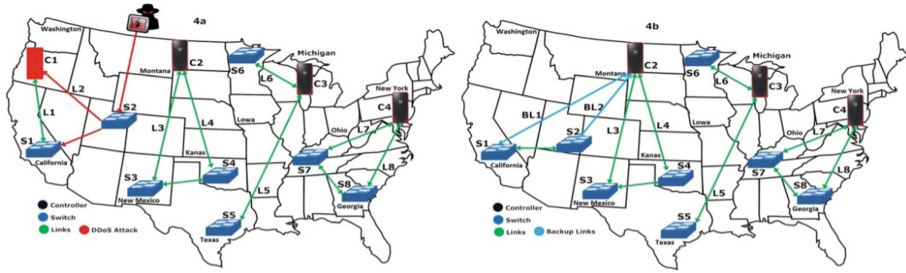
controllers. In [13] paper the authors investigated on partition of wide-area network into different small units of SDN in view of the placement of controller in each SDN unit and the location of the controller to be placed. They developed an approach for accurate and efficient evaluation of SDN controller's placement for wide-area networks using the Spectral Clustering controller placement algorithm. The Heller et al. early work [14] researched on controllers placement problem and then measuring the impact of controller placement on real topologies such that Internet2 and then several cases taken from the Internet Topology Zoo. In [15] the authors developed three heuristic algorithms namely primal-dual-based algorithm, network-partition-based algorithm and incremental greedy algorithm. They introduce the QoS Guaranteed controller placement as a criterion in a given network topology with a response time-bound to determine the number of controllers required subsequently, the suitable location for placing the controllers and which switch should be assigned to each controller are identified. Paper [16], provides a broad synopsis of SDN multi-controller architectures. It presents SDN and its main instantiation OpenFlow. The paper compares differences between two-types of multi-controller architectures, such that the distribution method and the communication system. The authors in [17] discussed and formulated the problems related to controller placement in multi-domain networks on the basis of network partitioning in their previous work. In the next section will provide a close understanding of the scenario when SDN based networks are under DDoS attack. Most of the work reported so far is on SDN controller placement strategies. Consideration of an attack-aware link assignment has yet to be reported. We have analyzed and identified the attack-aware link assignment and it's important in next section in detailed context.

### 3 Investigation of Link-Assignment

In this section we further elaborate on our investigation on attack-aware link-assignment in SDN networks under DDoS attack as shown in Fig. 4 the assumed data centers of the USA has been deployed at four different locations as region wise in USA. The first region of SDN data center is based at Washington with one controller and two switches. The second one is based at Montana region with one controller and two switches. The third one is at Michigan region with one controller and two switches. The fourth one is at New York region with one controller and two switches.

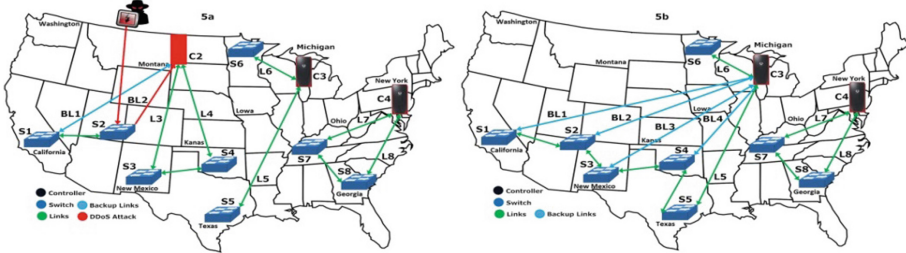
Figure 4a represents that due to DDoS traffic flooding attack controller C1 at Washington region become disabled. On the other hand the switch S1 and S2 activates its backup links BL1 and BL2 to the controller C2 at Montana region and network services restored its functions. In Fig. 4b which shows that again DDoS attacker has attacked on same switch S2. The switch S2 propagating the infection to the controller C2 and switch S1.

In Fig. 5a which represents that due to DDoS attack the controller C2 also become disabled while switch S1, S2, S3 and S4 activate its backup links BL1, BL2, BL3 and BL4 to the controller C3 at Michigan region. Provided backup links the SDN network services restored again. Figure 5b shows the scenario after two controllers C1 and C2 become disabled in Figs. 4 and 5 respectively. The DDoS attacker again makes an



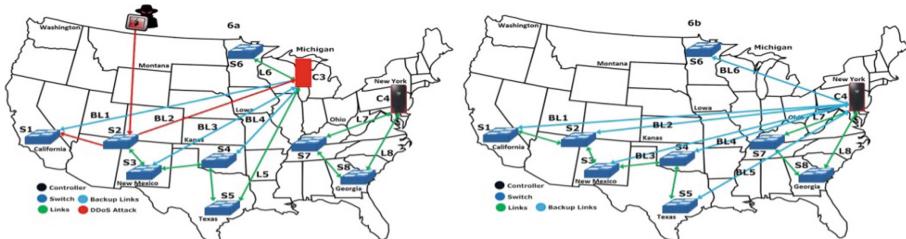
**Fig. 4.** (a) and (b) Represents the DDoS attack on switch S2 at Washington region and backup links activation to the controller C2 respectively.

attempt to attack the switch S2 at Michigan region. This attack is the reason to further spread the bug to neighbors because switch S2 is connected to controller C3 and switch S1 directly.



**Fig. 5.** (a) and (b) Represents the DDoS attack on same switch S2 at Montana and the backup links to the controller C3.

In Fig. 6a we can see that due to DDoS flooding attack three controllers C1, C2 and C3 become disabled. As switch S2 was under DDoS attack but provided backup links it again established its backup links to the controller C4 at Washington region in Fig. 6b. Here we notice that all switches are dependent on controller C4. Here provided backup links again restore network services and functions.



**Fig. 6.** (a) and (b) Shows the DDoS attack on switch S2 at Michigan region and the backup links activation to controller C4.

Table 1 shows the summary of our investigation in terms of SDN network performance and cost effectiveness. In scenario 4a controller C1 is connected to two switches S1 and S2. The cost of link is low because of less fiber used and controller is at nearby region. In scenario 5a controller C2 is connected to four switches S1, S2, S3 and S4 where link cost is medium. Whereas in scenario 6a controller C3 is connected to six switches S1, S2, S3, S4, S5 and S6. Here cost of link is high because of controller C3 is located in other region. In 6b controller C4 is connected to eight switches S1 – S8 here cost of link is also high due to long distance of controller C4. For future work we suggest to propose an attack-ware link assignment approach that will result in optimizing the cost of links and performance of SDN network under DDoS attack.

**Table 1.** Different cost of solutions under DDoS attack in SDN

Controllers	Switches	Location	Cost of link
C1-C2	S1 – S4	Scenario-1	Low
C2-C3	S1 – S6	Scenario-2	Medium
C3-C4	S1 – S8	Scenario-3	High

## 4 Conclusion

We have analyzed different existing SDN research works in the context of DDoS attacks. It is necessary to have smooth network operations to prevent or reduce the impacts of DDoS attacks. In this paper, we have investigated the attack-aware link assignment for the connectivity of switches to SDN controllers under DDoS attacks. The infected or broken link of switch to controller will create disconnection in network and interruption in services. As a result, such as reliability and availability will be impaired. Therefore, on the basis of our analysis report we have highlighted the importance of the attack-aware link-assignment for SDN switches to controllers. For our future research, we will concentrate on the approach to address attack-aware link-assignment issues to reduce the DDoS attacks propagation in SDN based networks and to minimize the cost of links.

**Acknowledgments.** This research work is fully supported by the research grant of TM R&D and Multimedia University, Cyberjaya, Malaysia. We are very thankful to the team of TM R&D and Multimedia University for providing the support to our research studies.

## References

1. Kreutz, D., Ramos, F.M., Verissimo, P.E., Rothenberg, C.E., Azodolmolky, S., Uhlig, S.: Software-defined networking: a comprehensive survey. *Proc. IEEE* **103**(1), 14–76 (2015)
2. Jammal, M., Singh, T., Shami, A., Asal, R., Li, Y.: Software defined networking: state of the art and research challenges. *Comput. Netw.* **72**, 74–98 (2014). Elsevier
3. Xia, W., Wen, Y., Foh, C.H., Niyato, D., Xie, H.: A survey on software-defined networking. *IEEE Commun. Surv. Tutorials* **17**(1), 27–51 (2015)

4. Bakhshi, T.: State of the art and recent research advances in software defined networking. *Wirel. Commun. Mob. Comput.* (2017)
5. Heller, B., Sherwood, R., McKeown, N.: The controller placement problem. In: Proceedings of the 1st Workshop on Hot Topics in Software Defined Networks, pp. 7–12. ACM, August 2012
6. Yan, Q., Yu, F.R., Gong, Q., Li, J.: Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: a survey, some research issues, and challenges. *IEEE Commun. Surv. Tutorials* **18**(1), 602–622 (2016)
7. Mallikarjunan, K.N., Muthupriya, K., Shalinie, S.M.: A survey of distributed denial of service attack. In: 2016 10th International Conference on Intelligent Systems and Control (ISCO), pp. 1–6. IEEE, January 2016
8. Dao, N.N., Park, J., Park, M., Cho, S.: A feasible method to combat against DDoS attack in SDN network. In: 2015 International Conference on Information Networking (ICOIN), pp. 309–311. IEEE, January 2015
9. Xu, Y., Liu, Y.: DDoS attack detection under SDN context. In: IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, pp. 1–9. IEEE, April 2016
10. Yan, Q., Yu, F.R.: Distributed denial of service attacks in software-defined networking with cloud computing. *IEEE Commun. Mag.* **53**(4), 52–59 (2015)
11. Wang, B., Zheng, Y., Lou, W., Hou, Y.T.: DDoS attack protection in the era of cloud computing and software-defined networking. *Comput. Netw.* **81**, 308–319 (2015)
12. Yan-nan, H., Wen-dong, W., Xiang-yang, G., Xi-rong, Q., Shi-duan, C
13. Xiao, P., Qu, W., Qi, H., Li, Z., Xu, Y.: The SDN controller placement problem for WAN. In: 2014 IEEE/CIC International Conference on Communications in China (ICCC), pp. 220–224. IEEE, October 2014
14. Borcoci, E., Badea, R., Obreja, S.G., Vochin, M.: On multi-controller placement optimization in software defined networking-based wans. *ICN* **2015**, 273 (2015)
15. Cheng, T.Y., Wang, M., Jia, X.: QoS-guaranteed controller placement in SDN. In: 2015 IEEE Global Communications Conference (GLOBECOM), pp. 1–6. IEEE, December 2015
16. Bliat, O., Ben Mamoun, M., Benaini, R.: An overview on SDN architectures with multiple controllers. *J. Comput. Netw. Commun.* (2016)
17. Aoki, H., Shinomiya, N.: Controller placement problem to enhance performance in multi-domain SDN networks. *ICN* **2016**, 120 (2016)
18. Manzano, M., Fagertun, A.M., Ruepp, S., Calle, E., Scoglio, C., Sydney, A., de la Oliva, A., Muñoz, A.: Unveiling Potential Failure Propagation Scenarios in Core Transport Networks (2014). arXiv preprint [arXiv:1402.2680](https://arxiv.org/abs/1402.2680)

# TIM: A Trust Insurance Mechanism for Network Function Virtualization Based on Trusted Computing

Guangwu Xu<sup>1</sup>, Yankun Tang<sup>1</sup>, Zheng Yan<sup>1,2(✉)</sup>, and Peng Zhang<sup>1</sup>

<sup>1</sup> The State Key Lab of Integrated Services Networks,  
School of Cyber Engineering, Xidian University, Xi'an 710071, China  
1328840210@qq.com, zyan@xidian.edu.cn,  
pengzhangzhang@gmail.com

<sup>2</sup> Department of Communications and Networking,  
Aalto University, 02150 Espoo, Finland

**Abstract.** As a new network architecture and key technology of 5G, Network Function Virtualization (NFV) is paid special attention in both industry and academia. In the context of NFV, Virtualized Network Functions (VNFs) located in different network computing platforms need collaborate to fulfill an expected networking service. Platform trust authentication and trusted collaboration should be highly ensured in order to support secure and trustworthy 5G mobile networks and wireless systems. However, such a mechanism is still missed in the literature. In this paper, we propose and develop a trust insurance mechanism, named TIM, based on Trusted Computing Platform (TCP) to realize trust authentication and achieve trusted collaboration among VNFs. Performance evaluation shows the effectiveness of TIM with regard to trust attestation accuracy, operation efficiency, and reasonable resource consumption.

**Keywords:** Virtualized Network Function · Trusted computing  
Trust management · Network Function Virtualization

## 1 Introduction

With the fast growth of mobile communications, 4G mobile networks have obtained wide acceptance and popularity with constant increase of mobile users. However, it is still difficult for network operators to upgrade network components in order to meet user expectation on high quality of networking services. Traditional network architecture is obviously unable to meet the challenges caused by the demands on large network bandwidth, flexible network function deployment and efficient network equipment update, under the circumstance of big data.

With the advantages of flexibility, lightweight, openness scalability and sustainability, Network Function Virtualization (NFV) and Software Defined Networking (SDN) can fundamentally solve the aforementioned problems of the traditional networks. They have become two of the key technologies in the next generation of mobile networks and wireless systems (i.e., 5G). NFV provides new opportunities to enable dynamic and flexible network function deployment, configuration, and management. It

brings great convenience for network architecture reconstruction. NFV makes use of virtualization technologies on general x86 servers to provide a platform and environment for deployment of Virtualized Network Functions (VNFs) and network element software. It abstracts hardware resources into a unified virtualized resource pool and allocates them for upper VNFs according to practical needs, so as to decouple network equipment and software. As shown in Fig. 1, the VNF layer above the virtualization layer can run various software-based VNFs. NFV infrastructure (NFVI) refers to the hardware and software components that build up an environment in which VNFs can be deployed and executed. An orchestrator is responsible for the management of NFV infrastructure and software resources. A VNF manager takes charge of instantiation, update, query, scaling of VNFs. Operation Support System and Business Support System (OSS/BSS) are used by operators to manage and operate the networks.

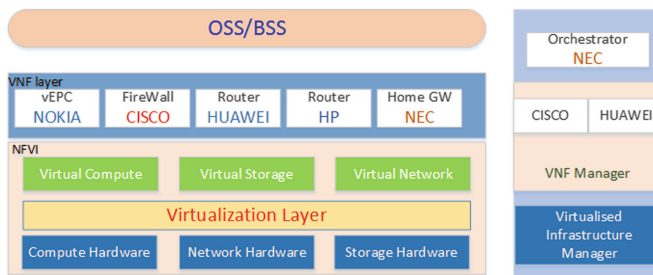


Fig. 1. NFV reference architecture

## 1.1 Motivations

Although significant advantages, NFV may bring with serious security issues, such as secure network service provision, trusted collaboration between platforms, and seamless and secure heterogeneous network resource integration. In the infrastructure of NFV, various network functions located in different platforms need to work together to fulfill an entire networking service in a customized way. However, one open issue is how to authenticate the trustworthiness of other platforms and establish trusted collaboration during the fulfillment of a networking service among multiple virtualized computing platforms. Obviously, different VNFs could be deployed in different platforms located in different positions of the networks. This means that an entire customized networking service should be fulfilled through the cooperation of different VNFs in different platforms, in which each platform could provide one or more functions. Different network functions in different platforms should authenticate with others, keep a trust relationship and collaborate to fulfill a common goal in a trustworthy way. However, existing schemes either focus on single-platform security or cannot support conditional trust among multiple platforms [7–12]. The literature still lacks such a secure and effective mechanism that is specially designed for NFV to ensure cross-platform conditional and sustainable trust and supports trusted collaboration among multiple platforms. How to ensure the trust relationships among VNF platforms during networking service fulfillment is still an open issue.

## 1.2 Main Contributions

In this paper, we propose a trust insurance mechanism, named TIM, for NFV by applying Trusted Computing Platform (TCP) technology [11, 13]. We make use of the advantages of Trusted Platform Module (TPM) by embedding a Root Trust Module (RTM) implemented by TPM into the NFV platform [12]. Through platform configuration challenge, a NFV platform can authenticate the trust status of another platform and initialize a platform-to-platform trust relationship. By embedding a trust insurance policy into the RTM of the challenged platform and performing runtime monitoring on its platform configuration changes, as well as controlling platform changes based on the embedded policy, trust insurance across different platforms can be achieved. With this proposed mechanism, one NFV platform can verify the trustworthiness of a remote platform and keep its trust in it during platform collaboration, thus realize trust management among VNFs. We implemented the proposed mechanism to validate its performance. The performance evaluation results show the effectiveness of TIM with regard to trust attestation accuracy, operation efficiency, and reasonable resource consumption.

The proposed trust insurance mechanism differs substantially from the existing work. It achieves trust authentication and trust sustainment among NFV platforms based on specified conditions. It extends initial trust built among NFV platforms during the fulfillment of an expected networking service, thus achieves on-going trust. Specifically, the main contributions of this paper can be summarized as below:

- We propose a trust insurance mechanism based on TCP technology to realize trustworthy collaboration among NFV platforms in a VNF service chain. It contains a detailed protocol for trust authentication and trust sustainment between NFV platforms.
- We implement the mechanism in a virtualized computing environment and conduct thorough performance evaluation to show its effectiveness.

The rest of the paper is organized as follows. Section 2 gives a brief overview of related work. Section 3 describes the details of the proposed mechanism. In Sect. 4, we introduce TIM implementation, followed by security analysis and performance evaluation in Sect. 5. Finally, a conclusion is presented in the last section.

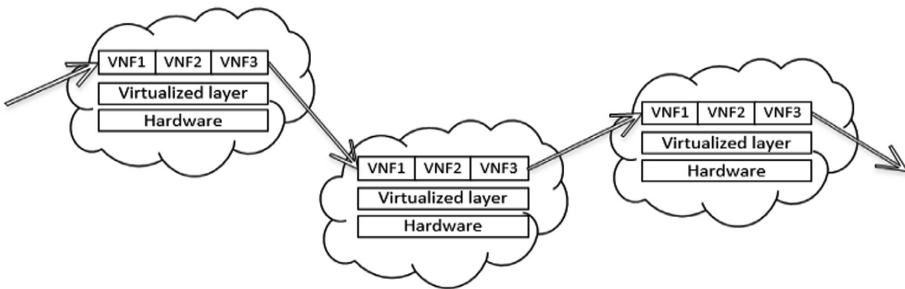
## 2 Related Work

### 2.1 Security Issues of NFV

As a key technology of 5G, NFV has been adopted by most network operators and manufacturers, such as Ericsson, Nokia and Huawei, ZTE, H3C and Fiberhom. However, NFV also brings new vulnerabilities along with its benefits. In the context of NFV, hardware resources are virtualized in a resource pool. All the resources could be rescaled and allocated to VNFs in a flexible and facile way. As a result, NFV can be customized according to corresponding practical demands, so that all the network components could be organized and accommodated efficiently and flexibly.



Various VNFs located at different platforms need work together to form a network service chain for fulfilling a networking service, as shown in Fig. 2. The provision of VNFs should be based on secure interaction and trusted collaboration. Incomplete, malicious or inconsistent configurations of NFV platforms in the service chain could cause breakdown of whole NFV supporting infrastructure. Moreover, automation of NFV process with integration of other technologies may lead the networking service to be error-prone. Note that those software-based VNFs are possibly provided by different vendors, thus they may create security holes due to integration complexity. In the case of executing a VNF in a virtual machine whose physical resources are shared with other network functions, security and trust threats should be resisted from its underlying infrastructure. All above potential security problems highly requests a trust insurance mechanism when building a networking system based on VNFs.



**Fig. 2.** A VNF-based service chain

We classify the security vulnerabilities and potential attacks in NFV into two categories and summarize them as below based on the statements pointed out in [1–5].

**Intra-Platform Attacks.** In the domain of VNFs, attackers may be able to gain unauthorized access to the NFV platform by taking the advantage of vulnerabilities in software-based VNFs, which could compromise the security of the whole platform. In the domain of NFVI, abuse of hypervisor and hardware resources by a malicious virtual machine could impact the security and QoS of other virtual machines. The whole platform could become the target of attacks. In the virtualized layer, following the principle of open-source, Application Programming Interfaces (APIs) are used to provide programmable orchestration for VNFs. These APIs could introduce serious security threats to NFV [6]. Therefore, it is essential to ensure the security of NFVI.

**Inter-Platform Attacks.** Inter-platform attack refers to the attack from one malicious platform towards another NFV platform in the service chain, e.g. the information of a competitor’s platform is extracted and misused to corrupt its services. So far, there is no any existing mechanism to validate the trustworthiness of every network device in the service chain to prevent this kind of security threats [3].

## 2.2 Security Schemes of NFV

Focusing on the aforementioned security issues, a number of solutions were proposed, however with some drawbacks. In [4], a framework of verification on NFV-enabled network services was proposed to provide a formal method to build and verify NFV components. A secure orchestrator was proposed in [8] to manage physical networks and VNFs in a multi-vendor and multi-tenant NFV environment. But they only focus on the secure management of VNFs in a single platform and cannot fundamentally solve the issue about trust authentication and collaboration across multiple NFV platforms. In [4, 5, 7, 9], solutions were proposed to manage the service chain and the process of booting a protected virtual machine. But the literature still lacks an effective scheme to authenticate the trust of VNFs and ensure its trustworthiness during VNF collaboration, especially among VNFs located in different NFV platforms. In [7], a scheme for privacy preservation in a VNF-enabled collaborative service delivery architecture was proposed, but it cannot fundamentally ensure the security of a NFV platform and multi-platform collaboration. Mathieu et al. proposed a scheme to monitor security components in NFV, but it cannot support trustworthy collaboration across platforms in a sustainable way [15]. In [9], a particular ETSI-NFV inter-DC (Data Center) architecture was proposed in a compatible optical network test-bed to provide enhanced security capabilities for VNF distribution across Data-Centers. But this scheme cannot achieve step-by-step verification and ensure conditional trust as our scheme.

Some existing studies explored security across multiple NFV platforms. A scalable remote attestation scheme applying TPM is proposed for NFV [14]. But its drawback is it cannot ensure conditional trust relationships and sustain them according to the policy of trustors. In [5] a scheme was proposed for integrating SDN and NFV in Openstack cloud to minimize network attacks and improve network service quality. A tenant-attested trusted cloud service framework was proposed to verify the security status of a remote cloud platform and its software at runtime [12]. Although it supports attestation across platforms, it cannot support conditional and sustainable trust relationship.

In [1], Yan et al. proposed a security framework to protect NFV based on trusted computing, which can solve both intra- and inter-platform security problems. Yan and Cofta presented a mechanism in [12] for sustaining trust among computing platforms. The mechanism can automatically inform the trustor about any distrusted behaviors of a trustee platform according to predefined conditions and trust policies. It can filter out the attacks that aim to compromise the platform or reject malicious codes. However, neither of above works conducted performance evaluation to show effectiveness.

## 3 Trust Insurance Mechanism

In this section, we propose the TIM to grab and monitor platform configurations by applying RTM to ensure security and trust among NFV platforms. In the mechanism, one NFV platform can verify and attest a remote cooperating platform at the run-time of VNFs by monitoring its configurations. By applying this mechanism, the remote platform can work as the expectation of the local platform.

### 3.1 Notations, Definitions and Abbreviations

**Platform Configuration Register (PCR):** PCR stores platform configuration information and records integrity measurement results. The PCR can be only reset or extended, so that malicious code cannot arbitrarily tamper it.

**Attestation Identity Key (AIK):** AIK is used to digitally sign the integrity reports and authenticate the values of PCRs. The integrity reports are used to determine the current configuration state of a platform. Through Direct Anonymous Attestation (DAA) or a trusted third party, anonymity and identity privacy preservation could be achieved according to TCP technologies [13].

**Root Trust Module (RTM):** RTM is a trusted module to measure the integrity of a platform and ensure secure data storage. It can be applied to monitor platform configuration and conduct platform measurement. In TIM, TPM and its supporting software are adopted to play the role of RTM.

### 3.2 Trust Chain Based on TCP

TCP aims to ensure the trustworthiness of a computing platform through authenticated booting and authenticated loading of other software components of a computing platform, e.g., BIOS, OS loader, OS kernel, legitimate software, and third party software. The TCP hardware keeps a tamper-evident log of the booting process, using a cryptographic hash function to detect any changes with the log. That is a loader calculates the hash code of the next software contributor and logs it in the PCRs. If the value derived from the log is same as the trusted value provided by the TPM, the check is passed. With this way, a trust chain is established inside the platform based on TPM [12] through verification one by one in a sequential order of booting from hardware level to software level. TIM is designed based on the above theory by applying TPM as RTM, which is embedded in the NFV platform as an independent module. The RTM can measure the integrity of the platform and test its changes through platform configuration monitoring. What is more, the TCP hardware can make the configuration known to others, thus realize the trust attestation on a remote platform by digitally certifying platform configurations on OS and its installed software.

However, an obvious shortcoming of TCP is it forces users to accept a preset strategy defined by a device manufacture, rather than a personalized strategy according to practical preferences and demands. This kind of ‘blind trust’ is one of the biggest barriers that hinder its wide usage and acceptance. Moreover, it cannot dynamically ensure the trust relationship among platforms in case that there are some configuration changes on a remote platform, thus shows no way to dynamically confront trust relationship broken.

### 3.3 Trust Insurance Protocol

In order to overcome the above problem, we propose a trust insurance protocol in TIM, as shown in Fig. 3. Platform A needs collaboration with Platform B that correspondingly

challenges A’s configuration for the purpose of trusted collaboration. Before B establishes a trust relationship with A for future cooperation, it should verify the security status of A in case of any intrusions and attacks happened on A. The RTM embedded in Platform A can measure its PCRs and report to B to prove its security. After trust attestation and validation, Platform B can embed the conditions of trusted collaboration into the RTM of A. The same way is applied for A if B wants to collaborate with A, thus mutual trust can be established. Then, they can work together in a service chain to complete an entire networking service. The trust relationship is controlled through the conditions defined by a trustor platform, which are executed by the RTM at a trustee platform. The RTM (implemented by TPM) at the trustee can be verified by the trustor as its expectation for some intended purpose and cannot be compromised by the trustee or other malicious entities. We hold such an assumption on the basis of sound security ensured by the TPM.

TIM applies the following trust formula: “Entity B trusts entity A under condition C based on RTM for purpose P”. The difference of this formula from others lies in the condition C, which is introduced to ensure a trust relationship from its initialization to the fulfillment of P from B to A. C is defined by challenger B to specify the rules and policies for sustaining the trust of B in A for purpose P. The condition can restrict any changes or distrust behaviors at A that may impact the initialized trust relationship. Similarly, A can embed its trust condition into B’s RTM at the same time to set up mutual trust insurance. With the help of the proposed mechanism, the trust relationship can be ensured with the support of RTM and sustained automatically based on pre-set conditions during the fulfillment of P. Note that additional attestation could be performed in order to embed upgraded conditions into a remote platform.

The proposed protocol is shown in Fig. 3 and described in details below:

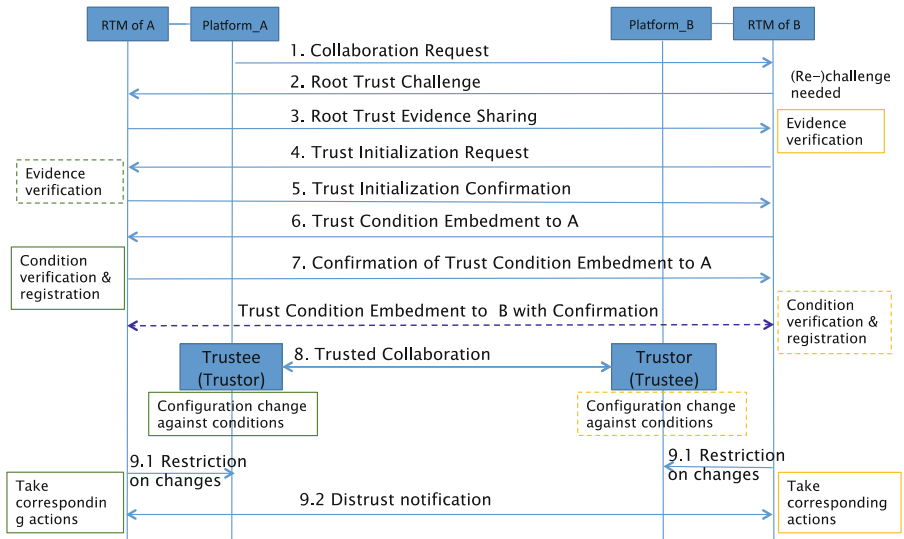


Fig. 3. Trust insurance protocol

- (1) **Collaboration Request:** Platform A requests to collaborate with Platform B for the purpose of fulfilling a networking service.
- (2) **Root Trust Challenge:** Platform B sends a non-predictable nonce to challenge Platform A by requesting its trust evidence to verify its identity, corresponding platform integrity measurements, and PCR values for trust initialization.
- (3) **Root Trust Evidence Sharing:** Platform A provides its platform security proof to Platform B by responding A's challenge after receiving the nonce. Concretely, Platform A retrieves the RTM to get its PCR value, gains platform measurement values and sends them to B together with the nonce and its signature.
- (4) **Trust Initialization Request:** In this step, Platform B verifies whether the received trust evidence of Platform A is as expected. If the verification is positive, Platform B would consider A as trusted and request to establish a trust relationship with it (by providing root trust evidence). Otherwise, a trust relationship cannot be initialized due to the distrusted platform configurations of Platform A.
- (5) **Trust Initialization Confirmation:** In this step, Platform A responds and confirms to initialize the trust relationship with B by challenging and verifying the trust evidence of Platform B if needed for establishing mutual trust.
- (6) **Trust Condition Embedment:** To ensure on-going trust after trust initialization, Platform B formulates trust conditions for trust sustainment in the future. The conditions specify the allowed actions and behaviors of Platform A with regard to ensuring B's trust in A. They are the premise of trust. Platform B informs A about the conditions that are signed with its AIK. The conditions are embedded into A's RTM to instruct its platform operation. If no any trust conditions are requested by B, B's trust in A will be "blind trust". That is, once the trust relationship is established, all operations performed at Platform A are treated acceptable and trusted by B.
- (7) **Confirmation of Trust Condition Embedment:** In this step, A informs B that it receives the conditions and agrees to follow them, or A refuses to accept the condition and give up the collaboration with B under its expectation if the conditions violate its security rules.
- (8) **Trusted Collaboration:** After successful confirmation of trust condition embedment, the two platforms can collaborate based on the trust conditions specified by B and cooperate together in a service chain to fulfill a networking service. Note that Step (6) and (7) can be applied by Platform A to control trust sustainment based on its policy issued to Platform B.
- (9) **Distrust Notification:** When violating trust conditions, RTM of A is able to notify A that its operations that cause platform configuration changes are restricted or inform Platform B about the configuration changes. The same operations are conducted in Platform B if A embeds its trust conditions into B.

After receiving the notification of configuration change of A (B), corresponding actions will be taken. If necessary, re-challenge on the trust status will be performed to re-establish trust relationship under new conditions.

## 4 Implementation

We implemented the proposed mechanism in desktops running 64-bit RedHat Enterprise Linux Server 6.5 with 3.30 GHz Intel Core i3-3320 CPU and 4.0G RAM since the NFV deploys VNFs in general x86 Servers. We built up two Linux servers in VMware to simulate remote communications between NFV platforms. Then, we installed routers and firewalls as different VNFs on it. By using the programming interface provided by Trust Software Stack (TSS) [10], we implemented the proposed mechanism in C/C++ language and demonstrated the main functions of TIM. Herein, we used TPM-Emulator [10] (an open-source development platform) to simulate the RTM and store a special purpose signature key named AIK, which was used for platform identity authentication, platform attestation and certification, as well as signing signature. The TPM-Emulator simulated 80% functions of a standard TPM chip. TrouSers [10] is used to realize TSS functions and to provide reliable APIs for upper applications. TDDL is a TPM device driver library, which is a middle layer between TPM-Emulator and TSS. Through TPM-Emulator and Integrity Measurement Architecture (IMA) kernel patch packages [10], PCR lists, events and logs are generated at the startup of the Linux server platform.

We validated all the features of TIM. The PCRs change with the platform configuration change (e.g., router network configuration, firewall configuration) in the Linux Server. Before trust collaboration is established between remote and local NFV platforms, TPM performs integrity measurement in its embedded platform. Figure 4 shows the procedure of TPM-Emulator and TSS startup.

<pre> tpmd.c:264: Info: initializing socket /var/run/tpm/tpmd_socket: 0 tpmd.c:299: Debug: initializing TPM emulator tpm_emulator_extern.c:101: Info: _tpm_extern_init() tpm_emulator_extern.c:104: Debug: opening random device /dev/urandom tpm_cmd_handler.c:4113: Debug: tpm_emulator_init(2, 0x00000000) tpm_startup.c:29: Info: TPM_Init() tpm_testing.c:243: Info: TPM_SelfTestFull() tpm_testing.c:39: Debug: tpm_test_prng() tpm_testing.c:69: Debug: Monobit: 10145 tpm_testing.c:70: Debug: Poker: 32.0 tpm_testing.c:71: Debug: run_1: 2513, 2513 tpm_testing.c:72: Debug: run_2: 1263, 1251 tpm_testing.c:73: Debug: run_3: 626, 575 tpm_testing.c:74: Debug: run_4: 351, 324 tpm_testing.c:75: Debug: run_5: 147, 175 tpm_testing.c:76: Debug: run_6+: 116, 178 tpm_testing.c:77: Debug: run_34: 0 tpm_testing.c:111: Debug: tpm_test_sha1() tpm_testing.c:157: Debug: tpm_test_hmac() tpm_testing.c:184: Debug: tpm_test_rsa_EK() tpm_testing.c:186: Debug: tpm_rsa_generate_key() tpm_testing.c:191: Debug: testing endorsement key tpm_testing.c:197: Debug: tpm_rsa_sign(RSA_SSA_PKCS1_SHA1) tpm_testing.c:200: Debug: tpm_rsa_verify(RSA_SSA_PKCS1_SHA1) tpm_testing.c:203: Debug: tpm_rsa_sign(RSA_SSA_PKCS1_DER) tpm_testing.c:206: Debug: tpm_rsa_verify(RSA_SSA_PKCS1_DER) tpm_testing.c:210: Debug: tpm_rsa_encrypt(RSA_ES_PKCSV15) tpm_testing.c:214: Debug: tpm_rsa_decrypt(RSA_ES_PKCSV15) tpm_testing.c:218: Debug: verify plain text tpm_testing.c:221: Debug: tpm_rsa_encrypt(RSA_ES_OAEP_SHA1) tpm_testing.c:225: Debug: tpm_rsa_decrypt(RSA_ES_OAEP_SHA1) tpm_testing.c:229: Debug: verify plain text tpm_testing.c:261: Info: Self-Test succeeded tpm_startup.c:43: Info: TPM_Startup(2) tpmd.c:309: Debug: waiting for connections... </pre>	<pre> [root@myplatform build]# tcstd -e -f TCSD TDDL ioctl: (22) Invalid argument TCSD TDDL Falling back to Read/Write device support. TCSD trousers 0.3.13: TCSD up and running. </pre>
(a) TPM running state	(b) TSS running state

Fig. 4. TPM startup and TSS startup

## 5 Security Analysis and Performance Evaluation

### 5.1 Security Analysis

In TIM, the insurance of trust relationship is achieved through the verification to the communication entity on the other side according to the trust attestation protocol, so as to the trust status of the remote entity. The sustainment of trust relationship is achieved by dynamically monitoring the trust status of the other entity with RTM.

For Intra-Platform attacks, the platform layer security can be ensured by the RTM. Based on the RTM, we can build additional trust in any components installed upon it, for example, the NFVI middleware components and VNFs. In addition, VNF security can be protected with certificate verification or PCR value verification, which can be handled by RTM to ensure that VNFs are sourced from a trustworthy party and as expectation.

For Inter-Platform attacks, firstly, the insecure factors in inter-platform can be resisted by RTM-based booting, installation, certificate verification and PCR value verification. Other outside attacks and intrusions, as well as other security threats on networking devices can be overcome by deploying various security related functions as VNFs in the platform.

Mostly importantly, by running the trust insurance protocol, TIM can help a platform to embed trust conditions into a remote platform, with which it aims to collaborate. By making use of the RTM to monitor the platform configuration changes and match them with the trust conditions, TIM insures the trust relationship to be sustained during the collaboration of a number of VNFs located in different NFV platforms to fulfill an expected networking service.

### 5.2 Performance Evaluation

We performed a number of simulations to test the effectiveness of TIM with regard to trust attestation accuracy, operation efficiency, and reasonable resource consumption.

Figure 5 shows the procedure of platform challenge and trust attestation with a success message. This test achieves trust verification on a remote platform by checking

```
[myplatform@myplatform ima_attest]$ ./emos_ima_client 10.170.32.231
Receive Nonce Success!
Nonce: 8e723214
MeasureList:
boot_aggregate e07baa9b6f70913a8fdef1bd762c8b0aa3edf984
emos_ima_conf 5e9c92cf10dfc0f120e1af4571ee55d0b940b92
rcs a83947734285fac3e5e439659854c319626374cc
tty1.conf 7fa86adf2b53162bbf76cabff62b7e3a0578c1
tty2.conf fed2e0c4f13bbf975540a9b96331f02db084429b
tty5.conf 3eeb4e4e5834e4de14eeb84e180b149c9c4180e
tty4.conf e70420aedc6ae1d41163f859e93ea2bfa6b417
tty5.conf da185f2439c5f738778e57b9175a3edbd4c73ac2
tty6.conf c56dac3dae7d7dd182973c741c3440574313446c
.vimrc 1df08788db97615d172c6831ea2a8acc92e14089
syslog.conf 5fac3e439659854c8394177fe2626374cce5319
sysctl.conf a5c3deb84e189c9c4180e3a6de14e6374c0b1462
inetd.conf c8e3ade044a9b9c9e1b47beb818a84b47091ac6f
PCR_INDEX: 15
PCR_VALUE: a5621634b42e5b3406098c5fe2e31d200c6510a9
Attesting Success!
Receive Trust Conditions!
Verify Trust Conditions Success!
Load Trust Conditions Success!
Trust Relationship Established!
```

```
[verifierhost@verifierhost ima_attest]$ ./emos_ima_server
Connection Request from 10.170.32.230
Verify Certificate Success!
Generate Nonce Success!
Nonce: 8e723214
Receive Trust Evidence Success!
Verify Signature Success!
PCR_INDEX: 15
Hash_Value: a5621634b42e5b3406098c5fe2e31d200c6510a9
PCR_Value: a5621634b42e5b3406098c5fe2e31d200c6510a9
Verify PCR Success!
Send Trust Conditions Success!
Trust Relationship Established!
```

(a) TPM running state

(b) TSS running state

Fig. 5. Platform attestation status when challenge success

the value of the nonce used for challenging, PCR index and PCR values with an expected hash value. Figure 6 shows the prompts to deny a platform configuration change (in our test, we changed the router network interface configurations) because it violates a pre-set trust condition (i.e., any platform configuration change is not allowed). Any configuration changes caused by new software installation, platform update, and VNF upgrade can be detected and rejected accordingly with 100% accuracy since they cause platform configuration changes.

We further compared TIM with a scheme about cloud service attestation [11] regarding operation time since it is a scheme for cloud computing based on trusted computing and has a similar purpose to ours. As shown in Fig. 7, we observe that the operation performance of TIM exceeds the scheme in [11]. AIK setup is to build AIK certificate, which only costs 113.793 milliseconds (ms). The total time shown in Fig. 7 includes integrity and PCR verification time and data transmission time. In our implementation, the size of nonce is 160-bit and the size of trust condition data is 681 bytes. As AIK is a 2048-bit RSA signature key, the signature size is within 2048-bit, thus the communication cost of TIM is trivial.

```

Verify Trust Conditions Success!
Load Trust Conditions Success!
Trust Relationship Established!

Current Configuration Change is Restricted for Violating Trust Conditions!
MeasureList:
boot_aggregate e07baa9b6f70918a8fdef1bd762c8b0aa3edf984
emos_ima_conf 5e9c92cf10dfc0f120e1af4571ee55d0b940b92
rcs          a83947734285fac3e5e439659854c319626374cc
tty1.conf    7fa486adf2b533162bbf76cabff62b7e3a0578c1
tty2.conf    fe2e60cfd78bbf875540a0b96831f02db084429b
tty3.conf    3eeb4e44e5834e4de14eeb84e180b149c9c4180e
tty4.conf    e70420aedc6aae1d411163f855e93ea2bfa6b417
tty5.conf    da185f2439c5f738778e57b9175a3edb4c73ac2
tty6.conf    c56dac3da67d7dd182973c741c3440574313446c
.vimrc       1df08788db97615d1f2c6831ea2a8acc92e14d89
syslog.conf  5fac3e439659854ca839477fec8626374ccea5319
sysctl.conf  2a7d8637d2c436ce73e1a61eb15fa80eda35b8e4
inetd.conf   c8e3ade044a9b9c9e1b47beb818a84b47091ac6f
[myplatform@myplatform ima_attest]$ █

```

**Fig. 6.** Attestor status when configuration change

We additionally tested the resource consumption of TIM, concretely memory occupancy and CPU occupancy. We used Virtual-Memory Size (VMS) that includes the memory consumed by a process and its shared libraries, Resident Set Size (RSS) that is the non-swapped physical memory that a process has used and is the total memory actually held in RAM for the process and can represent the total memory that the process is used. Daemon is a backend program used by TPM, which can receive remote connection requests, thus support TPM and TSS functions. Table 1 reports the resource consumption of main processes of TIM in the platform when applying TIM. We find that they only consume few CPU and memory resources.



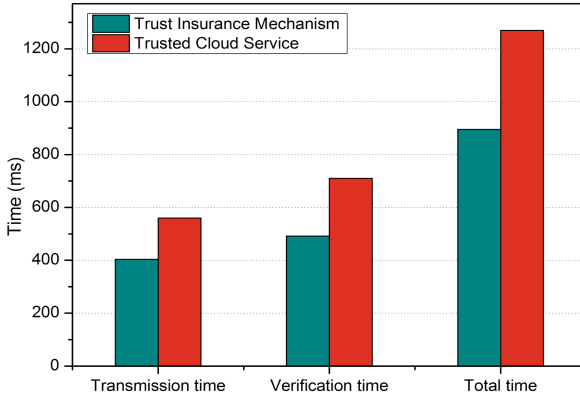


Fig. 7. Operation time

Table 1. Resource overhead (kb: KiloByte)

Process	Resource overhead			
	VMS (kb)	RSS (kb)	CPU%	Mem%
TPM	6968	1156	0.0%	0.0%
TSS	94720	1864	0.1%	0.0%
Daemon	168456	1200	0.1%	0.0%

In addition, we compared the platform CPU occupancy of TIM with the situation that TIM is not applied. The result is shown in Fig. 8. Comparing with the highest occupancy, we observe that applying TIM costs about 5% more CPU overhead, and on average TIM occupies about 10% more CPU resources. We can see that the resource consumption of TIM is reasonable and acceptable.

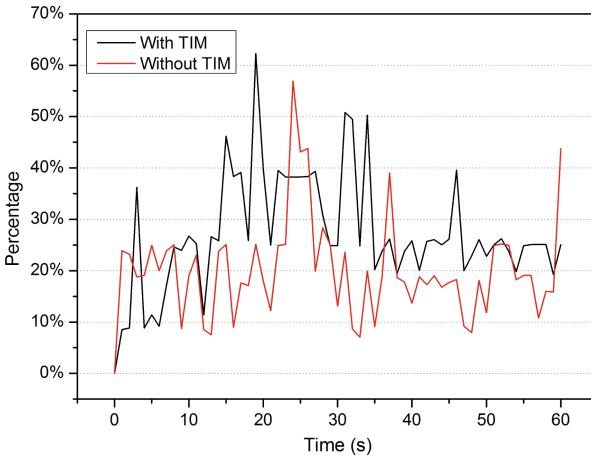


Fig. 8. CPU occupancy

## 6 Conclusions and Future Work

In this paper, we presented TIM, a trust insurance mechanism based on trusted computing in order to provide secure and trusted networking services in the infrastructure of NFV. This mechanism achieves trust authentication and trusted collaboration between platforms by applying RTM to monitor platform configurations and ensure their changes compatible with trust conditions. Performance evaluation results show the effectiveness of TIM with regard to trust attestation accuracy, operation efficiency, and reasonable resource consumption. Regarding the future work, we plan to optimize TIM in order to reduce resource consumption and propose a fine-grained method for detecting configuration changes in NFV platforms.

**Acknowledgments.** This work is sponsored by the National Key Research and Development Program of China (grant 2016YFB0800700), the NSFC (grants 61672410 and U1536202), the Project Supported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2016ZDJC-06), the 111 project (grants B16037 and B08038), the PhD grant of the Chinese Educational Ministry (grant JY0300130104), and Academy of Finland (grant 308087).

## References

1. Yan, Z., Zhang, P., Vasilakos, A.V.: A security and trust framework for virtualized networks and software-defined networking. *Secur. Commun. Netw.* **9**(16), 3059–3069 (2016)
2. Yang, W., Fung, C.: A survey on security in network functions virtualization. In: *Proceeding of IEEE NetSoft 2016*, pp. 15–19, Seoul (2016)
3. Aljuhani, A., Alharbi, T.: Virtualized network functions security attacks and vulnerabilities. In: *Proceeding of IEEE CCWC 2017*, pp. 1–4, Las Vegas (2017)
4. Shin, M.K., Choi, Y., Kwak, H.H., Pack S., Kang M., Choi, J.Y.: Verification for NFV-enabled network services. In: *Proceeding of IEEE ICTC 2015*, pp. 810–815, Jeju (2015)
5. Liu, Y., Guo, Z., Shou, G., Hu, Y.: To achieve a security service chain by integration of NFV and SDN. In: *Proceeding of IEEE IMCCC 2016*, pp. 974–977, Harbin (2016)
6. Jang, H., Jeong, J., Kim, H., Park, J.S.: A survey on interfaces to network security functions in network virtualization. In: *Proceeding of IEEE AINA 2015*, pp. 160–163, Gwangju (2015)
7. Biczók, G., Sonkoly, B., Bereczky, N., Boyd, C.: Private VNFs for collaborative multi-operator service delivery: an architectural case. In: *Proceeding of NOMS 2016*, pp. 1249–1252, Istanbul (2016)
8. Aguado, A., Hugues-Salas, E., Haigh, P.A., Marhuenda, J., Price, A.B., Sibson, P.: Secure NFV orchestration over an SDN-controlled optical network with time-shared quantum key distribution resources. *J. Lightwave Technol.* **35**(8), 1357–1362 (2017)
9. Bari, M.F., Chowdhury, S. R., Ahmed, R., Boutaba, R.: On orchestrating virtual network functions. In: *Proceeding of IEEE CNSM 2015*, pp. 50–56, Barcelona (2015)
10. Strasser, M., Sevcic, P.E.: A software-based TPM emulator for Linux. *ETH Zurich, Semesterarbeit* (2004)
11. Ren, J., Liu, L., Zhang, D., Zhang, Q., Ba, H.: Tenants attested trusted cloud service. In: *Proceeding. IEEE CLOUD 2016*, pp. 600–607, San Francisco (2016)

12. Yan, Z., Cofta, P.: A mechanism for trust sustainability among trusted computing platforms. In: Katsikas, S., Lopez, J., Pernul, G. (eds.) TrustBus 2004. LNCS, vol. 3184, pp. 11–19. Springer, Heidelberg (2004). [https://doi.org/10.1007/978-3-540-30079-3\\_2](https://doi.org/10.1007/978-3-540-30079-3_2)
13. Yan, Z.: Trust Management in Mobile Environments - Usable and Autonomic Models. IGI Global, Hershey (2013)
14. Lauer, H., Kuntze, N.: Hypervisor-based attestation of virtual environments. In: Proceeding of IEEE UIC/ATC/ScalCom/SmartWorld 2016, pp. 333–340, Toulouse (2016)
15. Mathieu, B., Doyen, G., Mallouli, W., et al.: Monitoring and securing new functions deployed in a virtualized networking environment. In: Proceedings of ICARS 2015, pp. 741–748, Toulouse (2015)

# Personalized Semantic Location Privacy Preservation Algorithm Based on Query Processing Cost Optimization

Mengzhen Xu, Hongyun Xu<sup>(✉)</sup>, and Cheng Xu

School of Computer Science and Engineering,  
South China University of Technology, Guangzhou 510006, China  
hongyun@scut.edu.cn

**Abstract.** Location-Based Services (LBSs) are playing an increasingly important role in our daily life with the development of GPS or WiFi enabled space positioning technologies and the popularization of mobile devices. While LBSs bring great conveniences to users, the exposure of users' location privacy becomes a growing concern. To address this issue, researchers propose several kinds of location preservation techniques such as location cloaking, dummies and etc. However, these methods are rather vulnerable when encounter a location semantic attack. Taking this into consideration, a few semantic location preservation methods are proposed. Based on the existing semantic location preservation frameworks, we propose a novel personalized semantic location privacy preservation method named Incremental Search (IS). In our method, an optimal anonymous location set is generated according to a certain rule, during which, two parameters are introduced to limit the number of locations in the final anonymous location set and the number of the anonymous location sets recorded temporarily so as to reduce the query processing cost. The evaluation on NGMO simulation platform validates that our method has a better performance than the two baseline algorithms.

**Keywords:** Semantic location · Privacy preservation · Incremental search  
Query processing cost · Road network

## 1 Introduction

With the popularization of mobile and locator devices, Location-based services (LBSs) have been used widely. However, people are suffering from the threats of leaking their identities and location information when they are enjoying LBSs. Attackers can infer user's ongoing activities or analyze his habits according to user's information such as current location and time.

Therefore, many researchers aimed at studying the privacy preservation methods in LBSs and proposed a lot of solutions mainly based on K-anonymity [1] and L-diversity [2]. The main idea of K-anonymity is guaranteeing at least other  $K - 1$  users in the obfuscation region together with the user, then the user's practical position will be

replaced with the obfuscation region and then sent to the LBS server, so the probability for the attacker to infer the user's location correctly is  $\frac{1}{K}$ . K-anonymity can protect user's identity and location information, but it does not consider the distribution of the  $K$  users in the obfuscation region. L-diversity is a supplementary to K-anonymity which requires anonymity area to contain at least  $L$  different locations so as to prevent attackers from linking the user to a location.

However, L-diversity may not always be so effective for all semantic location privacy preservation in some cases. On the one hand,  $L$  locations may be the same type, thus user's location information is revealed, on the other hand, the probability of user being in each location may not be equal, thus the attacker can filter out the locations with low probabilities. To the former problem, researchers extract L-diversity to  $L$  different types of locations [3] rather than L-occurrence locations. To the latter problem, numerical values are used to quantify the probability of user being in each location, for example, popularity is used in [4, 5].

Generally speaking, a user may be sensitive to some locations while not to other locations. If users stay at sensitive locations, they need privacy preservation, otherwise, they do not need. Yigitoglu et al. [4, 5] use privacy profile to define user's sensitive locations, use popularity to define all kinds of locations and use the ratio of the aggregated popularity of the sensitive locations to the aggregated popularity of all locations to judge whether the anonymous area meets the privacy requirement. However, the algorithms [4, 5] just aim to find an anonymous area satisfying privacy requirement without considering the number of locations, which may lead to high query processing cost. Obviously, on the condition of meeting the privacy requirement, fewer locations will lead to lower query processing cost.

Our paper proposes a novel personalized semantic location privacy preservation algorithm named incremental search (IS), which aims to reduce query processing cost.

The main contributions of our paper are listed below:

- (1) We propose to select anonymous locations based on global optimization and local optimization to protect users' location privacy under semantic environment.
- (2) Two parameters are introduced to limit space cost and time cost.
- (3) We compare our algorithm with two existing algorithms based on a real map, the result shows that our algorithm can improve anonymous success rate and reduce query processing cost.

The rest parts are organized as follows. In Sect. 2 we review the related work. Section 3 introduces the system model. Section 4 describes the algorithm in detail. The experimental results are shown in Sect. 5. And we conclude the paper in Sect. 6.

## 2 Related Work

Previous work in location privacy preservation mainly includes dummies [6], mix-zone [7], K-anonymity [1, 8–11], obfuscation and coordinate transformation [12], cryptography based method [13]. K-anonymity is widely used which means user cannot be

recognized with at least other  $K - 1$  users by using location or identity. Based on  $K$ -anonymity, many other approaches have been proposed, such as Interval-cloak [8], Casper [9], Clique-Cloak [10] and HC [11].

However, these approaches do not consider spatial context. Therefore, Bamba et al. propose PrivacyGrid [14], a system which can support both  $K$ -anonymity and  $L$ -diversity. In their approach, an anonymous region contains at least  $K$  users and  $L$  locations. However, the region in [14] is not  $L$ -type diverse. For example,  $L$  locations may all be the same type (e.g. hospital), which is vulnerable to location similarity attack [5]. Therefore, Xue et al. [3] define  $L$ -diversity to be  $L$  different types. These approaches do not consider that different types of locations have different probabilities of being visited; Probe [15–17] takes this situation into consideration by introducing the concept of popularity to measure the visiting probabilities of each type of locations, what's more, it classifies locations into sensitive locations and non-sensitive locations, then uses obfuscation method to complete privacy preservation. Byoungyoung [18] proposes a method that uses EMD (earth mover's distance) to mine location semantic information to avoid privacy being revealed, but this is only suitable for Euclidean space. Yigitoglu et al. [4] and Li [5] extend these approaches into road networks. In [4], they generate the city network from a map on which each location is a node, then use breadth first search algorithm to find a location set which makes the proportion of the popularity of the sensitive locations less than a certain threshold. In [5], they select all non-sensitive semantic locations from the neighbor locations of the anonymous area and add them into the anonymous area in turn, until it satisfies the privacy requirement or the terminal conditions. However, all of those approaches do not consider the number of locations in the anonymous set, which may lead to high query processing cost.

This paper proposes an algorithm to protect user's location privacy under the semantic environment, which is suitable for the road-network environment with higher success rate and lower query processing cost.

### 3 System Model

#### 3.1 Semantic Location City Network

**Definition 1 (Semantic location city network).** A semantic location city network is modeled as a connected and undirected weighted graph  $G$ ,  $G = (V, E, pt, pop)$ , where:

- (1)  $V$  is the set of vertices,  $V = V_p \cup V_i$ ,  $V_p$  is the set of semantic locations and  $V_i$  is the set of road intersections.
- (2)  $E$  is the non-empty set of edges,  $E \subseteq V \times V$ ,  $E = \{e \mid e = (u, v), \text{ where } u, v \in V\}$ .
- (3)  $pop$  is the popularity of  $v$  ( $v \in V$ ), the popularity of  $v$  is the probability that a user visit the location  $v$ .  $pt$  is the type of  $v$ , each location's type is represented by function  $pt: V \rightarrow PT$  ( $PT$  is the set of location types).

For simplicity, we assume that locations with the same type have the same popularity, represented by function  $pop: PT \rightarrow [0, 1)$ , the popularity of road intersection is 0 (In this paper, we assume a road intersection belongs to a location which is not semantic). The popularity of semantic location  $v$  is denoted as  $p(v)$ , the popularity of semantic location set  $s$  is denoted as  $p(s)$ .

### 3.2 System Structure

Figure 1 depicts the Central Server Structure designed to provide location privacy preservation. We add a Location Privacy Server (LPS) between mobile users and the LBS server. LPS stores users' privacy profiles ( $pp$ ). The process for user to complete a query is as follows: 1. User sends the query request. 2. LPS receives the query and judges whether the user's location is sensitive; if it is non-sensitive, LPS sends the real location to LBS server; if it is sensitive, LPS runs location privacy preservation algorithm and generates a location set based on the user's location, then sends the location set to LBS server. 3. LBS server calculates the candidate sets based on the real location or the location set and sends the candidate sets to LPS. 4. LPS filters the candidate sets and sends the appropriate results to the user. Our work focuses on step 2, namely providing privacy preservation for user's requests from sensitive locations.

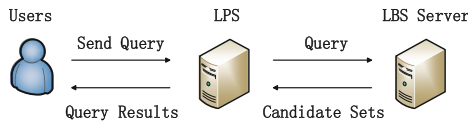


Fig. 1. Central server structure

### 3.3 Privacy Requirement

Let  $PT = PTs \cup PTns$ , where  $PTs$  represents the sensitive type set,  $PTns$  represents the non-sensitive type set. Let  $Vp = Vsp \cup Vnsp$ ,  $vsp$  represents a sensitive semantic location,  $vsp \in Vsp$ ,  $vns$  represents a non-sensitive semantic location,  $vns \in Vnsp$ .

**Definition 2 (Anonymous location set, ALS).** *ALS* is a location set used to achieve anonymity. The element in *ALS* can be a sensitive semantic location, a non-sensitive semantic location or a road intersection.

**Definition 3 ( $\theta$ -secure ALS).** If an *ALS* satisfies Eq. (1)

$$\frac{\sum_{VSP \in VS \cap VSP \in ALS} P(VSP)}{\sum_{VP \in VS \cap VP \in ALS} P(VP)} \leq \theta \tag{1}$$

We denote this *ALS* as a  $\theta$ -secure *ALS*. In other words,  $\theta$ -secure *ALS* satisfies the following condition: the ratio of the aggregated popularity of all sensitive semantic locations (*APSSL*) to the aggregated popularity of all semantic locations (*APSL*) in an *ALS* should be less than or equal to  $\theta$ .

**Definition 4 (*Lmax*).** It represents the maximal number of locations in an *ALS*. Although more locations in an *ALS* will provide better privacy preservation for users, it will also result in more query processing cost to the LBS server. To balance privacy preservation and query processing cost, we propose parameter *Lmax* which is decided by users.

**Definition 5 (*Cmax*).** It represents the maximal number of *ALS* to be recorded. Although recording more *ALS* will make the final *ALS* easier to contain fewer locations, it will result in larger space and time cost to LPS server. To balance the number of locations in the final *ALS* and the cost of LPS, we propose parameter *Cmax* which is also decided by users.

**Definition 6 (Deficient popularity, *DP*).** It represents the least popularity that an *ALS* needs to be a  $\theta$ -secure *ALS*. For example, an *ALS*'s *APSSL* is 0.3 and its *APSL* is 0.5. If  $\theta = 0.3$ , we can calculate the *DP* of the *ALS* is  $\frac{0.3}{0.3} - 0.5$ , we got 0.5 *DP* is used to achieve local optimization.

## 4 Algorithm

We describe IS algorithm in this section. Before executing IS algorithm, the user should provide *pp* and the values of *Cmax* and *Lmax*. *pp* contains: (1) Sensitive location type set *PTs* = {*pts1*, *pts2*, ..., *ptsn*}. (2) The value of  $\theta$ .

The pseudo-code of this algorithm is given in Algorithm 1. The algorithm is accomplished by looping. During each loop we search new *ALS* by adding an adjacent location to the recorded *ALS*; we keep two sets *curSet* and *nextSet*, *curSet* records the *ALS* achieved in last loop, *nextSet* records new *ALS* searched through *ALS* in *curSet*. Two hash tables *sp* and *tp* are defined to record the *APSSL* and *APSL* for every *ALS* in *curSet* and *nextSet*. Both *sp* and *tp* will be updated when a new *ALS* is found, *curSet* will be updated in the end of the loop. The algorithm terminates when a  $\theta$ -secure *ALS* is found or the number of locations of an *ALS* in *curSet* reaches *Lmax*.

As mentioned before, parameter *Cmax* is defined to limit the number of *ALS* being recorded (In Algorithm 1, the *ALS* are recorded in *curSet* and *nextSet*). Therefore, if the number of *ALS* in *nextSet* is larger than *Cmax*, we will select the top *Cmax* local optimal *ALS*



---

**Algorithm 1**

**Input:** semantic location city network  $G = (V, E, pt, pop)$ , user's location  $loc_u$ , privacy profile  $pp = \{PTs, \theta\}$ , parameter  $Lmax$ , parameter  $Cmax$ .

**Output:** a  $\theta$ -secure  $ALS$ .

---

```

1:   $curSet = \{\{loc_u\}\}$ ,  $nextSet = \{\}$ ,  $sp.put(\{loc_u\}, p(\{loc_u\}))$ ,  $tp.put(\{loc_u\}, p(\{loc_u\}))$ 
2:  While (true) do
3:    For each  $seti$  in  $curSet$  do
4:      If  $seti.size > Lmax$  then
5:        return false;
6:      For each  $setii$  in  $seti$  do
7:         $Temp = seti$ 
8:         $Vadj = getAdj(setii)$  //  $Vadj$  is the set of all adjacent nodes of node  $setii$ 
9:        For each  $node$  in  $Vadj$  do
10:       If  $!Temp.contains(node)$  then
11:          $Temp.add(node)$ 
12:          $nextSet.add(Temp)$ 
13:          $tp.put(Temp, tp.get(seti) + p(node))$  //update  $tp$ 
14:         If  $PTs.contains(node.type)$  then //update  $sp$ 
15:            $sp.put(Temp, sp.get(seti) + p(node))$ 
16:         Else
17:            $sp.put(Temp, sp.get(seti))$ 
18:         End
19:         If  $(sp.get(Temp) / tp.get(Temp) \leq \theta)$  then
20:           return  $Temp$ ;

```

```

21:         End
22:     End
23:     Temp = seti;
24:     End
25: End
26: sp.remove(seti)
27: tp.remove(seti)
28: End
29: curSet.clear();
30: curSet = Get_Cmax(nextSet, sp, tp,  $\theta$ );
31: nextSet.clear();
32: End
33:
34://Select Cmax local optimal ALS from nextSet
35: Set Get_Cmax(nextSet, sp, tp,  $\theta$ ) {
36:     If nextSet.size > Cmax then
37:         For each seti in nextSet do           //calculate DP for every ALS in nextSet
38:             dp = sp.get(seti) /  $\theta$  - tp.get(seti)
39:         End
40:         nextSet.sort_by(dp)                       // sort nextSet as the value of DP increase
41:         newSet = nextSet.first(Cmax)             // select the first Cmax values in nextSet
42:         return newSet
43:     Else
44:         return nextSet
45:     End
46: }

```

---

Figure 2 depicts a semantic location city network. Each vertex has two attributes, the first one represents identity and the second one represents type. Assuming that the privacy profile  $pp = \{\{H,O\}, 0.5\}$ , the popularities of each location type are as follows, {School (S): 0.2, Hospital (H): 0.15, Office (O): 0.25, Entertainment (E): 0.15, Mall (M): 0.15, Park (P): 0.1, Intersection (I): 0}. User locates at  $v1$ ,  $Lmax = 3$ ,  $Cmax = 1$ . In the beginning,  $curSet = \{\{v1\}\}$ ,  $sp = \{\{v1\} \rightarrow 0.15\}$ ,  $tp = \{\{v1\} \rightarrow 0.15\}$ ,  $nextSet = \{\}$ , there is only one ALS ( $\{v1\}$ ) in  $curSet$ .

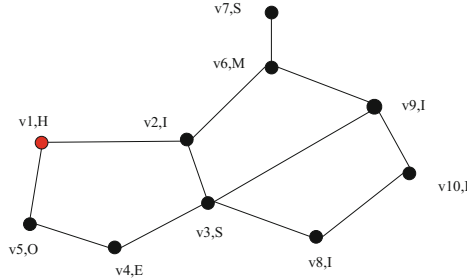


Fig. 2. An example of IS algorithm

In the first round, the number of locations for each ALS is 1, which is smaller than  $Lmax$ , so continue searching. The adjacent locations of ALS are  $v2$ ,  $v5$ . select  $v2$  first, then we obtain  $ALS = \{v1, v2\}$ ,  $sp(\{v1, v2\}) = 0.15$ ,  $tp(\{v1, v2\}) = 0.15$ . Since  $\frac{0.15}{0.15} > 0.5$ , it does not meet the privacy requirement, add  $\{v1, v2\}$  to  $nextSet$ . Then select  $v5$ , and we obtain  $ALS = \{v1, v5\}$ ,  $sp(\{v1, v5\}) = 0.4$ ,  $tp(\{v1, v5\}) = 0.4$ . Since  $\frac{0.4}{0.4} > 0.5$ , it does not meet the privacy requirement, add  $\{v1, v5\}$  to  $nextSet$ . We obtain  $nextSet = \{\{v1, v2\}, \{v1, v5\}\}$ ,  $sp = \{\{v1, v2\} \rightarrow 0.15, \{v1, v5\} \rightarrow 0.4\}$ ,  $tp = \{\{v1, v2\} \rightarrow 0.15, \{v1, v5\} \rightarrow 0.4\}$ . The number of sets in  $nextSet$  is 2, which is larger than  $Cmax$ , so we need to select the top  $Cmax$  local optimal ALS. The DP of  $ALS\{v1, v2\}$  is  $\frac{0.15}{0.5} - 0.15 = 0.15$  and the DP of  $ALS\{v1, v5\}$  is  $\frac{0.4}{0.5} - 0.4 = 0.4$ . Because  $0.15 < 0.4$ , so we select set  $\{v1, v2\}$ . Thus  $curSet = \{\{v1, v2\}\}$ ,  $sp = \{\{v1, v2\} \rightarrow 0.15\}$ ,  $tp = \{\{v1, v2\} \rightarrow 0.15\}$ ,  $nextSet = \{\}$ .

In the second round, the number of locations for each ALS in  $curSet$  is 2, which is smaller than  $Lmax$ , so continue searching. The adjacent locations of  $\{v1, v2\}$  are  $v5$ ,  $v3$ ,  $v6$ . Select  $v5$  first, then we obtain  $ALS = \{v1, v2, v5\}$ ,  $sp(\{v1, v2, v5\}) = 0.4$ ,  $tp(\{v1, v2, v5\}) = 0.4$ . Since  $\frac{0.4}{0.4} > 0.5$ , it does not satisfy the privacy requirement, add  $\{v1, v2, v5\}$  to  $nextSet$ . Then select  $v3$ , and we obtain  $ALS = \{v1, v2, v3\}$ ,  $sp(\{v1, v2, v3\}) = 0.15$ ,  $tp(\{v1, v2, v3\}) = 0.35$ . Since  $\frac{0.15}{0.35} < 0.5$ , it satisfies the privacy requirement, so  $\{v1, v2, v3\}$  is sent as the final ALS to LBS server. The algorithm terminates.

Assuming that the number of adjacent locations for each location is a constant value  $A$ , the number of recorded ALS is  $C$  and the number of locations in an ALS is  $L$ . For simplicity, the case that two ALS become equal after each of them adding one location respectively is ignored.

When  $C < C_{max}$  and  $L < L_{max}$ : if the number of locations in *ALS* changes from  $m$  to  $m + 1$ , the number of *ALS* that are needed to be recorded increases  $A^*m$  times, so the space complexity and the time complexity are both  $O(A^L * L)$ .

When  $C = C_{max}$  and  $L < L_{max}$ : the number of *ALS* that are needed to be recorded is a constant, thus the space complexity is  $O(L)$  and the time complexity is  $O(L^2 \log L)$ .

In IS algorithm, both  $L_{max}$  and  $C_{max}$  may affect the running time. Larger  $L_{max}$  or  $C_{max}$  may result in longer running time and larger storage space, so it is important to choose the appropriate values of  $L_{max}$  and  $C_{max}$ . Generally speaking,  $L_{max}$  and  $C_{max}$  can be larger if the machine’s performance is good enough. More discussions about these two parameters will be shown in next section.

## 5 Experimental Analysis

In this section, we evaluate the performance of the proposed algorithm. Two different *PTs* are used where  $PTs_1 = \{\text{Entertainment}\}$  and  $PTs_2 = \{\text{Hospital, Office}\}$ . Overlapping algorithm in [4] and SA algorithm in [5] are implemented for comparison.

### 5.1 Experiment Setting

The algorithms are realized using Java and the coding environment is Eclipse. The experimental platform consists of a desktop PC equipped with an Intel(R) Pentium (R) 4 2.66 GHz CPU and 2 GB of RAM. The famous Network-based Generator of Moving Objects (i.e., NGMO) simulation platform is used to accomplish the experiments. We use the map of Oldenburg city in Germany, which contains 6105 nodes and 7035 edges. Raw data is processed according to Definition 1 to generate semantic location city network. We traverse every location and send query request. In this experiment, we just discuss six types of semantic location whose popularity are the same as we assumed in Sect. 4. Experiment parameters are shown in Table 1.

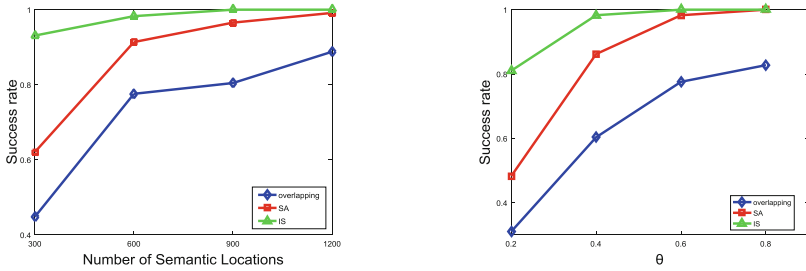
**Table 1.** Parameter settings

Parameter	Default value	Range
$\theta$	0.4	[0.2, 0.8]
Number of semantic locations	600	[300, 1200]
Location types (counts)	School (S:64), Office (O:174), Hospital (H:72), Market (M:100), Entertainment (E:58), Park (P:132)	/
$L_{max}$	30	[20, 40]
$C_{max}$	1000	[100, 1000]

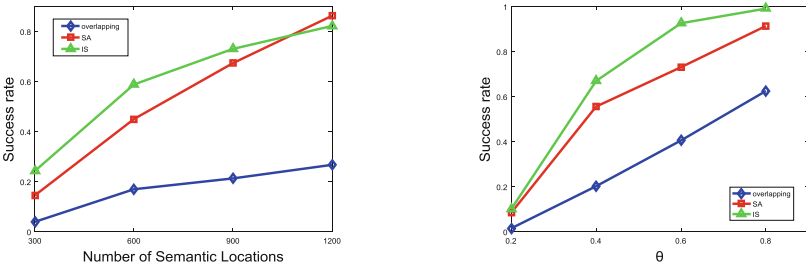
### 5.2 Experimental Result

Success rate is the ratio of the number of successful anonymous sensitive semantic locations to the number of total sensitive semantic locations. It is an important metric

that reflects the effectiveness of a privacy preservation algorithm. If the success rate is higher, the privacy preservation algorithm is more effective. Figure 3 depicts the tendencies of success rate with varying number of semantic locations and  $\theta$  when sensitive location set is  $PTS_1$  and  $PTS_2$  respectively.



(a)  $\theta$  takes default value ,using  $PTS_1$  (b) Number of semantic locations takes default value, using  $PTS_1$

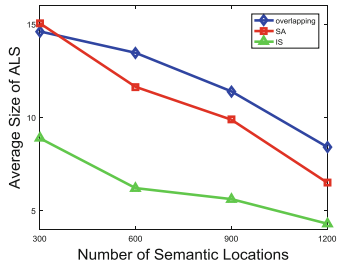


(c)  $\theta$  takes default value,using  $PTS_2$  (d) Number of semantic locations takes default value, using  $PTS_2$

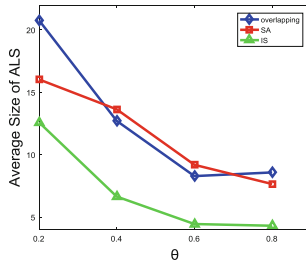
**Fig. 3.** Relation between success rate and the number of semantic locations or  $\theta$

From Fig. 3, we can see when the number of semantic locations or  $\theta$  increases, success rates of three algorithms all increase. This is because when the number of semantic locations increases, both the number of sensitive and the number of non-sensitive semantic locations increase, which makes it easier to find a  $\theta$ -secure ALS; when  $\theta$  increases, privacy requirement is easier to be satisfied. IS has the highest success rate among three algorithms, since it records many (no more than  $C_{max}$ ) ALSs and is more possible to find a  $\theta$ -secure ALS. For SA, it records only one ALS, thus it is harder to get a  $\theta$ -secure ALS compared with IS. For overlapping, it fails when the new searched location is not a non-sensitive location, which makes its success rate lower.

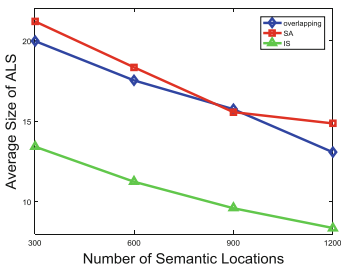
Average size of ALS is the average number of locations in an ALS during a successful query, it can be used to measure query processing cost at LBS server. Figure 4 depicts the tendencies of average size of ALS with varying number of semantic locations and  $\theta$  when sensitive location set is  $PTS_1$  and  $PTS_2$  respectively.



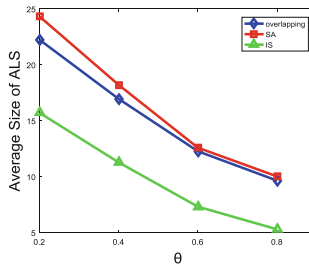
(a)  $\theta$  takes default value, using  $PTs_1$



(b) Number of semantic locations takes default value, using  $PTs_1$



(c)  $\theta$  takes default value, using  $PTs_2$



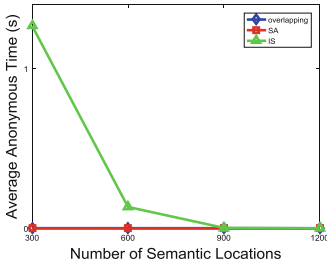
(d) Number of semantic locations takes default value, using  $PTs_2$

**Fig. 4.** Relation between average size of  $ALS$  and the number of semantic locations or  $\theta$

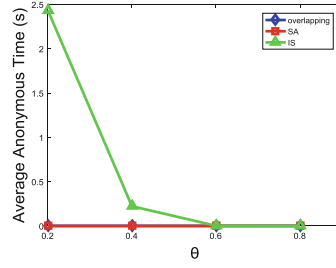
From Fig. 4, we can see when the number of semantic locations or  $\theta$  increases, the average sizes of  $ALS$  of three algorithms all decrease. This is because when the number of semantic locations increases, it is easier to find an adjacent location which is non-sensitive, thus fewer locations are needed to get a  $\theta$ -secure  $ALS$ ; when  $\theta$  increases, privacy requirement is easier to be satisfied, which leads to a smaller average size of  $ALS$ . IS algorithm has the smallest average size of  $ALS$ , since it combines global optimization and local optimization which will return an  $ALS$  with smaller average size compared with SA and overlapping.

Average anonymous time is a metric to evaluate the effectiveness of an anonymity algorithm. It refers to the average time of a successful query's anonymization process. An anonymity algorithm is thought to be more effective if its average anonymous time is lower. Figure 5 depicts the tendencies of average anonymous time with varying number of semantic locations and  $\theta$  when sensitive location set is  $PTs_1$  and  $PTs_2$  respectively.

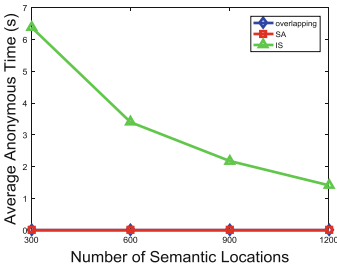
From Fig. 5, we can see, IS does not perform so well as overlapping and SA, because it records more than one  $ALS$  and needs more time to find a  $\theta$ -secure  $ALS$ . However, with the number of semantic locations or  $\theta$  increasing, the average anonymous time is getting closer to SA and overlapping.



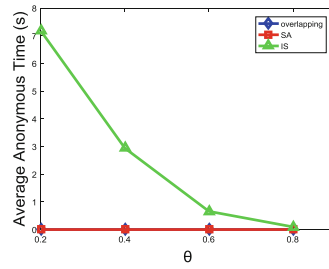
(a)  $\theta$  takes default value, using  $PTs_1$



(b) Number of semantic locations takes default value, using  $PTs_1$



(c)  $\theta$  takes default value, using  $PTs_2$



(d) Number of semantic locations takes default value, using  $PTs_2$

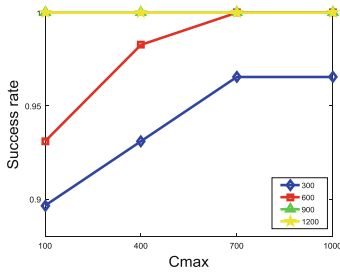
**Fig. 5.** Relation between average anonymous time and the number of semantic locations or  $\theta$

### 5.3 Parameters Discussion

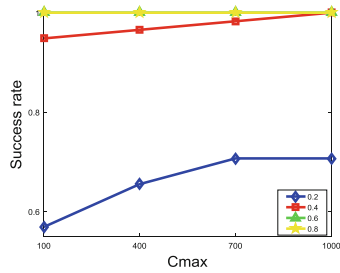
In our algorithm, both the values of  $Cmax$  and  $Lmax$  are very important. In this section we discuss them in detail. For simplicity, we only discuss the case that the sensitive location set is  $PTs_1$ .

Figure 6 depicts the tendencies of success rate, average size of  $ALS$ , average anonymous time with varying  $Cmax$  when the number of semantic locations or  $\theta$  takes default value. From Fig. 6(a) and (b), we can see success rate increases and the growth rate gradually decreases when  $Cmax$  increases. From Fig. 6(c) and (d), we can see the average size of  $ALS$  decreases gradually and tends to be a stable value. This is because when  $Cmax$  increases, the number of recorded  $ALS$  enlarges which makes it easier to get a  $\theta$ -secure  $ALS$ . From Fig. 6(e) and (f), we can see average anonymous time is high when  $\theta$  is 0.2 or the number of semantic locations is between 300 and 600, but it is low in other conditions. This is because when  $\theta$  is 0.2 or the number of semantic locations is between 300 and 600, the success rate is low. The failed anonymous process cost much time since IS terminates until the size of  $ALS$  exceeds  $Lmax$ . While in other conditions, success rate is high, so the value of average anonymous time is low. When

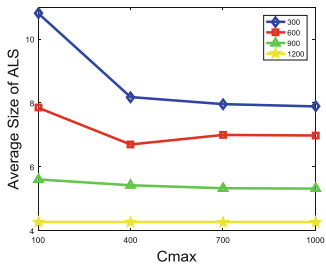
$C_{max}$  is smaller than 400, success rate is low and average size of ALS is large; when  $C_{max}$  is bigger than 700, the average anonymous time becomes very large, so  $C_{max}$  is suggested to be between 400 and 700.



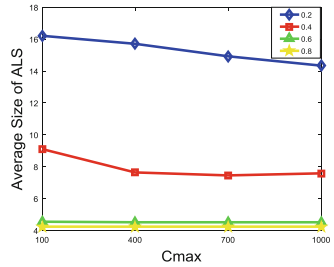
(a)  $\theta$  takes default value



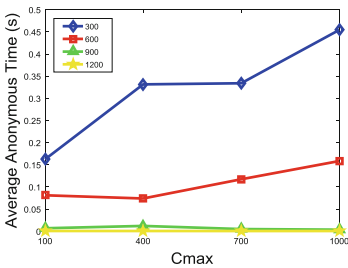
(b) Number of semantic locations takes default value



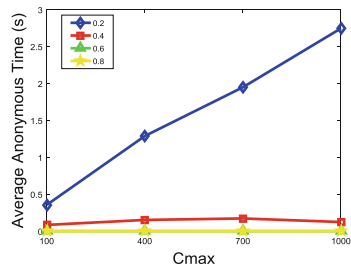
(c)  $\theta$  takes default value



(d) Number of semantic locations takes default value



(e)  $\theta$  takes default value

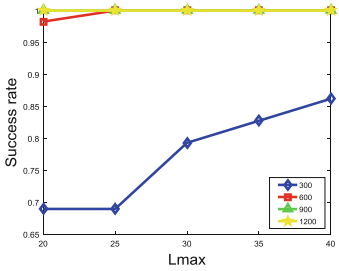


(f) Number of semantic locations takes default value

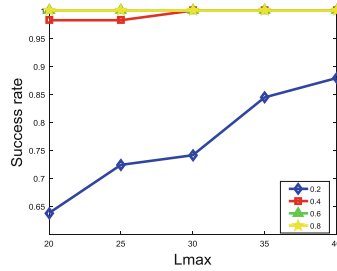
**Fig. 6.**  $C_{max}$  value analysis



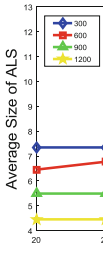
Figure 7 depicts the tendencies of success rate, average size of ALS, average anonymous time with varying  $L_{max}$  when the number of semantic locations or  $\theta$  takes default value. From Fig. 7, we can see when the number of semantic locations is 300 or  $\theta$  is 0.2, success rate is increasing with the raise of  $L_{max}$ . The growth rate is high when  $L_{max}$  is smaller than 30 and low when  $L_{max}$  is bigger than 35. Both the average size of ALS and the average anonymous time increase with  $L_{max}$  increasing. Success rate



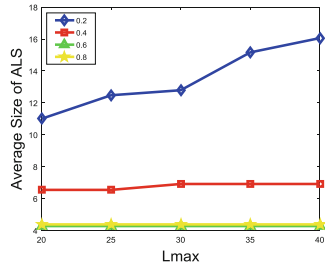
(a)  $\theta$  takes default value



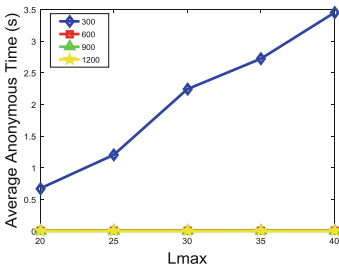
(b) Number of semantic locations takes default value



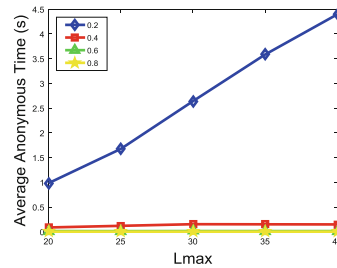
(c)  $\theta$  takes default value



(d) Number of semantic locations takes default value



(e)  $\theta$  takes default value



(f) Number of semantic locations takes default value

Fig. 7.  $L_{max}$  value analysis

increases because a bigger  $L_{max}$  makes it easier to satisfy the  $\theta$ -secure  $ALS$  requirement. Some semantic locations which fail to find the  $\theta$ -secure  $ALS$  with small  $L_{max}$  eventually find a  $\theta$ -secure  $ALS$  after  $L_{max}$  enlarges, thus the average size of  $ALS$  increases. At the same time, average anonymous time becomes longer because these semantic locations cost much time to find a  $\theta$ -secure  $ALS$ . When  $L_{max}$  is larger than 30, the average size of  $ALS$  is large and average anonymous time is long, so it is better to choose 30 as the value of  $L_{max}$ . When the number of semantic locations or  $\theta$  takes other values, success rate is very close to 1, and the average size of  $ALS$  and average anonymous time almost keep the same, which means that  $L_{max}$  has few impacts on IS when  $\theta$  or the number of semantic locations is large.

## 6 Conclusion

This paper proposed IS algorithm to achieve personalized semantic location privacy preservation. According to combining global optimization with local optimization, IS can improve anonymous success rate and reduce query processing cost. Two parameters  $C_{max}$  and  $L_{max}$  are introduced to limit space and time cost. By comparing IS with overlapping and SA in three aspects, the experimental results show that IS has good performance.

**Acknowledgements.** This work was partially supported by the Natural Science Foundation of China (No. 61272403), by the Fundamental Research Funds for the Central Universities (No. 10561201474).

## References

1. Sweeney, L.: K-anonymity: A model for protecting privacy. *Int. J. Uncert. Fuzz. Knowl. Syst.* **10**(5), 557–570 (2002)
2. Machanavajjhala, A., Kifer, D., Gehrke, J., Venkitasubramaniam, M.: L-diversity: privacy beyond k-anonymity. *Proc. ACM Trans. Knowledge Discov. Data (TKDD)*, 152 (2007)
3. Xue, M., Kalnis, P., Pung, H.K.: Location diversity: enhanced privacy protection in location based services. In: Choudhury, T., Quigley, A., Strang, T., Suginuma, K. (eds.) *LoCA 2009*. LNCS, vol. 5561, pp. 70–87. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-01721-6\\_5](https://doi.org/10.1007/978-3-642-01721-6_5)
4. Yigitoglu, E., Damiani, M.L.: Privacy-preserving sharing of sensitive semantic locations under road-network constraints. In: *Proceedings of International Conference on Mobile Data Management*, pp. 186–195 (2012)
5. Li, M., Qin, Z., Wang, C.: Sensitive semantics-aware personality cloaking on road-network environment. *Int. J. Secur. Appl.* **8**(1), 133–146 (2014)
6. Kido, H., Yanagisawa, Y., Satoh, T.: An anonymous communication technique using dummies for location-based services. In: *Proceedings of IEEE International Conference on Pervasive Services, ICPS (2005)*
7. Beresford, A.R., Stajano, F.: Mix zones: user privacy in location-aware services. In: *Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops (PerCom 2004 Workshops)*, pp. 127–131 (2004)

8. Gruteser, M., Grunwald, D.: Anonymous usage of location based services through spatial and temporal cloaking. In: Proceedings of the 1st International Conference on Mobile Systems, Applications and Services (MobiSys 2003), San Francisco, California, pp. 31–42 (2003)
9. Mohamed, F.M., Chow, C.Y., Walid, G.A.: The new casper: query processing for location services without compromising privacy. In: Proceedings of 32nd International Conference on Very Large Data Bases, pp. 763–774. ACM Press (2006)
10. Gedik, B., Liu, L.: Protecting location privacy with personalized k-anonymity: architecture and algorithms. *IEEE Trans. Mob. Comput.* **7**(1), 1–18 (2008)
11. Kalnis, P., Ghinita, G., Mouratidis, K.: Preventing location-based identity inference in anonymous spatial queries. *Proc. IEEE Trans. Knowl. Data Eng.* **19**(12), 1719–1733 (2007)
12. Ardagna, C.A., Cremonini, M., Damiani, E., De Capitani di Vimercati, S., Samarati, P.: Location privacy protection through obfuscation-based techniques. In: Barker, S., Ahn, G.-J. (eds.) DBSec 2007. LNCS, vol. 4602, pp. 47–60. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-73538-0\\_4](https://doi.org/10.1007/978-3-540-73538-0_4)
13. Ghinita, G., Kalnis, P., Khoshgozaran, A., Shahabi, C., Tan, K.L.: Private queries in location based services: anonymizers are not necessary. In: Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data (SIGMOD 2008), Vancouver, Canada, pp 121–132 (2008)
14. Bamba, B., Liu, L., Pesti, P.: Supporting anonymous location queries in mobile environments with Privacy Grid. In: Proceedings of the 17th International Conference on World Wide Web, New York, pp. 237–246 (2008)
15. Damiani, M.L., Bertino, E., Silvestri, C.: The PROBE framework for the personalized cloaking of private locations. *ACM Trans. Data Priv.* **3**(2), 123–148 (2010)
16. Damiani, M.L., Bertino, E., Silvestri, C.: Protecting location privacy against spatial inferences: the PROBE approach. In: ACM SPRINGL 2009 (2009)
17. Damiani, M.L., Silvestri, C., Bertino, E.: Fine-grained cloaking of sensitive positions in location-sharing applications. *IEEE Pervasive Comput.* **10**(4), 64–72 (2011)
18. Byoungyoung, L., Jinoh, O., Hwanjo, Y.: Protecting location privacy using location semantics. In: Proceedings of the 17th ACM SIGKDD, pp. 1289–1297 (2011)

# Smartphone Bloatware: An Overlooked Privacy Problem

Haroon Elahi, Guojun Wang<sup>(✉)</sup>, and Xu Li

School of Computer Science and Educational Software, Guangzhou University,  
Guangzhou 510006, China  
csgjwang@gzhu.edu.cn

**Abstract.** Smartphone bloatware or pre-installed applications are included in smartphones by the Original Equipment Manufacturers and the telecom carriers partly to remain ahead of competitors and partly to subsidize the costs of smartphones through app developers. Such applications are also known to be used for maintaining device control and for collecting user data. There are a number of other privacy and security costs associated with these freely available apps. However, costs and benefits of these apps for the smartphone users have not been assessed. In this paper, we provide findings of a user-study that was conducted to investigate the practical utility of bloatware in personal and professional lives of users. We surveyed a representative sample of 53 smartphone users for this purpose. We try to learn their dependence on such applications, identify frequently used and computationally wanted applications, and assess relevant user expectations. We also review the privacy and security implications of such apps and discuss whether these are such valuable assets that users should surrender control of their personal data and privacy, in exchange. We introduce the concepts of ‘dormant profiles’, and ‘on-demand and flexible’ apps to reduce the privacy problems introduced by these apps, and to put users in control of their data and privacy.

**Keywords:** Personal data privacy · Data over-collection  
Smartphone bloatware · Security · Privacy enhancing models

## 1 Introduction

‘Feature richness turns into bloat when there are more features than you want to use’ [1]. Desktop applications have long been criticized for having too many features, introducing complexity and usability issues, and burdening the computing resources [3, 4, 10]. Despite long running criticism, this practice has not stopped, and rather has been adopted by smartphone manufacturers or Original Equipment Manufacturers (OEMs). However, as against the problem of overly-featured software applications in the personal computers (PCs), we face an aggravated situation in smartphones. Nowadays, powerful smartphones are being manufactured, with larger memories, and faster processors, and come packed with a large number of so-called utility apps, many of which might never be used by the smartphone users [2, 4, 6]. Users are, however, bound to live with them with little or no choice at all. Thus the problem has

transformed from bloated application software in PCs to bloated devices – the app swarmed smartphones.

Historically, bloatware apps have been included in the smartphones partly to remain ahead of competitors and partly to subsidize the costs of smartphones through app developers, for an increase in the profits [2]. However, recently developed data mining methods and techniques like user-profiling have introduced new technological and business opportunities for app developers and OEMs. This has resulted in data-collection marathons among third-party app providers, OEMs, telecommunication operators and other related parties [5, 12]. Such practices are termed as data over-collection when these data collections exceed the functional needs of apps and services [26]. The richness and sensitive nature of smartphone data make data over-collection a serious problem and a serious privacy threat [18, 26]. OEMs, third-party app providers, telecommunication operators, and malicious parties have used bloatware apps to over-collect user data [8, 22].

Bloatware has also introduced numerous security and trust issues in smartphones. For example, these apps can also enable security breaches through intra-application information sharing, and because of security holes in their large code base [7, 14, 15]. Similarly, it is natural that users demand control over their data and resources [5]. However, users have very little choice and control over selection or rejection of such apps, which breeds mistrust. Furthermore, the presence of a large number of apps introduces complexity of use in smartphones, and uncertainty, feature- fatigue and cognitive loads among smartphone users [10]. Smartphone users have to deal with a large number of privacy settings of smartphone apps and often ignore bloatware apps assuming that they are system apps [26]. Hence, such apps can be used for unobtrusive data over-collection. Such possibilities have introduced the ‘bloat attacks’ where malicious parties introduce malware in smartphones and tablets before they are sold to users [6]. We provide a detailed account of security and privacy threats of bloatware in Sect. 8.

Despite having serious privacy, security and usability implications, smartphone bloatware has gained little attention from the research community in the recent past. Although, many studies have recognized them as security and privacy threats in smartphones [8, 19–21], OEMs are continuously packing their smartphones with these apps. In this paper, we compare costs and benefits of such apps for the smartphone users in order to learn whether these apps are such valuable assets that security and privacy of user personal data should be compromised to retain them. In order to achieve this goal, we use the findings of our user-study that investigates the practical utility of such apps in the personal and professional lives of the users, and review the potential privacy and security issues caused by these apps. We also make propositions that can help in dealing with such issues and threats, and to put users in control of their data and privacy.

## 2 Contributions

The contributions of this study are multi-fold. This study makes following primary contributions, with the aim to evaluate the practical utility of smartphone bloatware for users and the problems posed by such apps to justify alternative models for improved usability, secure devices, and improved privacy control:

1. We present the findings of a user-study that investigates the practical utility of smartphone bloatware in personal and professional lives of smartphone users.
2. We review the usability, privacy, and security problems introduced by such apps, assess them against the findings of our study and try to evaluate costs and benefits of bloatware apps.
3. Our study focuses on privacy and security implications of smartphone bloatware as against few past studies that investigated the usability issues.
4. We introduce the concepts of dormant profiles, and on-demand and flexible apps to overcome the relevant problems and to put users in control of their data and privacy.

The rest of this paper is structured as follows. In Sect. 3, we provide a review of related work. Section 4 explains study design. Section 5 discusses the methods employed to conduct this study. Section 6 provides and insights into results of our study. Section 7 provides details of missing responses. Section 8 provides analysis of the findings. Section 9 reviews major privacy risks of smartphone bloatware. In Sect. 10, we discuss privacy implications of smartphone bloatware in the light of our results. Section 11 proposes alternatives to the existing practices of bloating smartphone with large number of apps. Section 12 reports limitations of this study, and Sect. 13 contains concluding remarks and future work.

### 3 Related Work

Kaufman and Weed [1], reported different aspects of bloat almost three decades ago. They suggested that although a device with many features could be used to accomplish many things, this feature richness would turn into bloat if there were more features than needed. They explained that bloat would result in an inability to figure out how to accomplish a task, presence of unnecessary information and complexity, visual clutter, interface design anomalies, excessive learning times, and system fragility. Baecker et al. [3], conducted a study to investigate complexity MS Word in terms of functional complexity, data complexity, and the complexity of learning about software. They discovered that the probed application was bloated as it had a large number of unused functions that users never utilized. They suggested that the complex functionality of software systems should be encountered by personalization. McGreenre [4], introduced the notions of ‘objective’ and ‘subjective’ bloat in their research. They termed functions of an application that were used only by a few users as ‘objective bloat’ and those used by certain classes of users and not by others as ‘subjective bloat’. They urged elimination and relocation such features. They also suggested that personalization was a solution to reduce the system complexity resulting from too many features. They recommended the use of digital personae, and social roles and activities, etc. to achieve these goals.

McDaniel [2], explored the problem of bloatware in smartphones and found that an increased competition, a constant demand for latest and greatest phone hardware, and increased costs of newly introduced network technologies were the major factors that led smartphone manufacturers to seek subsidies through bloatware. They pointed out that bloatware had opened doors to loss of privacy and security and suggested

investigations of security issues introduced by bloatware. Kim et al. [20], conducted a study that provided insights about how the presence of bloatware on the smartphone might backfire. They discovered that it generated a negative user experience and users associated negative attributes such as product sluggishness with it. One important study that presented the challenges to the safe use of smartphones was conducted by Vrhovc [8]. This study found that although bloatware was meant to provide user-friendliness and improved interaction with various services, it used large amounts of device resources, engaged in personal data collection and introduced security vulnerabilities.

Spensky et al. [21], surveyed numerous components of mobile devices, giving particular attention to those that collected, processed, or protected users' private data. They identified bloatware as one of the threats to user privacy. They explained that bloatware could even open door to security vulnerabilities in the operating system. In their study of privacy issues in smartphones, Lee [22], discussed the privacy breaches in smartphones. One of the major reported breaches was by Carrier IQ that affected more than 150 million smartphones. Carrier IQ is a bloatware app that is used by mobile carriers to collect user data [27].

## 4 Study Design

This study evaluates the practical value of bloatware apps for smartphone users, identifies the privacy and security problems caused by such apps, and thereafter tries to synthesize whether the price that users pay to acquire such apps is worth having them in the first place.

In order to achieve this goal, we use a conduct a user-study to learn the practical value of bloatware apps in personal and professional lives of users, and conduct literature review to find privacy and security threats of bloatware apps. By practical value, we mean the assessment of satisfaction and attractiveness after a user acquires and uses such apps [9, 31]. Studies show that users believe in more are better, and factors like perceived risks, perceived benefits and trust play a primary role while choosing apps [1, 30]. They acquire apps when perceived benefits of such apps exceed perceived risks. Bloatware apps are among characterizing features of smartphones that manufacturers use to compete in the market and attract a larger customer base [2, 19, 32]. We assume that such features must be designed keeping in view user requirements, as well as, competitors' feature apps. Thompson et al. [10] suggest that mobile manufacturers include new features to provide greater functionality to smartphone users. Keeping in view these research findings we proceed while following these assumptions:

1. Users derive large utilities in their private and professional lives through functionalities provided by bloatware apps.
2. Users conduct necessary research before buying their smartphones, use bloatware apps frequently, and hence are closely familiar with them.
3. As users have different academic and professional domains and their requirements differ, they use different bloatware apps.

4. Keeping in view the variety of functionalities provided by bloatware apps in smartphones, users do not need to install a large number of third-party apps.
5. Users want advance feature apps to be shipped with the smartphones.

We identify security and privacy risks of bloatware. We compare the results of our user-study with the potential threats of bloatware and synthesize whether it is worth losing control over personal data and privacy to have such free utility apps.

## 5 Method

### 5.1 Survey Content

A user survey was designed to collect the data needed for evaluating practical value of bloatware apps for smartphones users. We want to examine whether users include use of bloatware apps among the major uses of their smartphones; how many and which bloatware apps do they use on regular basis; whether people find such apps useful in their professional activities; and whether professional different backgrounds imply using different bloatware apps. We also want to find out user-perceived must-have bloatware apps, again to explore their relation with user backgrounds and whether the presence of these apps reduces users' dependence on third-party apps from app markets. In the end, we want to know how users will react if there are no preinstalled apps in new smartphones. We included following questions in our survey:

1. What is the general use of your smart phone?
2. Does your smart phone help you in accomplishing your professional duties?
3. How many pre-installed smart phone apps do you use?
4. Please name three pre-installed smart phone apps that you use frequently.
5. How useful are the pre-installed smart phone apps for you?
6. Do the pre-installed smart phone apps help you in your professional life?
7. Please name five pre-installed apps that must be present in the phone when you buy it?
8. How many smart phone apps did you need to install after buying your smart phone?
9. If your mobile-phone comes with no pre-installed apps, how will it affect you?
10. Please describe briefly one must have pre-installed app, that would be really useful for you.

### 5.2 Survey Method

We used Google Forms to setup an online survey questionnaire. We requested our social media contacts to participate in the study. The response was slow, therefore we needed to send personal messages. In some cases we called our contacts to remind them.



### 5.3 Sample Information

Being an online survey, cross-border audiences were involved. The social, cultural and professional backgrounds of the participants were quite diverse. Despite having a large number of social-network contacts, only 53 people volunteered to participate in the study. Influential contacts were also requested to further promote the survey among their contacts.

## 6 Results

In this section, we present our findings based upon the data collection through our online survey.

### 6.1 Demographics

We received 53 responses for our survey. Out of 53 respondents, 29 were males and 24 were females. Respondents belonged to a wide range of age groups. Two participants were between 18 and 20 years of age; fifteen were between 21 and 25 years of age; seventeen were between 26 and 30 years of age and nineteen respondents were above 30 years of age. 81.1% users possessed Android phones, 15.1% were iPhone users, 1.9% were Windows Phone users, and 1.9% had other phones. The respondents were located in twelve different countries and for the majority of the respondents, English was not their first language. Participants of our survey included professionals from technical and non-technical domains, students, and house-wives (Table 1).

In the following subsections, we provide the responses to the questions included in our survey.

#### 1. *What is the general use of your smart phone?*

*Findings:* The given options included: connecting with contacts through calls, connecting with contacts through messaging, using social networks, playing games, internet browsing, availing different services through pre-installed apps, availing different services through third-party apps, listening music, using navigation services in the phone, and all of above. Only 20.8% respondents acknowledged that they availed different services through pre-installed apps. Whereas, 37.7% respondents marked that one of the major uses of their smartphones was to derive utility through the use of third-party apps.

#### 2. *Does your smartphone help you in accomplishing your professional duties?*

*Findings:* The majority of respondents (i.e., 88.7%), said that smartphones helped them in accomplishing their professional duties while 11.3% denied such role of smartphones in their professional lives.

#### 3. *How many pre-installed smart phone apps do you use?*

*Findings:* 50.9% respondents said they use less than five apps in this category, 35.8% said they used six to ten apps, 9.4% said they used between eleven to fifteen apps and 3.8% said they used between sixteen to twenty such apps.

**Table 1.** Professional backgrounds

Profession	Frequency
Student	13
School teacher	10
University teacher	6
Researcher	3
Banking	2
Big Data engineer	1
Software engineer	2
House wife	3
Business owner	1
IT professional	2
Sales manager	1
Buyer	1
Software developer	2
Designer	1
Psychologist	1
Management	1
Other	1
Did not disclose	2

4. *Please name three pre-installed smart phone apps that you use frequently.*

*Findings:* We received only 48 responses against this question. One of the questions was treated wrong, as the respondent provided names of special purpose third-party apps that are not shipped with smartphones. In order to understand which types of apps were most frequently used by the respondents, we grouped apps into categories. We used the app categories used by Google, for this purpose. Responses showed that most frequently used apps were related to communication that was used by 77% of the users. Productivity, business, and social apps were used by 29.17, 29.17% and 22.9% users respectively. Photography, maps & navigation, entertainment, and tools were used by 22.9%, 18.75%, 18.75% and 14.58% of the users. Music & audio 8.34% (4/48), News & Magazines 4.17% (2/48), Weather 2.1% (1/48) were the apps used by least used apps, mentioned by 8.34%, 4.17%, and 2.1% users respectively.

5. *How useful are the pre-installed smart phone apps for you?*

*Findings:* Respondents were given four options, namely: very useful, quite useful, somewhat useful, and not useful. 54.7% respondents said such apps were very useful, 26.4% chose quite useful, 17% said somewhat useful and 1.9% said not useful.

6. *Do the pre-installed smart phone apps help you in your professional life?*

*Findings:* When we asked them how many pre-installed apps helped them in their professional lives, 7.5% respondents said all of them, 37% said most of them, 49.1% said some of them and 5.7% said none of them.

7. *Please name 5 pre-installed apps that must be present in the phone when you buy it?*

*Findings:* Only 40 responses were received for this question. Two responses were invalid. Hence, in total 38 valid responses. We summarized the responses using app categories used by Google [33]. The responses were as follows: Communication: 76.3% (29/38), social: 44.74% (17/38), entertainment: 15.7% (6/38), tools: 10.5% (4/38), productivity: 39.47% (15/38), music & audio: 13.15% (5/38), photography: 34.2% (13/38), Maps & Navigation: 23.68% (9/38), news & magazines: 2.6% (1/38), business: 34.2% (13/38), weather: 2.6% (1/38), games: 7.89% (3/38), personalization: 7.89% (3/38), and books & reference: 2.6% (1/38).

8. *How many smart phone apps did you need to install after buying your smart phone?*

*Findings:* 32.7% of the respondents said they had installed five or less apps, 32.7% said they had installed six to ten third-party apps, 17.3% said they had installed between eleven to fifteen such apps, 13.5% respondents installed between sixteen to twenty apps, and 3.8% respondents had to install more than twenty third-party apps to meet their functional needs.

9. *If your mobile-phone comes with no pre-installed apps, how will it affect you?*

*Findings:* 22.6% respondents said that they would be totally disappointed, 35.8% said they would be somewhat disappointed, 30.2% said it did not matter and 11.3% said they would be happy to have smartphones with no pre-installed apps at all.

10. *Please briefly describe one must have pre-installed app, that would be really useful for you.*

*Findings:* We included one open ended question regarding one dream app that users would like to have pre-installed in their smartphones. We received 40 responses for this question. One response was invalid, one respondent wanted no apps, and one respondent was happy with available apps. Apps asked for in the valid responses fell in the following categories: communication: 7, social: 4, tools: 5, productivity: 2, maps and navigation: 2, Business: 2, Health and Fitness: 4, Language and Dictionary: 2. Furthermore, one respondent wanted a security app, one respondent wanted an app that would tutor on how to use other apps, and one respondent asked for an auto-answering app.

## 7 Missing Values

Two respondents did not disclose their professions. One respondent did not supply the country of location, and five did not provide names of three frequently used bloatware apps. Only 40 respondents provided names of five must have bloatware apps. Number of third-party app installation was not provided by one respondent. Only 40 responses were received for the last app where description was needed to learn whether users associated their computing needs with their professions and whether they had any exceptional needs.

## 8 Analysis

In this section, we discuss the findings of our study while looking at our assumptions.

1. *Users derive large utilities in their private and professional lives through functionalities provided by bloatware apps.*

The results of our study suggest that only a 20.8% of the users consider deriving utility from bloatware as a primary use of their smartphones. A large number of the users (88.7%) agreed that smartphones help them in their professional lives. However, when it comes to bloatware apps, almost half of them (49.1%) said that only some of such apps are helpful and a small number of the users (5.7%) even denied seeking any help from such apps in their professional lives. Furthermore, as the results of our study show, only 22.6% respondents expressed total disappointment in case smartphones come with no bloatware apps. 30.2% of the respondents said that for them it didn't matter whether such apps were included in the new smartphones or not; and 11.3% of the participants said they would be happy in such case.

2. *Users conduct necessary research before buying their smartphones, use bloatware apps frequently, and hence are closely familiar with them.*

In contrast with our assumption, 50.9% of the users said they used less than five bloatware apps. 35.8% users said they used between six to ten bloatware apps. We tried an alternative route by asking users the most frequently used bloatware apps. The responses put Calls, Messaging and Social Media apps on the top. Only 29.17% of the users used business apps. Mostly these business apps included email and calculators. Photography and navigation apps were used by one fifth of users respectively.

3. *As users have different academic and professional domains and their requirements differ, they use different bloatware apps.*

When asked to list five bloatware apps that they would like to come with their smartphones, we received only 71% valid responses. Again, communication apps (call and messaging) were on the top, followed by social media, productivity, business and photography apps.

4. *Keeping in view the variety of functionalities provided by bloatware apps in smartphones, users do not need to install a large number of third-party apps.*

37.7% of the users marked availing services through apps acquired from open mobile markets as one of the primary uses of their smartphones. 34.7% of the users installed more than ten third-party apps. For 32.7% of the users the number of such apps was between six and ten. In contrast, 50.9% of the users said they used less than five bloatware apps. 35.8% users said they used between six to ten bloatware apps.

5. *Users want advance feature apps to be shipped with the smartphones.*

In the end, we wanted to know whether users want advanced featured apps to be shipped with phones. So we asked users to describe one must-have app that they would like to have. The responses were very diverse and mostly resembled what they have on

their smartphones. Only exceptional responses included health and fitness apps, language and dictionary features, and security features. None of the users put apps that could be associated with their professional domains.

## 9 Privacy Risks

In desktop applications, the concept of bloatware mainly comprised of applications having excessive number of features. Such properties mainly caused usability issues [1, 3, 4]. Although concerns were also raised about privacy violations of such applications, we do not find much research around such problems [23]. However, in case of smartphones, the nature of bloatware changed. The excessive number of features turned into a large number of creepy apps. In this section, we provide the major privacy risks caused by bloatware apps in smartphones.

### 9.1 Lack of Control

Bloatware is used by original equipment manufacturers (OEMs) to maintain control over smartphones [21]. In parallel, an increased control of device by OEMs and telecom operators means less control for users. Smartphone bloatware features this lack of control for users who cannot choose or remove such apps. The large number of privacy invasive features also introduce the creep factor among users [2, 21, 22]. Such characteristics make these apps unwanted and undesirable for the users.

### 9.2 Negative User Experience

This research provides insights about how providing pre-installed applications on smartphone may backfire; consumers may think that bloatware makes their product sluggish. Marketers might want to be cautious to communicate about many installed features. In order to overcome the limitations of scenario-based experiment, field studies with actual stimuli using a real smartphone are suggested to confirm our findings in the future. Moreover, users don't have any control over such apps. They cannot uninstall the apps.

### 9.3 Broken Expectations

Studies show that even though users buy feature-rich gadgets, when they are lost into the features and the devices are hard to use, it reduces their product satisfaction [10]. The broken expectations, breed untrustworthiness [24]. Such untrustworthiness cannot be afforded in smartphones, keeping in view the sensitive nature of information and operations involved. This loss of trust raises privacy concerns among users and subsequently either they do not install apps where they have privacy concerns or uninstall them even after installation [2, 9]. This loss of trust, resulting from privacy concerns, has been identified as one of the hurdles in acceptance of pervasive technologies [6, 8].

## 9.4 Instable Operating System

While, the combinatorial nature of smartphone related technologies also cautions to operationalize features beyond synergies between technical specifications, bloatware apps introduce misconfigurations and excessive resource consumption [19]. Bloatware apps not only introduce creep factor and fears among users about stability of their devices and device platforms but practically damage operating systems. Kaufman and Weed suggested that too many unknown features could cause users to worry about whether their actions were correct and if the system was liable to break [1]. Continuous updates and efforts to uninstall these apps by the users cause configuration damage and leave security holes. These holes can be used by the attackers to carry out root attacks.

## 9.5 Bloatware App Updates

There are two kinds of privacy issues caused by bloatware updates. First, are caused by excessive updates and others are introduced by delayed updates. While large number of bloatware apps are not used by smartphone users, they are updated regularly which takes long time and consume large part of storage space. This is particularly important in the case of smartphones where resources are scarce. Such phenomena have been termed as *Zombieware* in the literature due to their characteristics of eating resources [23]. Often, users have concerns regarding what these updates actually do and that the app providers do not actually drain data to their servers in the guise of updates [1]. Second type of privacy issues are caused by security holes resulting from slow updates. Especially in the case of smartphones, it has been discovered that third-parties take longer times to update their applications, often leaving vulnerabilities unattended for long periods of time [21]. Such vulnerabilities enable privacy violations.

## 9.6 Larger Attack Surface

With the large number of bloatware apps included in smartphones by OEMs and telecommunication operators, users' dependence on third-party apps should reduce. However, studies show that, on average, non-rooted devices have twice the number of apps as compared with rooted devices [34]. This means larger code base, larger number of bugs, more Inter-Component Communication leaks and larger pressure on limited resources of smartphones [28].

## 9.7 Too Many Privacy Settings

Users are in general incapable and illiterate when it comes to privacy settings [26]. Users can perceive the bloatware apps as system apps and ignore their privacy settings. Large number of unattended privacy settings can cause continuous flow of personal data to third-party servers.

This is not limited to end-users. App developers also need to take care of lots of sensitive privacy settings such as careful selection of permissions for their apps. App developers need to be careful while designing and developing smartphone apps. A little

carelessness can cause big privacy implications for personal user data. Such as matching permissions can enable smartphone apps in Android phones to access data collected by other apps [17].

### 9.8 Collusion Attacks

Bloatware apps are known to send user data to third-party servers [23]. Memory usage mechanism in android environments makes it possible for different apps to share memory content and thus access information that is beyond their functional scope [16, 17, 29]. Most of the bloatware apps are personal utility apps and deal with personal data. Inter-component information leaks are a known problem and thus such apps enable exposure of personal data to other apps [15].

### 9.9 Third-Party APIs

Smartphone apps use third-party APIs to perform different tasks. Such APIs gain access to data and resources that their parent applications have [13, 21]. Thus they collect personal data and send to third-party servers without any knowledge of the users. This data is used for purposes unknown to users and without their permissions. The larger the number of apps, the larger is the number of such back-doors.

## 10 Discussion

After looking at the results of our study, when we look at privacy and security threats caused these apps, their costs and benefits are not comparable. Lack of control over device resulting from inability to remove such apps, an overall negative user experience leading to loss of trust, broken privacy expectations, instable operating system, memory and processing consumption leading to memory leaks and battery drainage, mistrust resulting from updates or operating system fragmentation from no updates at all, cognitive loads resulting from too many privacy settings, collusion attacks, and privacy invasion by third-party APIs are not trivial to be ignored.

Additionally, as most of such apps are utility apps and deal with personal data, they are known to send these data to third-party servers and to telecommunication operators [11, 12, 22, 23]. The results of our study show that practical utilities of such apps are not outstanding, and users depend more on third-party apps acquired from app markets. Keeping in view the low practical utility of such apps as suggested by the results of our study, and inability of users to remove such apps, their data collection cannot be justified and can be treated as data over-collection. We believe that data over-collection is a serious problem and if not handled effectively, it will pose bigger challenges in future technologies. We believe that although there is a need for pre-installed apps to provide essential functionality to the users, bloating mobile devices with such apps is not appropriate. Keeping in view the results of our study and privacy threats caused by such apps, we suggest that their privacy implications are much bigger than the utility that users derive from them.

In the following section we provide alternative models that can be used to encounter the privacy and security issues identified in this paper.

## 11 Proposed Alternatives

As the results of our study show, the computational needs of users are quite varied and unpredictable. Even the users with similar computing requirements choose different apps. For example, Facebook is a very famous social networking app in most of the countries, however, in China WeChat is more popular and preferred by the users. Therefore, we argue that swarming smartphones with large number of bloatware is not a good model to provide functionality to the users. As discussed in the previous sections, it introduces usability, privacy, trust and security issues. We propose the following alternatives to the existing fashion of shipping a large number of apps with every new smartphone. These alternatives can help to encounter the problems posed by the smartphone bloatware.

### 11.1 On-demand Apps

Instead of pre-installing large number of apps, we suggest that OEMs should make the free apps available on product sites and users should be able to install the apps that they need without extra costs. Users should be able to evaluate the privacy and performance implications of such apps. The information regarding such apps can be provided in the product documentation provided with the sold devices or through push notices. This will let users choose what they want and what they don't want. In addition to increased control over app choice, users will be able to accept and reject apps based on their privacy implications.

### 11.2 Flexible Apps

Taking a step ahead we suggest use of flexible apps for feature-level control. To avoid feature fatigue, users should be able to select the features that they need or want. Users should be transparently told about the resources and data needed for such extra features. The current descriptions of the apps on Google play store do not take into account the privacies. The details of functionality and features or the apps are not comprehensive and related privacy implications are totally missing. Users should have detailed access to such details before they choose apps and their features.

### 11.3 Dormant Profiles Approach

The use of user profiles is not new in computer science. In a recent study, Liu et al. [35] suggested the use of privacy profiles to model users' privacy preferences in order to reduce burdens caused by large number of privacy settings in smartphone apps. Moreover, it is known that different users value different products differently depending on the features of such products [25]. Hence, there is a possibility to learn users' app selection preferences. OEMs can classify their potential consumers and can create app



profiles for each class based on historical data. Such user classes can be identified following the functional needs and app usage patterns of users. We call them dormant profiles because only users should have control over choosing a profile by looking at the set of apps provided in different profiles. They should have essential details of functions and privacy implications of these apps. In a newly purchased smartphone, only the links to dormant profiles should be provided to the user, who in turn, can activate desired profile. In this model, none of the smartphones should contain the actual code and all apps in a selected profile should be downloaded after users' approval.

## 12 Limitations

There are a number of limitations in this study. First of all, our user-study has certain limitations. The participants of this study are well educated. However, keeping in view the nature of problem this factor does not bias the results in our favor. Second, the participants of this study come from twelve different countries. Cultural differences and infrastructural differences can have their impact. But, we do not see any exceptional variations in responses. Furthermore, our research addresses mainly android phones. We have not covered technical issues faced in other smartphones available in the market. Still, when it comes to generic privacy implications such smartphones are no exception and our results can be used with extra verification. It is also important to acknowledge that to review the privacy and security threats, we have mainly depended upon the findings in literature. However, this does not reduce the importance of our research as the referred studies were published at different times and were published in different reputed forums. We, however, believe that there is a need to conduct system level investigations to verify the data over-collection practices of bloatware apps.

## 13 Conclusion and Future Work

Android security and resulting privacy has been in the spotlight ever since the first Android Phone was launched in 2008 and we see extensive research efforts put in these areas. However, little efforts have been paid to change the models that enable such violations and threats. Bloating smartphones with large number of personal utility apps is one of such models. In this paper, we have explained in detail that issues caused by bloatware apps in smartphones are far beyond usability. We tried to assess costs and benefits of bloatware apps by evaluating their practical utility for users. In this regard, we presented the results of a user-study that we conducted to evaluate the utility of bloatware apps for smartphone users. The results of our study suggest that majority of the smartphone users use only a small number of bloatware apps in their daily lives. Rest of such apps that are never used by users work as sleeper cells, thwarting the privacies of users by continuously collecting personal information and sending it to third-party servers, violating the privacies of users. Such apps also cause instability in the operating system through misconfigurations and continuous updates. Such apps challenge the control of users over the devices and cause root attacks after voluntary

rooting of devices by users to regain device control. We believe that these are too many and too serious privacy hazards caused by these apps. We argue that the privacy costs that users pay to acquire and keep such apps are not reasonable and introduce a need to change the existing culture of swarming smartphones with bloatware apps. As a first step, we introduced the concepts of dormant profiles, and on-demand and flexible apps that can be successfully used to counter the privacy problems introduced by bloatware and put users back in control of their personal data and privacy.

In future, we intend to conduct thorough analysis of app data to verify privacy invasion practices of such apps. We also plan to provide detailed implementation models of dormant profiles, and on-demand and flexible apps.

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation of China under Grants 61632009 & 61472451, in part by the Guangdong Provincial Natural Science Foundation under Grant 2017A030308006 and High-Level Talents Program of Higher Education in Guangdong Province under Grant 2016ZJ01.

## References

1. Kaufman, L., Weed, B.: Too much of a good thing? Identifying and resolving bloat in the user interface. In: Conference Summary on Human Factors in Computing Systems - CHI 1998, vol. 30, pp. 207–208 (1998). <https://doi.org/10.1145/286498.286693>
2. McDaniel, P.: Bloatware comes to the smartphone. *IEEE Secur. Priv.* **10**(4), 85–87 (2012). <https://dl.acm.org/citation.cfm?id=2377522>
3. Baecker, R., Booth, K., Jovicic, S., McGrenere, J., Moore, G.: Reducing the gap between what users know and what they need to know. In: Proceedings on the 2000 Conference on Universal Usability - CUU 2000, pp. 17–23 (2000). <https://doi.org/10.1145/355460.355467>
4. McGrenere, J.: Bloat: the objective and subject dimensions. In: CHI 2000 Extended Abstracts on Human Factors in Computing Systems, pp. 337–338 (2000). <https://doi.org/10.1145/633292.633495>
5. Van Kleek, M., Liccardi, I., Binns, R., Zhao, J., Weitzner, D.J., Shadbolt, N.: Better the devil you know: exposing the data sharing practices of smartphone apps. In: CHI (2017). <https://doi.org/10.1145/3025453.3025556>
6. <https://goo.gl/kEKs3y>. Accessed 6 June 2017
7. Shu, J., Zhang, Y., Li, J., Li, B., Gu, D.: Why data deletion fails? A study on deletion flaws and data remanence in Android systems. *ACM Trans. Embed. Comput. Syst.* **16**, 1–22 (2017). <https://doi.org/10.1145/3007211>
8. Vrhovec, S.L.R.: Safe use of mobile devices in the cyberspace. In: 2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics, pp. 1397–1401 (2016). <https://doi.org/10.1109/MIPRO.2016.7522358>
9. Yu, P., Yeung, C.A.: App mining: finding the real value of mobile applications. In: Proceedings of WWW 2014 Companion, pp. 417–418 (2014). <https://doi.org/10.1145/2567948.2577297>
10. Thompson, D.V., Hamilton, R.W., Rust, R.T.: Feature fatigue: when product capabilities become too much of a good thing. *J. Mark. Res.* **42**, 431–442 (2005). <https://doi.org/10.1509/jmkr.2005.42.4.431>

11. Chen, X., Zhu, S.: DroidJust: automated functionality-aware privacy leakage analysis for Android applications. In: *WiSec 2015*, 22–26 June 2015, New York City, NY, USA, pp. 5:1–5:12 (2015). <https://doi.org/10.1145/2766498.2766507>
12. Liu, X., Wang, W., Liu, J.: POSTER: the popular apps in your pocket are leaking your privacy. In: *Proceedings of the CCS 2015*, 12–16 October 2015, Denver, Color. USA, pp. 1653–1655 (2015). <https://doi.org/10.1145/2810103.2810127>
13. Reaves, B., Bowers, J., Gorski III, S.A., Anise, O., Bobhate, R., Cho, R., Das, H., Hussain, S., Karachiwala, H., Scaife, N., Wright, B., Butler, K., Enck, W., Traynor, P.: \* droid: assessment and Evaluation of Android application analysis tools. *ACM Comput. Surv.* **49**, 1–30 (2016). <https://doi.org/10.1145/2996358>
14. Witten, B., Landwehr, C., Caloyannides, M.: Does open source improve system security? *IEEE Softw.* **18**, 57–61 (2001)
15. Li, L., Bartel, A., Bissyandé, T.F., Klein, J., Le Traon, Y., Arzt, S., McDaniel, P.: IccTA: detecting inter-component privacy leaks in Android apps. In: *Proceedings of the 37th International Conference on Software Engineering*, vol. 1, pp. 280–291 (2015). <https://doi.org/10.1109/ICSE.2015.48>
16. Lu, L., Li, Z., Wu, Z., Lee, W., Jiang, G.: CHEX: statically vetting Android apps for component hijacking vulnerabilities. In: *ACM Conference on Computer and Communications Security*, pp. 229–240 (2012). <https://doi.org/10.1145/2382196.2382223>
17. Dimitriadis, A., Efraimidis, P.S., Katos, V.: Malevolent app pairs: an Android permission overpassing scheme. In: *Proceedings of CF 2016*, pp. 431–436 (2016). <https://doi.org/10.1145/2903150.2911706>
18. Gisdakis, S., Giannetsos, T., Papadimitratos, P.: Android privacy C(R)ache: reading your external storage and sensors for fun and profit. In: *2nd MobiHoc International Workshop on Privacy-Aware Mobile Computing PAMCO 2016*, pp. 1–10 (2016). <https://doi.org/10.1145/2940343.2940346>
19. Han, Q., Cho, D.: Characterizing the technological evolution of smartphones. In: *Proceedings of the 18th Annual International Conference on Electronic Commerce: e-Commerce in Smart connected World - ICEC 2016*, pp. 1–8 (2016). <https://doi.org/10.1145/2971603.2971635>
20. Kim, S., Choe, Y., Lee, Y., Kim, S., Choe, Y., Users, S., Puntoni, S.: How heavy is your smartphone? Imaginary weight perception of smartphone users and its impact on product evaluation. *Adv. Consum. Res.* **44**, 512–513 (2016)
21. Spensky, C., Stewart, J., Yerukhimovich, A., Shay, R., Trachtenberg, A., Housley, R., Cunningham, R.K.: SoK: privacy on mobile devices – it’s complicated. In: *Proceedings on Privacy Enhancing Technologies*, pp. 96–116 (2016). <https://doi.org/10.1515/popets-2016-0018>
22. Lee, N.: Smartphones and privacy. In: Lee, N. (ed.) *Facebook Nation*. Springer, New York (2013). [https://doi.org/10.1007/978-1-4614-5308-6\\_3](https://doi.org/10.1007/978-1-4614-5308-6_3)
23. Neumann, P.G.: Risks to the public. In: *ACM SIGSOFT*, vol. 36, pp. 17–23 (2011). <https://doi.org/10.1145/1988997.1989002>
24. Neumann, P.G.: Risks of untrustworthiness. In: *Proceedings of the Computer Security Applications Conference ACSAC*, pp. 321–326 (2006). <https://doi.org/10.1109/ACSAC.2006.45>
25. Koski, H., Kretschmer, T.: Innovation and dominant design in mobile telephony. *Ind. Innov.* **14**, 305–324 (2007). <https://doi.org/10.1080/13662710701369262>
26. Li, Y., Dai, W., Member, S., Ming, Z., Qiu, M., Member, S.: Privacy protection for preventing data over-collection in smart city. *IEEE Trans. Comput.* **65**, 1339–1350 (2016). <https://doi.org/10.1109/TC.2015.2470247>
27. [https://en.wikipedia.org/wiki/Carrier\\_IQ](https://en.wikipedia.org/wiki/Carrier_IQ). Accessed 6 June 2017

28. Xu, M., Song, C., Ji, Y., Shih, M.W., Lu, K., Zheng, C., et al.: Toward engineering a secure Android ecosystem: a survey of existing techniques. *ACM Comput. Surv.* **49**(2), 38 (2016). <https://doi.org/10.1145/2963145>
29. Marforio, C., Francillon, A., Capkun, S.: Application collusion attack on the permission-based security model and its implications for modern smartphone systems. Technical report 724, pp. 1–16 (2011). <https://doi.org/10.3929/ethz-a-006936208>
30. Harris, M.A., Brookshire, R., Chin, A.G.: Identifying factors influencing consumers' intent to install mobile applications. *Int. J. Inf. Manag.* **36**, 441–450 (2016). <https://doi.org/10.1016/j.ijinfomgt.2016.02.004>
31. Yin, P., Luo, P., Lee, W.-C., Wang, M.: App recommendation: a contest between satisfaction and temptation. In: Proceedings of the 6th ACM International Conference Web Search Data Mining, pp. 395–404 (2013). <https://doi.org/10.1145/2433396.2433446>
32. Riikonen, A., Smura, T., Kivi, A., Töyli, J.: Diffusion of mobile handset features: Analysis of turning points and stages. *Telecomm. Policy* **37**, 563–572 (2013). <https://doi.org/10.1016/j.telpol.2012.07.011>
33. <http://support.google.com/googleplay/android-developer/answer/11345?hl=en>. Accessed 7 July 2017
34. Shen, Y., Evans, N., Benameur, A.: Insights into rooted and non-rooted Android mobile devices with behavior analytics. In: Proceedings of the 31st Annual ACM Symposium on Applied Computing - SAC 2016, pp. 580–587 (2016). <https://doi.org/10.1145/2851613.2851713>
35. Liu, B., Lin, J., Sadeh, N.: Reconciling mobile app privacy and usability on smartphones: could user privacy profiles help? In: Proceedings of the 23rd international conference on World Wide Web, pp. 201–212 (2014). <https://doi.org/10.1145/2566486.2568035>

# An ECC-Based Off-line Anonymous Grouping-Proof Protocol

Zhibin Zhou<sup>1,2</sup>, Pin Liu<sup>1</sup>, Qin Liu<sup>3</sup>, and Guojun Wang<sup>4</sup>(✉)

<sup>1</sup> School of Information Science and Engineering, Central South University, Changsha 410083, China

zzbzm1031@gmail.com, jiandanglp@csu.edu.cn

<sup>2</sup> College of Physics and Information Science, Hunan Normal University, Changsha 410012, China

<sup>3</sup> School of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China

gracelq628@hnu.edu.cn

<sup>4</sup> School of Computer Science and Educational Software, Guangzhou University, Guangzhou 510006, China

csgjwang@gmail.com

**Abstract.** As more and more items tagged with RFID tags, the grouping-proof technology which is used to provide a coexistence evidence for a group of related items has become widely utilized. Due to the wireless channel used in RFID systems, security risk exists in the communication between the reader and tags. How to ensure the tag's information security and generate reliable grouping-proof becomes a hot research topic. In this paper, an ECC-based off-line anonymous grouping-proof protocol (EAGP) is proposed. Different from the traditional methods which leave the verification process completely to background server, the EAGP authorizes the reader to verify the validity of group-proof without knowing the identity of tags. From the security analysis, the EAGP can protect the security and privacy of RFID tags, defense impersonations and replay attacks. Furthermore, it has the ability to prevent the system overhead caused by invalid submission of grouping-proofs from reader. As a result, the proposed EAGP equips practical application values.

**Keywords:** Privacy · Security · Grouping-proof · Anonymous Elliptic curve cryptograph

## 1 Introduction

RFID grouping-proof technology is a mechanism that can prove a group of tagged items appeared at the same time and the same place [1]. The grouping-proof protocol can be widely used for many applications which need coexistence proof to guarantee the items have been scanned simultaneously, such as supply-chain, health care, and evidence in law, etc. [2–4]. For example, in the logistics management, we can generate a proof to guarantee the integrity of the container and

the goods in it by scanning their tags simultaneously. In the intelligent health care environment, we can prove the correction of the medicine taking through scanning the patients and their unit-dose medications at the same time and the same place [5].

According to the connection method between the reader and verifier, there are two different modes: online and offline [6]. The online mode requires a stable connection between the reader and verifier, the verifier can send and receive messages from specific tag (via the reader) during the whole protocol execution. This mode has good real-time performance and high security, but the application condition requirement are relatively high, the reliability and efficiency of communication has to be considered [7]. On the other hand, in the offline mode, the reader can collect tag information and generate grouping-proof without the participation of verifier. After these processes, the reader send the proof to verifier at last. In this vein, the verifier in offline mode don't need to communicate with any specific tag (via the reader), the mode only needs the connection between reader and verifier before and after the generation of grouping-proof. The connection requirement is more flexible during the protocol, however, there are many security problems need be solved in this mode, which has become the research focus in many literatures [4, 6, 8–15].

Figure 1 shows a common offline mode of RFID grouping-proof system. The tags are divided into  $N$  groups and each group represents a set of related items with RFID tags. The reader receives group information from the verifier and communicates with tags. If it can simultaneously scans all tags in the  $i$ th group, the reader generates a grouping-proof  $G_i$ . After every group has been scanned, the reader sends  $\{G_1, G_2, \dots, G_N\}$  to the verifier. The verifier checks these proofs and stores them as a record. In the grouping-proof protocol, the simultaneously scanning means all tags are scanned in a short time interval.

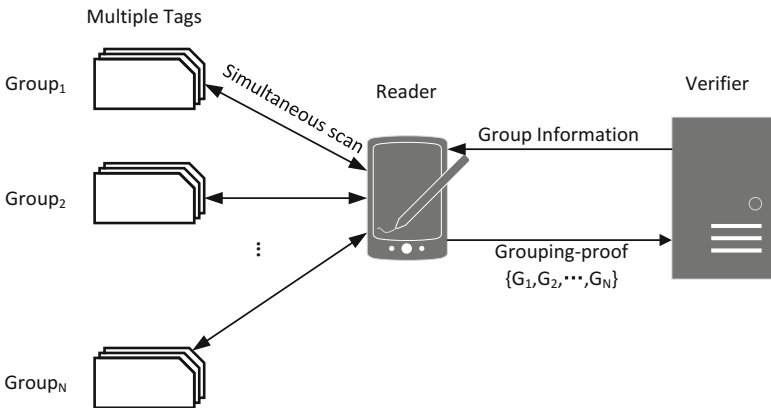


Fig. 1. The offline mode

## 1.1 Motivation

In this study, we are interested in the offline mode of grouping-proof protocols. There are many work to do in this mode. At first, The grouping-proof can show the presence of all group tags, which belong to the same single item. Note that each item intends to be sold or transported to other owners. To protect the location privacy of these items, the anonymous should be considered as an important security property. In order to do this, the authentication should be anonymous that any unauthorized third party cannot obtain tag's identity during the protocol execution. The second point is the secret key distribution. Considering the large number of tags in RFID system, the use of symmetric encryption schemes is not practical. So the PKI systems are considered. Since the encryption and decryption in RSA need performs modular exponentiation of great number, to guarantee the security, the length of modulus always larger than 1024 bits which makes the multiplication and division become the time-consuming calculation, it is not possible to be applied in RFID tags in reality. The Ellipse Curve Cryptography (ECC) method is used instead. Compared with RSA, the point or scalar multiplication is the basic operation for cryptographic protocols based on ECDLP; it is easily performed via repeated group operations. The ECC encryption can guarantee the same security level with a shorter key and is used in many RFID security protocols. The third problem of offline grouping-proof protocol is that the validity check can only be performed by the verifier. The solution to prevent invalid grouping-proof is to check the tag's identity before submitting the grouping-proof to verifier, which can only be performed by the reader. However, this solution needs the reader to store the tag's identity, which may bring potential safety hazard about the tag's privacy information. Therefore, it is essentially necessary to find a way to guarantee the legality of grouping-proof without reveal the secret information of tags.

## 1.2 Our Contributions

The main contributions of this paper are shown as follows.

- (1) We investigate the K.H's protocol [16] and provides improvements in key distribution and resistance to impersonation attack and DoP (Denial of Proof) Attack.
- (2) We establish a scheme to seal the identity of tag into the grouping-proof message by group key and session key, which makes the proof data include two types of tag information: the group member identity and the individual identity.
- (3) We propose an ECC Based offline Anonymous Grouping-proof Protocol (EAGP) which has two verification stages. The first stage is used to verify the legality of tag's group member and check the grouping-proof briefly. The second stage is used to verify the identity of tag and further confirm the grouping-proof.

- (4) We carry out the correctness proof and security analysis about the EAGP, and obtain a conclusion that this protocol can resist DoP attack [17], impersonate attack and protect the tag's information when the reader was compromised.

The rest of paper is organized as follows. An overview of related RFID grouping-proof protocols is presented in Sect. 2. Section 3 introduces the K.H's protocol. The system model and definition are described in Sect. 4. Section 5 shows the EAGP protocol. The correctness validation and security analysis about EAGP are described in Sects. 6 and 7. Section 8 provides a conclusion about this work.

## 2 Related Work

The idea of grouping-proof is first introduced in [1], and the protocol is called as yoking-proof, which only involves two tags grouping-proof in protocol. In succeeded studies, there are many promotions to enhance the security and practicality of this protocol [4, 10, 13, 15]. Burmester et al. in [18] pointed out that there are some problems in yoking-proof protocols: (1) Vulnerable to replay attack; (2) Unrelated tags can participate in a yoking session, and that the failure can only be found by the verifier; (3) The protocol does not take the presence of a rogue reader into account. To mitigate these drawbacks, the authors improves the protocol by using group key, proposing the grouping-proof protocol with forward security. In [19], the authors used the code scheme to check the tag information and improve the protocol security. In order to further improve safety of RFID systems, the work in [20] discussed the feasibility of the ECC in RFID systems. In [21], the authors proposed a RFID chip scheme to support ECC. After that, a RFID mutual authentication protocol based on ECC (ID-Transfer) was proposed. Based on the ID-Transfer, Batina proposed the first grouping-proof protocol based on the ECC technology in [22] and proved it can provide proof validation and privacy protection in the presence of untrusted tags or reader. The literature [23] showed that Batina's protocol is vulnerable to malicious tracking and proposed the improvement scheme. Hong-yan in [16] further showed that the Batina's protocol cannot resist impersonate attack and they proposed to using the authentication of reader during the grouping-proof procedure to solve this problem.

## 3 Preliminaries

In this section, we introduce the Ellipse Curve Cryptography (ECC) and the related hardness problems. The details are described as follows.

### 3.1 The Ellipse Curve Cryptography

Elliptic curves are algebraic structures that constitute a basic class of cryptographic primitives which rely on a mathematical hard problem. An elliptic



curve  $E$  over a finite field  $\mathbb{F}_q$  with characteristic  $q > 3$  can be defined by the Eq. (1):

$$y^2 = x^3 + ax + b \quad (1)$$

where  $a, b, x, y \in \mathbb{F}_q$  and  $4a^3 + 27b^2 \neq 0 \pmod{p}$ . The point  $(x, y)$  is a point on the elliptic curve. Let  $P$  be a fixed point on the curve  $E(\mathbb{F}_q)$  with prime order  $n$  and  $k$  is a large integer scalar in  $[1, n - 1]$ . Due to the hardness of Elliptic Curve Discrete Logarithm Problem [26], it is easy to compute the scalar multiplication  $Q = kP$  but hard to find  $k$  by knowing only  $Q$  and  $P$ .

### 3.2 Elliptic Curve Discrete Logarithm Problem (ECDLP)

**ECDLP Definition:** Given an elliptic curve  $E$  defined over a finite field  $\mathbb{F}_q$ , a point  $P \in E(\mathbb{F}_q)$  of order  $n$ , and a point  $Q = kP$  where  $0 \leq k \leq n - 1$ , determine  $k$ .

The well-known hardness of the ECDLP is crucial for the security of our elliptic curve scheme.

## 4 Investigation of Hong-yan K's Protocol

Literature [16] puts forward the grouping-proof protocol based on ECC. The framework of this protocol is shown in Fig. 2.

The protocol has four stages: (1) Initialization stage, (2) Authentication stage, (3) Grouping-proof generation stage, (4) Verification stage. In initialization stage, the server writes the tag's private key, the reader's public key and verifier's public key into a tag, expressed as  $\{s_t, K, Y\}$ . The authentication stage is used to authenticate the identity of a reader. It can prevent the reader impersonation attack. In the grouping-proof generation stage, according to the random number broadcasted by the reader, tag A and tag B calculate  $\{T_{a,1}, T_{a,2}\}$  and  $\{T_{b,1}, T_{b,2}\}$ . Finally, the reader passes these data as grouping-proof to the verifier for validation.

K.H's protocol uses authentication to solve the impersonation attack, and there are some flaws need to be pointed out.

- (1) The key distribution: In K.H's protocol, tag A and tag B need to store the reader's public key. If the reader is changed, the new public key needs to be written into all the tags. If the amount of tags is very big, the overhead is too serious.
- (2) The DoP attack: the reader in K.H's protocol can't validate the proof and is unable to check the legality of tags. If the reader suffered from DoP attack or some unrelated tags taken part into the proof process, before the proof be sent to the verifier, the failure can't be identified immediately which will reduce the system real-time performance.
- (3) Communication overhead: the using of authentication stage increases the number of communication times between the tag and the reader, which leads to the additional overhead of communication.

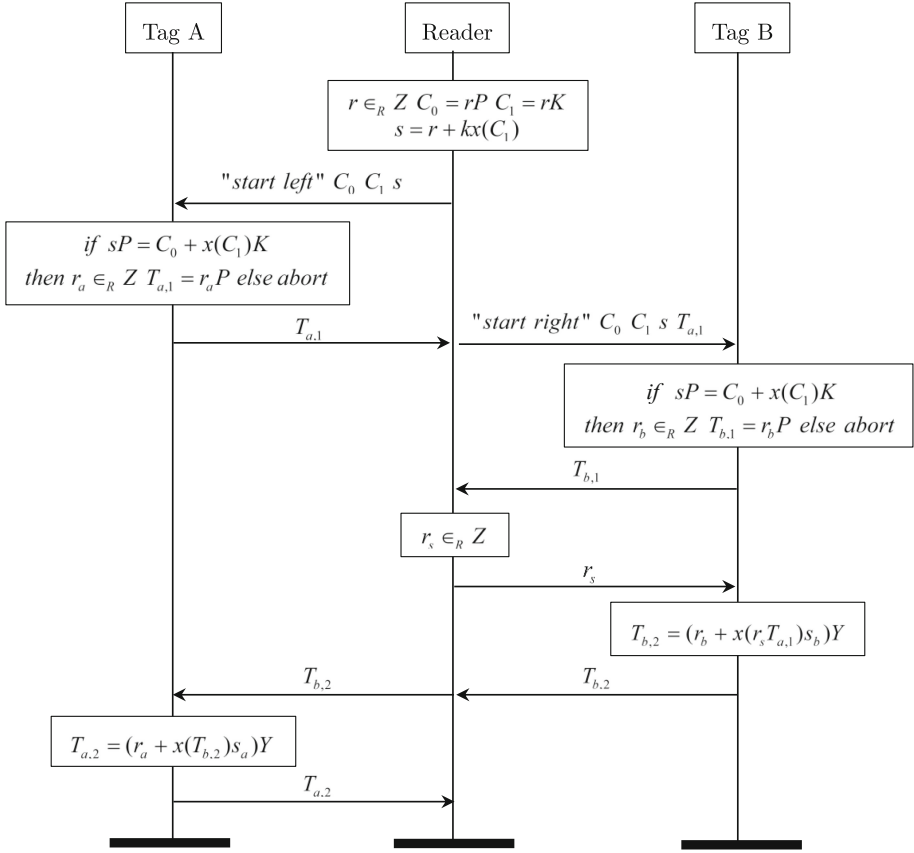


Fig. 2. The Hong-yan K's protocol

## 5 The System Model and Security Requirement

### 5.1 The System Mode

In our work, The RFID grouping-proof system is consist of three parts: reader, RFID tags and Verifier.

- Tag: The tags in our protocol is passive low-cost device which has a relative small storage and limited computational capacity. The tags are divided into several groups.
- Reader: The RFID reader is a powerful device which is controlled by an untrusted third party. For security reasons, the privacy information of tag and verifier is unknown to reader.
- Verifier: An offline trusted third party (TTP) which maintains all the keys and identities of groups.

There are two type of channels in our protocol. The channel between the tag and the reader and the channel between the reader and the verifier. We assume the former is not secure and can be attacked by the adversary. The second channel is secure and the message transferred in this channel can not be eavesdropped.

## 5.2 The Adversary Mode

In grouping-proof protocols, the adversary has two purposes: (1) Forge the grouping-proof which can pass the validation of verifier; (2) Get the privacy information of the reader and tags. According to the attacker described in [24], the adversary in our protocol can completely control the communication channel between the reader and tags, in terms of modifying, delaying and replaying any message in protocol. In addition, the adversary can also hack the tag and fully control it.

## 5.3 The Security Requirement of Grouping-Proof System

The security requirements include these parts:

- **Anonymity**  
The anonymous of tags and readers, which means the adversary can't get the identity of a tag or a reader by eavesdropping the protocol message.
- **Location Privacy**  
The adversary can't track the location of a reader and tags through the protocol messages.
- **Resist to replay attack**  
The adversary can't use the message in previous sessions to cheat the reader or tags to generate grouping-proof.
- **Defense the DoP Attack**  
The adversary can't use illegal tag involved in the protocol to disturb the proof validation execute by verifier [17].
- **Reader secret information protection**  
If the reader is hacked in, the adversary can't use the information stored in it to extract any secret information of tags.

## 6 Describe of EAGP

To overcome the weakness of the grouping-proof protocol which is put forward in [16]. We come up with the improvement protocol EAGP.

The simultaneous scan is the basic requirement in grouping-proof protocols. To ensure this, the EAGP uses the timeout mechanism to guarantee the tags are scanned by a reader in a very short interval. When the protocol starts, both the reader and tag activate a timer. If a session of grouping-proof don't complete before the timeout, then the protocol is terminated. For simplicity, we assume

**Table 1.** Summary of notations

Notation	Description
$r_s, r_a, r_b$	The random number generated by reader, Tag A and Tag B
$P$	The base point on the elliptic curve $E(\mathbb{F}_q)$
$Y, y$	The public/private key of group $G$
$k_a, k_b$	Temporary grouping-proof key of Tag A, B
$k_{ai}, k_{bi}$	Secret key of Tags A and B
$PK_A, PK_B$	Public key of Tags A and B

each group has two tags. It’s not difficult to expand our protocol to multiple tags. We assume the verifier can be trusted. The reader and tag are untrusted and can be impersonated or even controlled by an adversary. The notations used in EAGP are summarized in Table 1.

In EAGP, without losing any security characteristics, we cut down the times of communication between the reader and tags to reduce the communication overhead. The proposed protocol is consisted of three phases: Initial phase, Grouping-proof generation phase and Verification phase.

The descriptions of the protocol are as follows:

### 6.1 Initial Phase

The verifier divides the Tag A and Tag B into one group, allocates group parameters as: the verifier chooses a random number  $y \in \mathbb{Z}$  and computes  $Y = y \cdot P$  as its public key. The group’s public key  $Y$  is stored in the tag, while keeping the private key  $y$ . Both tags share their secret keys  $k_{ai}$  or  $k_{bi}$  with verifier; In addition, The verifier stores the public key  $PK_A$  and  $PK_B$ . Reader gets the group key  $y$  from the verifier.

### 6.2 Grouping-Proof Generation Phase

The framework is demonstrated in Fig. 3.

- (1) Reader generates a random number  $r_s$ , calculate  $C_0 = r_s P$ ,  $C_1 = r_s Y$ , and  $s = r_s + yx(C_1)$ . Then, the  $\{s, C_0, C_1, r_s\}$  is sent to the tag A along with the message of “start left”.
- (2) Tag A verifies the equation  $sP = C_0 + x(C_1)Y$ . If it does not hold, the protocol is terminated. Otherwise, it generates a random number  $k_1$ , calculates  $r_a = x(k_1 P)$ , generates the session secret key  $k_a = x(Y) \oplus r_a$ . Then, it seals its secret key  $k_{ai}$  into message  $m_a$  as follows:

$$m_a = k_1^{-1}(r_s + k_{ai} \cdot r_a) \tag{2}$$

Finally, tag A sends  $\{m_a, r_a\}$  to the reader.

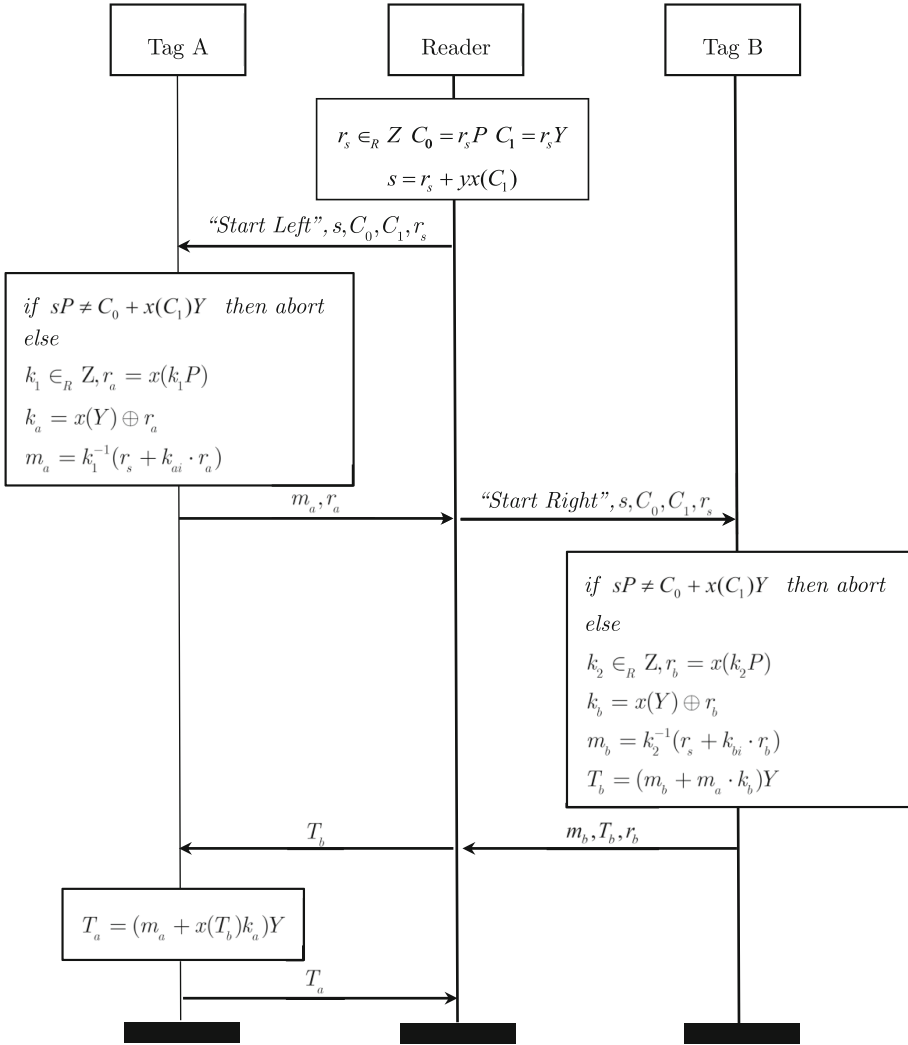


Fig. 3. The EAGP

- (3) Reader sends  $\{m_a, s, C_0, C_1, r_s\}$  along with the message of "start right" to Tag B.
- (4) Tag B verifies the equation  $sP = C_0 + x(C_1)Y$ . If it does not hold, the protocol is terminated. Otherwise, it generates a random number  $k_2$ , calculates  $r_b, k_b, m_b, T_b$  and sends  $\{m_b, T_b, r_b\}$  to the reader.
- (5) Reader sends the message  $T_b$  to Tag A.
- (6) Tag A calculate  $T_a = (m_a + x(T_b)k_a)Y$ , and send it to reader.
- (7) The reader generates the grouping-proof  $G$  shown in Eq. (3)

$$G = \{m_a, T_a, m_b, T_b, r_a, r_b, s\} \tag{3}$$

### 6.3 The Verification Phase

There are two steps in verification phase, (1) Reader verification step. (2) Verifier verification step.

(1) Reader verification step:

Reader calculates  $Y' = yP$ ,  $k'_a = x(Y') \oplus r_a$ ,  $k'_b = x(Y') \oplus r_b$  and validates the Eqs. (4) and (5):

$$(y^{-1}T_a - m_aP) \cdot x(T_b)^{-1} = k'_aP \quad (4)$$

$$(y^{-1}T_b - m_bP) \cdot m_a^{-1} = k'_bP \quad (5)$$

The utilization of group key  $y$  can prove that Tag A and B belong to the same group and be scanned by reader simultaneously.

(2) Verifier verification step:

The second verification stage is executed by the verifier to authenticate the tag's identity in grouping-proof. The procedure is described as follows:

– Calculate the following equations

$$w = m_a^{-1} \bmod n \quad (6)$$

$$u_1 = s \cdot w \bmod n \quad (7)$$

$$u_2 = r_a \cdot w \bmod n \quad (8)$$

$$x_a = x(u_1P + u_2PK_A) \quad (9)$$

– If  $x_a = r_a$  is valid, the validation is successful, and the verifier stores the proof in the server as a record. Otherwise, the validation fails and the proof is abandoned.

## 7 Correctness Proof

### 7.1 The Correctness of Reader Verify

If  $Y = y \cdot P$  is known, we have

$$y^{-1}T_a = m_aP + x(T_b)k_aP \quad (10)$$

Then, the left side of Eq. (4) can be simplified as  $x(T_b)k_aP \cdot x(T_b)^{-1} = k_aP$ . Therefore, the Eq. (4) is proved.

In a similar way, we have

$$y^{-1}T_b = (m_b + m_a \cdot k_b)P \quad (11)$$

In this vein, the left side of Eq. (5) can be simplified as  $m_a \cdot k_bP \cdot m_a^{-1} = k_bP$ . Then, the Eq. (5) is proved.

## 7.2 The Correctness Verifier Authentication

According to Eq. (2), we have:

$$k_1 = m_a^{-1}(s + k_{ai} \cdot r_a) \quad (12)$$

According to Eqs. (7) and (8), we have: According to Eqs. (7) and (8), we have:

$$\begin{aligned} k_1 &= m_a^{-1} \cdot s + m_a^{-1} \cdot k_{ai} \cdot r_a \\ &= u_1 + u_2 \cdot k_{ai} \end{aligned} \quad (13)$$

Then, we can obtain:

$$\begin{aligned} x_a &= x(u_1P + u_2PK_A) \\ &= x(u_1P + u_2k_{ai} \cdot P) \\ &= x(k_1P) = r_a \end{aligned} \quad (14)$$

The correctness proof of verifier authentication is completed.

## 8 Security Analysis and Comparison

### 8.1 Security Analysis

#### The anonymous of tag and reader

During the execution of the protocol, the communication message set can be expressed as  $\{m_a, r_a, r_b, s, r_s, m_b, T_a, T_b\}$ . Among them,  $r_a, r_b, r_s$  are the random numbers generate by tags and reader, while the other messages are calculated from these random numbers, The adversary can't get any information of protocol participants from the communication messages.

#### The location privacy of tag and reader

All the messages sent from the EAGP are random numbers or generated from random numbers. In each protocol session, the session key and random numbers are different. Adversary can't figure out the protocol participants by the messages they send. Therefore, it is difficult for the adversary to track any tag or reader, the locations of reader and tags are protected.

#### Defense of DoP attack

The EAGP adds the reader verification in protocol. When the reader sends the proof to a verifier, the reader can verify the tag's group member identity and proof data before hand. If the adversary does not know the group key, the message  $\{m_a, T_a, m_b, T_b\}$  will not satisfy the Eqs. (4) and (5), it is impossible to cheat the reader to sending invalid grouping-proof to the verifier.

#### Reader secret information protection

In EAGP, the reader only stores the group's private key  $y$ . No tag information is stored in reader's memory. Even if the adversary gets the group's private key by hacking the reader, it still can't get any secret information about tag, which makes sure the information security of tags.

### Resist to impersonation attack

The impersonation attack includes two methods: Impersonate Tag, and Impersonate Reader. In the first type, the adversary impersonates the tag, tries to cheat the reader to pass the grouping-proof verification, and further cheats the verifier. In the second type, the adversary impersonates the reader to collect the tag's information, or generates the valid grouping-proof without scan to the real tag. The attack process is described as follows.

#### – Impersonate Tag

There are two situations where the adversary impersonate a tag: (1) The adversary don't know any secret key, that means it can't deduce legal  $\{T_a, T_b\}$ . In this situation, the grouping-proof generated in present of attack can't pass the reader validation Eqs. (4) and (5). This attack can be detected before the proof is sent to verifier, protecting the system from DoP attack. (2) The adversary get the group's public key  $Y$ . From  $Y$ , he can deduce the session key  $k_a$  and  $k_b$ . Then he can generate the grouping-proof that can satisfy Eqs. (4) and (5). But due to the lack of tag A, B's authentication secret key  $k_{ai}, k_{bi}$ , to forge the legal  $\{m_a, m_b\}$  need solve the ECDLP described in Sect. 2, the probability is negligible. So it is nearly impossible to pass the verifier validation. In conclusion, EAGP can resist the tag impersonate attack in both situations.

#### – Impersonate Reader

If the adversary impersonate the reader, the message set it can collect is  $\{m_a, r_a, T_a, m_b, r_b, T_b\}$ , all the information are transferred in ciphertext. Without knowing the secret key of tag, the reader can't deduce the tag's identity, neither can he forge valid grouping-proof without scanning legal tags.

### Resist to Replay Attack

The replay attack is the adversary uses a tag's response to a rogue reader's challenge to impersonate the tag. Suppose the adversary collected the message of tag A:  $\{m_a^1, r_a^1, r_s^1, s^1, T_a^1\}$  in EAGP session  $p1$ , collected the message of tag B:  $\{m_b^2, r_b^2, T_b^2, r_s^2, s^2\}$  in EAGP session  $p2$ . He send the message  $T_b$  to tag A, get the response  $\hat{T}_a^2 = (m_a^1 + x(T_b^2)k_a^2)Y$ , try to forge the grouping-proof about tag A and tag B:  $\hat{G} = \{m_a^1, \hat{T}_a^2, m_b^2, T_b^2, r_a^1, r_b^2, s^2\}$ . Because the message is get in two different sessions, there are  $m_a^1 \neq m_a^2, k_a^1 \neq k_a^2$ , so the  $\hat{T}_a^2 \neq T_a^2$ , neither Eqs. (4) or (5) are satisfied. The forged grouping-proof can't pass the reader validation.

## 8.2 Security Comparison

Table 2 lists the comparison of the existing grouping-proof schemes and EAGP. It can be seen from the comparison that the EAGP basically satisfied the security requirements of the grouping-proof protocol.



**Table 2.** The comparison of grouping-proof protocols

	Anonymous	Location privacy	DoP attack	Reader information protection	Tag impersonate attack	Reader impersonate attack	Replay attack
Juels [1]	×	×	×	√	×	×	×
Burmester [18]	√	√	×	√	√	×	√
Burmester [19]	√	√	×	×	√	√	√
Batina [22]	√	√	×	√	×	×	×
Chao [23]	√	√	×	√	×	×	√
Lin [25]	√	√	×	√	×	×	×
Hong-yan [16]	√	√	×	√	×	√	√
EAGP	√	√	√	√	√	√	√

## 9 Conclusion

Batina et al.’s protocol is proved vulnerable to impersonate attack, the Hong-yan K’s protocol add the authentication phase to solve this problem, but the solution can’t resist DoP attack and has problem in secret key distributions and impersonation attack. In this paper, we proposed an offline grouping-proof protocol EAGP. Through the security analysis, this protocol can resist impersonation, DoP and replay attack, and protect the security and privacy of tag’s secret information.

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation of China under Grant Numbers 61632009, 61472451 and 61402161, the High Level Talents Program of Higher Education in Guangdong Province under Grant Number 2016ZJ01, the Hunan Provincial Education Department of China under Grant Number 2015C0589, the Hunan Provincial Natural Science Foundation of China under Grant Number 2015JJ3046, the Fundamental Research Funds for the Central Universities of Central South University under Grant Number 2016zzts058.

## References

1. Juels, A.: “Yoking-Proofs” for RFID Tags. In: PerCom Workshops, pp. 138–143 (2004)
2. Peris-Lopez, P., Orfila, A., Mitrokotsa, A., Van der Lubbe, J.C.: A comprehensive RFID solution to enhance inpatient medication safety. *Int. J. Med. Inform.* **80**(1), 13–24 (2011)
3. Chen, Y.-Y., Tsai, M.-L.: An RFID solution for enhancing inpatient medication safety with real-time verifiable grouping-proof. *Int. J. Med. Inform.* **83**, 70–81 (2014)
4. Chen, C.-L., Wu, C.-Y.: Using RFID yoking proof protocol to enhance inpatient medication safety. *J. Med. Syst.* **36**, 2849–2864 (2012). *Informatics*, vol. 80, pp. 13–24 (2011)
5. Zhou, Z., Liu, Q., Wang, G., Jia, W.: Secure medication scheme using the grouping-proof technology. *J. Chin. Comput. Syst.* **36**, 2349–2353 (2015)

6. Peris-Lopez, P., Orfila, A., Hernandez-Castro, J.C., van der Lubbe, J.C.A.: Flaws on RFID grouping-proofs. Guidelines for future sound protocols. *J. Netw. Comput. Appl.* **34**, 833–845 (2011)
7. Xie, K., Cao, J., Wang, X., et al.: Optimal resource allocation for reliable and energy efficient cooperative communications. *IEEE Trans. Wirel. Commun.* **12**, 4994–5007 (2013)
8. Sundaresan, S., Doss, R., Zhou, W.: Offline grouping proof protocol for RFID systems. In: 2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 247–252 (2013)
9. Liu, H., Ning, H., Zhang, Y., He, D., Xiong, Q., Yang, L.: Grouping-proofs-based authentication protocol for distributed RFID systems. *IEEE Trans. Parallel Distrib. Syst.* **24**(7), 1321–1330 (2013)
10. Chien, H.-Y., Liu, S.-B.: Tree-based RFID yoking proof. In: 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing, NSWCTC 2009, pp. 550–553 (2009)
11. Lien, Y., Leng, X., Mayes, K., Chiu, J.-H.: Reading order independent grouping proof for RFID tags. In: IEEE International Conference on Intelligence and Security Informatics, ISI 2008, pp. 128–136 (2008)
12. Piramuthu, S.: On existence proofs for multiple RFID tags. In: ACS/IEEE International Conference on Pervasive Services, pp. 317–320 (2006)
13. Li, N., Mu, Y., Susilo, W., Varadharajan, V.: Anonymous yoking-group proofs. In: Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security, pp. 615–620 (2015)
14. Ma, C., Lin, J., Wang, Y., Shang, M.: Offline RFID grouping proofs with trusted timestamps. In: 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications, pp. 674–681 (2012)
15. Cho, J.S., Yeo, S.S., Hwang, S., Rhee, S.Y., Kim, S.K.: Enhanced yoking proof protocols for RFID tags and tag groups. In: 22nd International Conference on Advanced Information Networking and Applications - Workshops, AINAW 2008, pp. 1591–1596 (2008)
16. Hong-yan, K.: Analysis and improvement of ECC-based grouping-proof protocol for RFID. *Int. J. Control Autom.* **9**, 343–352 (2016)
17. Lo, N.-W., Yeh, K.-H.: Anonymous coexistence proofs for RFID tags. *J. Inf. Sci. Eng.* **26**, 1213–1230 (2010)
18. Burmester, M., De Medeiros, B., Motta, R.: Provably secure grouping-proofs for RFID tags. In: International Conference on Smart Card Research and Advanced Applications, pp. 176–190 (2008)
19. Burmester, M., Munilla, J.: An anonymous RFID grouping-proof with missing tag identification. In: 10th IEEE International Conference on Radio-Frequency Identification, pp. 3–5 (2016)
20. Wolkerstorfer, J.: Is elliptic-curve cryptography suitable to secure RFID tags. In: Handout of the Ecrypt Workshop on RFID and Lightweight Crypto (2005)
21. Batina, L., Guajardo, J., Kerins, T., Mentens, N., Tuyls, P., Verbauwhede, I.: An elliptic curve processor suitable For RFID-tags. IACR Cryptology ePrint Archive, vol. 2006 p. 227 (2006)
22. Batina, L., Lee, Y.K., Seys, S., Singele, D., Verbauwhede, I.: Extending ECC-based RFID authentication protocols to privacy-preserving multi-party grouping proofs. *Pers. Ubiquit. Comput.* **16**, 323–335 (2012)
23. Li, H., Lv, C., Ma, J., et al.: Security analysis of a privacy-preserving ECC-based grouping-proof protocol. *J. Convergence Inf. Technol.* **6**, 113–119 (2011)

24. Canetti, R.: Universally composable security: a new paradigm for cryptographic protocols. IEEE Computer Society (2001)
25. Lin, Q., Zhang, F.: ECC-based grouping-proof RFID for inpatient medication safety. *J. Med. Syst.* **36**, 3527–3531 (2012)
26. Menezes, A.: Evaluation of security level of cryptography: the Elliptic Curve Discrete Logarithm Problem (ECDLP). University of Waterloo (2001)

# PCSD: A Tool for Android Malware Detection

Bo Leng<sup>1</sup>, Jianbin Li<sup>2</sup>(✉), Yang Xu<sup>1</sup>, Liang She<sup>1</sup>, Wuqiang Gao<sup>1</sup>,  
and Quanrun Zeng<sup>1</sup>

<sup>1</sup> School of Information Science and Engineering, Central South University,  
Changsha 410083, China

{boleng, xuyangcsu, sheliang, forrestgao, zengquanrun}@CSU.EDU.CN

<sup>2</sup> Information Security and Big Data Research Institute, Central South University,  
Changsha 410083, China  
lijianbin@CSU.EDU.CN

**Abstract.** The increasing amount and diversity of malicious applications are reducing efficiency of conventional defenses and it is necessary to create novel method for detection. Consequently, we propose PCSD, a lightweight tool for detection of Android malware by extracting statistical features from applications. As the influence of individual difference, PCSD performs cluster algorithm to reduce particularity. Meanwhile, it minimizes effect of random cluster by selecting cluster, which has minimum volatility on size per cluster, for improving detection accuracy. In our work, we collect statistical features from 5,553 malicious applications and 3,000 benign applications and build train model for detecting on the basis of different machine learning algorithms, like Bayesian ridge, Random forests, etc. Our results show that accuracy is 99.02% and *AUC* (*Area Under Curve*) is 99.51% in experiment. These results demonstrate the efficacy of PCSD to distinguish malicious and benign android applications.

**Keywords:** Android · Malware detection · Statistical features  
Machine learning · Individual difference

## 1 Introduction

A recent data from IDC [1] shows that Android system has occupied 85.0% of market in the first quarter of 2017. Application market such as Google Play Store is playing an crucial role in the popularity of Android devices and drive the economy of Android applications. However, efficient Internet connectivity and availability of individual information such as contacts, browsing history, messages and social network access have attracted the enormous interest of malware developers. Obviously, how to detect and keep the large number of malwares out of the android system is an emerging, crucial, but challenging issue. Previous work on the detection of malware primarily focused on static analysis combined features and machine learning algorithm.

Permission mechanism is one of the major Android security strategies. Permissions always show what applications want to do and can be extracted expediently. For instance, when some apps want to write external storage, they have

to apply for permission named `WRITE_EXTERNAL_STORAGE`, and all permissions developers requested need to be declared in *AndroidManifest.xml* and included in special label named `uses-permission`. The information of permissions can be utilized to detect android malwares [2]. Simultaneously, combining permissions and diverse machine learning algorithms [3, 4] is also a good way to differentiate malignant and innocuous applications. These approaches are similar on employing permissions from *AndroidManifest.xml* for identifying android malware and pay attention to content permissions reveal. Android components supply user interface and background service. Android app is composed of four types of component: *Activity*, *Service*, *Broadcast Receiver*, and *Content Provider*. Components are necessary for android applications and essential property of android programs. Components are able to reveal different characters from malwares and be used for detection [5, 6]. Plenty of smali files which have same format can be extracted from android application packages and clearly reflect procedure of program. Hoffmann et al. and Tang et al. made use of data-flow and intent-flow analysis in smali code respectively to discriminate malicious and benign applications [7, 8]. Techniques, which leverage smali info to detect harmful behavior, aim at gleaning method call graphs and then differentiate android applications incorporate with call graphs. However, those approaches absolutely focus on finding out what each feature means and what information behind each permission has. In addition, these researchers ignore a simple fact that various applications have disparate peculiarity and there is no unified strategy to mining characters of all applications. On the basis of this, PCSD, which combines permissions, components and smali information to detect, analyzes and extracts features in another point of view and makes use of individual differences adequately to differentiate malicious and benign applications.

Our approach is based on static analysis of source code of Android applications for obtaining statistic characters. Due to individual differences and the large amount of samples, clusters algorithm is used to minimize diversity between samples before training model. Moreover, different machine learning methods, like *Support Vector Regression*, *Random Forest Regression*, etc., are utilized for detection. The major contributions of this work include three points as follows: 1: Extracting statistic features of apk code rather than what property of data set means or represents for detection. 2: Reducing influence of individual differences in entire data set by cluster algorithm and minimizing affect of random cluster by selecting cluster number on the basis of volatility of size per cluster. 3: Utilizing two-layer learning mechanism for better accuracy.

Rest of the paper is summarized as follows. Section 2 describes the related work. Section 3 contains design and implementation. Section 4 gives observation and analysis. Section 5 makes the conclusion.

## 2 Related Work

Android malware detection contain dynamic analysis, static analysis and method combined dynamic and static analysis. Dynamic detection depends on monitoring behavior of android application when application is running. For example,

Tam et al. reconstructed the behaviors of Android malwares for detection by dynamic analysis [9]. Enck et al. utilized taint tracking in dynamic analysis to trail multiple sources of sensitive data for differentiate malevolent and innocuous examples [10]. Static analysis of android malwares examines a program without executing any code and potentially exhibits all possible paths of execution. Generally, it extracts plenty of features from android applications and always utilizes machine learning algorithm to train model. For instance, a great deal of analogous approaches in static malware detection have used diverse features, such as permissions [11, 12], components [13–15], API calls [16], system calls, with different classifiers such as Bayesian classification [17], support vector machine (SVM) [18, 19], and k-Nearest Neighbor [20]. Crowdroid [21] and DroidAPIMiner [22] also leveraged machine learning techniques to analyze features for detecting malware. Combining dynamic and static analysis is a method, which can leverage advantages of each other sufficiently. Yuan et al. proposed a ML-based method that utilizes more than 200 features extracted from both static analysis and dynamic analysis of Android apps for malware detection [23]. Spreitzenbarth et al. used results of static analysis to guide dynamic analysis and extended coverage of executed code [24]. Lindorfer et al. composed static analysis with dynamic analysis on both Dalvik VM and system level, as well as several stimulation techniques to increase code coverage [25]. Our work is related to static approaches and makes use of permissions, components, smali info. However, it differs in two primary aspects from previous work: First, we focus on statistic significance of feature, like how many permission one application has and how many static fields, final fields one apk has. Second, we significantly reduce influence of individual difference in training.

### 3 Design and Implementation

#### 3.1 Architecture

Our PCSD system is composed by three layers, namely, collection layer, extraction layer and Learning layer. Figure 1 shows our entire structure.



**Fig. 1.** Architecture of PCSD.

### 3.2 Collection

**Malicious Aps:** In our work, we used a Drebin Dataset [18], which contains real world Android malwares. The dataset consists of 5,560 applications from 179 different malware families, which size over 20, and consists of spywares, Adwares, information stealers, Trojans, etc.

In the light of our approach, we need to assemble file named *AndroidManifest.xml* and suffixed by *.smali* from each of these applications (APK files). We ignore 7 cases which can not be disassembled. Then we finally aggregate 5,553 malignant specimens.

**Benign Aps:** Although there are plenty of application stores, like *1moblie*, *Amazon store* and other things, *Google Play* is the most formal application sphere in which you are anticipated to install or download applications. As a result of the limitation of downloading times from *Google Play* and the slowness of downloading by oneself, we make use of four terminals and identifications to request APK. For embodying typical of samples sufficiently, we obtain applications from 27 categories in store and specific size per category shown in Table 1.

**Table 1.** Benign APKs.

Category	Size	Category	Size	Category	Size	Category	Size
Food and beverage	80	Software and demonstration	100	Personalise	84	Finance	130
Shopping	100	Vehicles and traffic	90	The company	94	Education	121
Map and navigation	98	Activity	52	Physical education	135	The weather	29
Travel navigation	36	Beauty fashion	75	Tool	79	Photography	104
Health and fitness	80	Home renovation	91	Social dating	82	Parenting	96
Music and audio	111	Art and design	214	Social	190	Entertainment	119
Life fashion	137	Business office	209	Medical care	264		

### 3.3 Extraction

Actually, APK file is compressed file which can be opened by unzip software. The Apktool [26] is a tool for reverse engineering Android apk files and can decode resources to nearly original form. In this paper, Apktool is applied to decompile android applications. After disassembling APK, there is a lot of files and folders. Figure 2 is a snippet of this files and folders. *AndroidManifest.xml*: it represents all permissions and components application requests. *apktool.xml*: it shows some additional info, like sdk Info, version and version info and so forth. *res folder*: it contains assorted resource. For example, all kinds of pictures and layout files. *smali*: Every smali file, which comes from Java bytecode, belongs to this folder. *original*: it includes some binary information. From what can be uncovered in *AndroidManifests.xml*, we will acquire a couple of characters, like how many activities on apk has, what kind of permission application has, etc. In smali folder, there is plenty of smali files. Whatever the name of smali file is, it's format always be same. Every smali file include field, method, interface



Fig. 2. A snippet of files and folders from decompiled APK.

that implemented by itself, class that inherited by itself. In this paper, we census total number of every feature information from smali file, like final field, static field and max register number, which belongs to identical application.

### 3.4 Learning

Learning shown in Fig. 3 is a significant part of this paper and consists of two sections: 1: In machine learning algorithm, sample difference usually results in enormous variety of train model. According to this, First Learning utilize cluster algorithm to divide malicious samples into different clusters for reducing effect of individual difference in training. Due to influence of diverse number of cluster in random cluster algorithm, size per cluster are going to fluctuate and vary. Then it will mostly effect accuracy and error of experiment. Consequently, we select final cluster number, which has minimum volatility of size per cluster. Subsequently, we build train set and test set on the basis of cluster result and utilize multiple machine learning algorithms to train on train set. Finally, we assess every model by test set and obtain optimum model per cluster, which get best accuracy. 2: Second Learning build temporary train set and test set. Then we leverage optimum model per cluster to calculate temporary set and constitute ultimate set with result per model. Finally, we take advantages of different algorithm to train model and select best model as final model for detection.

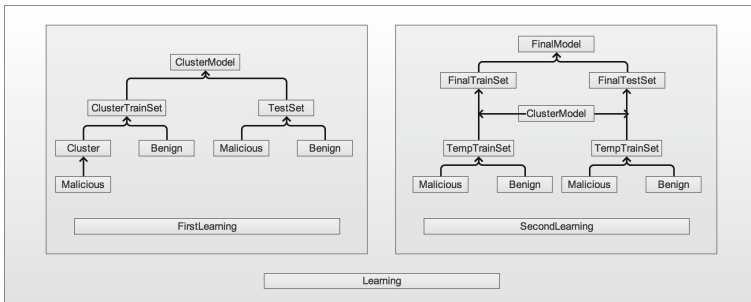


Fig. 3. Framework of Learning layer.



## 4 Observation and Analysis

### 4.1 Cluster

We leverage *KMeans* algorithm to generate clusters in 2,000 random malignant samples of entire malicious set. On account of significance of cluster number which mainly influences the complexity and accuracy of work, we execute 10 times trial when cluster number varying from 3 to 8 and the result show in Fig. 4.

In every subfigure of Fig. 4, abscissa shows how many clusters malevolent samples are divided into and ordinate represents the magnitude per cluster. As we all know, disparate clusters bring about undulant consequence of experiment. Hence, we finally decide that malevolent application will be split into five clusters on the basis of five clusters, which is provided with minimum volatility.

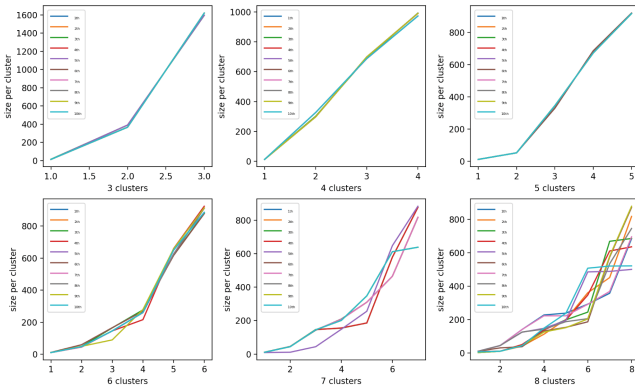


Fig. 4. Volatility of cluster.

### 4.2 Multiple Algorithm

On the basis of empirical analysis, which disparate algorithms used on train set will generate various effect. We select a few of machine learning algorithms, which include *Support Vector Regression (SVR)*, *MLPRegressor*, *Random Forests Regression*, *KNeighborsRegressor*, *RidgeRegression*, *LinearRegression*, *Bayesian Ridge Regression*, for obtaining best effect.

### 4.3 Model per Cluster

Under training of multiple algorithms, we calculate the *True Positive Ratio (TPR)* of each algorithm per cluster:

$$\begin{cases} TPR = \frac{TP}{TP+FN}. & (1) \\ y_i = FRP(Cluster_i, Algorithm_j) \quad (i = 1, 2, 3, 4, 5; j = 1, 2, 3, 4, 5, 6, 7). & (2) \end{cases}$$

where  $TP$  is the number of malware cases correctly checked (true positives) and  $FN$  is the number of malware cases misclassified as benign applications (false negatives). We also measured the *False Positive Ratio (FPR)* of each algorithm per cluster:

$$\begin{cases} FPR = \frac{FP}{FP+TN}. & (3) \\ y_i = FRP(Cluster_i Algorithm_j) \quad (i = 1, 2, 3, 4, 5; j = 1, 2, 3, 4, 5, 6, 7). & (4) \end{cases}$$

where  $FP$  is the number of benign software cases incorrectly checked as malware and  $TN$  is the number of legitimate executables correctly classified. Furthermore, we calculate the *Accuracy* of each algorithm per cluster, i.e., the total number of the classifier’s hits divided by the number of instances in the whole dataset:

$$\begin{cases} Accuracy = \frac{TP+TN}{TP+FP+TN+FN}. & (5) \\ y_i = Accuracy(Cluster_i Algorithm_j) \quad (i = 1, 2, 3, 4, 5; j = 1, 2, 3, 4, 5, 6, 7). & (6) \end{cases}$$

Besides, we measured the *Area Under the ROC Curve (AUC)* of each algorithm per cluster which establishes the relation between false negatives and false positives. The ROC curve is obtained by plotting the  $TPR$  against the  $FPR$ .

$$y_i = AUC(Cluster_i Algorithm_j) \quad (i = 1, 2, 3, 4, 5; j = 1, 2, 3, 4, 5, 6, 7). \quad (7)$$

Then we draw the *ROC* per cluster as Fig. 5, which takes the *False Positive Rate (FPR)* as the horizontal axis and the *True Positive Rate (TPR)* as the vertical axis.

From Fig. 5 we can see that disparate methods result in diverse curves and then we elected optimum algorithm, which get the maximum *AUC (Area Under Curve)*. At the same time, we record *FPR, TPR, AUC, Accuracy* as Table 2 belongs to optimum algorithm.

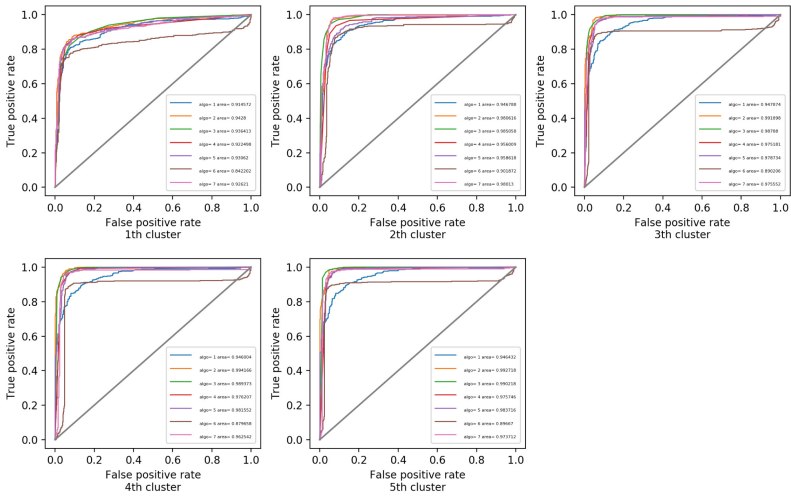


Fig. 5. ROC per cluster.

**Table 2.** Result per cluster.

Cluster	Optimum algorithm	TPR	FPR	AUC	Accuracy
1th cluster	MLPRegressor	0.82	0.05	0.9428	0.863333333333
2th cluster	MLPRegressor	0.951	0.054	0.980616	0.949333333333
3th cluster	MLPRegressor	0.976	0.052	0.991898	0.966666666667
4th cluster	MLPRegressor	0.982	0.056	0.994166	0.969333333333
5th cluster	MLPRegressor	0.986	0.05	0.992718	0.974

#### 4.4 Final Model

We utilize model per cluster to transform temporary train set and test set into final set. Each model calculates temporary set to obtain one column data as:

$$y_i = Model_i(temporary\_set) \quad (i = 1, 2, 3, 4, 5). \tag{8}$$

Then we combine five column data with label column as:

$$Z = \{y_i, label\} \quad (i = 1, 2, 3, 4, 5). \tag{9}$$

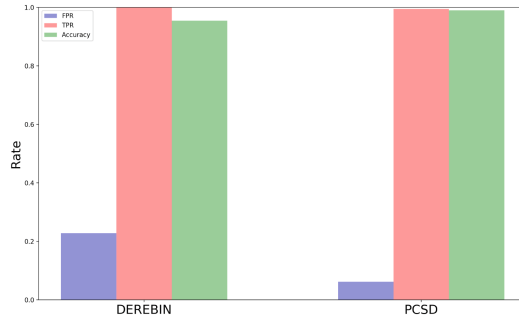
Finally, we execute different algorithms to train and test final set and take notes as Table 3. What from experiment data we can know is that *MLPRegressor* get the maximum *AUC* and *Accuracy*. Consequently, we decide *MLPRegressor* as final algorithm and leverage it to train model and then detect malware.

**Table 3.** Final result per algorithm.

Algorithm	TPR	FPR	AUC	Accuracy
SVR	0.9914529914529915	0.07	0.99029819563152899	0.9716677398583387
MLPRegressor	0.9905033238366572	0.062	0.9951396011396012	0.99029819563152899
RandomForestRegressor	0.98005698005698	0.048	0.99226590693257366	0.9710238248551192
KNeighborsRegressor	1.0	1.0	0.98795821462488131	0.6780424983902125
Ridge	1.0	1.0	0.99433618233618248	0.6780424983902125
LinearRegression	1.0	1.0	0.99140170940170946	0.6780424983902125
BayesianRidge	0.9886039886039886	0.046	0.99493447293447301	0.9774629748873149

#### 4.5 Comparison

In DREBIN [18], they used *SVM (Support Vector Machines)* algorithm and a great deal features from android applications to distinguish malevolent and innocuous applications. We compare DREBIN with our experiment and the result is shown as Fig. 6. With same test sample, we get a lower false positive, which is 0.062 and a higher accuracy concurrently, which is 0.99. From estimated data, we conclude that PCSD is effective tool for distinguishing malicious and benign applications.



**Fig. 6.** Comparison analysis.

## 5 Conclusion

Permissions, components and smali files, which come from android applications, have been utilized to detect malware by plenty of people, but almost all researchers pay little attention to statistic significance of those feature.

In our work, we leverage cluster algorithm to minimize influence of individual differences and reduce effect of random cluster by select cluster number, which has minimum volatility on size per cluster. In the future work, we are exploring more relevant influence between features and individual and extracting features from different angle to improve the detection accuracy of Android malicious applications.

## References

1. IDC. <http://www.idc.com/promo/smartphone-market-share/os>
2. Liang, S., Du, X.: Permission-combination-based scheme for Android mobile malware detection. In: 2014 IEEE International Conference on Communications (ICC), Sydney, NSW, pp. 2301–2306 (2014)
3. Huang, C.Y., Tsai, Y.T., Hsu, C.H.: Performance evaluation on permission-based detection for Android malware. In: Pan, J.S., Yang, C.N., Lin, C.C. (eds.) *Advances in Intelligent Systems and Applications - Volume 2. SIST*, vol. 21, pp. 111–120. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-35473-1\\_12](https://doi.org/10.1007/978-3-642-35473-1_12)
4. Aung, Z., Zaw, W.: Permission-based android malware detection. *Int. J. Sci. Technol. Res.* **2**, 228–234 (2013)
5. Shen, T., Zhongyang, Y., Xin, Z., Mao, B., Huang, H.: Detect Android malware variants using component based topology graph. In: 2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, Beijing, pp. 406–413 (2014)
6. Li, L., et al.: IccTA: detecting inter-component privacy leaks in Android apps. In: 2015 IEEE/ACM 37th IEEE International Conference on Software Engineering, Florence, pp. 280–291 (2015)
7. Hoffmann, J., Ussath, M., Holz, T., Spreitzenbarth, M.: Slicing droids: program slicing for smali code. In: *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC 2013)*, pp. 1844–1851. ACM, New York (2013)

8. Tang, J., et al.: NIVAnalyzer: a tool for automatically detecting and verifying next-intent vulnerabilities in Android apps. In: 2017 IEEE International Conference on Software Testing, Verification and Validation (ICST), Tokyo, pp. 492–499 (2017)
9. Tam, K., Khan, S.J., Fattori, A., et al.: CopperDroid: automatic reconstruction of Android malware behaviors. In: NDSS (2015)
10. Enck, W., Gilbert, P., Chun, B.-G., Cox, L.P., Jung, J., McDaniel, P., Sheth, A.N.: TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones. In: Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation (OSDI 2010), pp. 393–407. USENIX Association, Berkeley, CA, USA (2010)
11. Liu, X., Liu, J.: A two-layered permission-based Android malware detection scheme. In: 2014 2nd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering, Oxford, pp. 142–148 (2014)
12. Sanz, B., Santos, I., Laorden, C., Ugarte-Pedrero, X., Bringas, P.G., Álvarez, G.: PUMA: permission usage to detect malware in Android. In: Herrero, Á., et al. (eds.) International Joint Conference CISIS'12-ICEUTE'12-SOCO'12 Special Sessions. AISC, vol. 189, pp. 289–298. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-33018-6\\_30](https://doi.org/10.1007/978-3-642-33018-6_30)
13. Li, L., Bartel, A., Klein, J., Traon, Y.L.: Automatically exploiting potential component leaks in Android applications. In: 2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, Beijing, pp. 388–397 (2014)
14. Xu, K., Li, Y., Deng, R.H.: ICCDetector: ICC-based malware detection on Android. *IEEE Trans. Inf. Forensics Secur.* **11**(6), 1252–1264 (2016)
15. Chin, E., Felt, A.P., Greenwood, K., Wagner, D.: Analyzing inter-application communication in Android. In: Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (MobiSys 2011), pp. 239–252. ACM, New York (2011)
16. Wu, D.J., Mao, C.H., Wei, T.E., Lee, H.M., Wu, K.P.: DroidMat: Android malware detection through manifest and API calls tracing. In: 2012 Seventh Asia Joint Conference on Information Security, Tokyo, pp. 62–69 (2012)
17. Yerima, S.Y., Sezer, S., McWilliams, G., Muttik, I.: A new Android malware detection approach using Bayesian classification. In: 2013 IEEE 27th International Conference on Advanced Information Networking and Applications (AINA), Barcelona, pp. 121–128 (2013)
18. Arp, D., Spreitzenbarth, M., Hubner, M., et al.: DREBIN: effective and explainable detection of Android malware in your pocket. In: NDSS (2014)
19. Sahs, J., Khan, L.: A machine learning approach to Android malware detection. In: 2012 European Intelligence and Security Informatics Conference, Odense, pp. 141–147 (2012)
20. Sharma, A., Dash, S.K.: Mining API calls and permissions for android malware detection. In: Cryptology and Network Security, pp. 191–205 (2014)
21. Burguera, I., Zurutuza, U., Nadjm-Tehrani, S.: Crowdroid: behavior-based malware detection system for Android. In: Proceedings of the 1st ACM Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM 2011), pp. 15–26. ACM, New York (2011)
22. Aafer, Y., Du, W., Yin, H.: DroidAPIMiner: mining API-level features for robust malware detection in Android. In: Zia, T., Zomaya, A., Varadharajan, V., Mao, M. (eds.) SecureComm 2013. LNICSSITE, vol. 127, pp. 86–103. Springer, Cham (2013). [https://doi.org/10.1007/978-3-319-04283-1\\_6](https://doi.org/10.1007/978-3-319-04283-1_6)

23. Yuan, Z., Lu, Y., Wang, Z., Xue, Y.: Droid-Sec: deep learning in android malware detection. In: SIGCOMM Computer Communication Review, vol. 44, no. 4, pp. 371–372 (2014)
24. Spreitzenbarth, M., Freiling, F., Echtler, F., Schreck, T., Hoffmann, J.: Mobile-sandbox: having a deeper look into android applications. In: Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC 2013), pp. 1808–1815. ACM, New York (2013)
25. Lindorfer, M., Neugschwandtner, M., Weichselbaum, L., Fratantonio, Y., van der Veen, V., Platzer, C.: ANDRUBIS – 1,000,000 apps later: a view on current Android malware behaviors. In: 2014 Third International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS), Wroclaw, pp. 3–17 (2014)
26. A tool for reverse engineering Android APK files. <https://ibotpeaches.github.io/Apktool/>

# Authorship Analysis of Social Media Contents Using Tone and Personality Features

Athira Usha<sup>1,2(✉)</sup> and Sabu M. Thampi<sup>1</sup>

<sup>1</sup> Indian Institute of Information Technology and Management Kerala,  
Trivandrum, India

{athira.res, sabu.thampi}@iiitmk.ac.in

<sup>2</sup> CSE, Faculty of Engineering and Technology,  
University of Kerala, Trivandrum, India

**Abstract.** Online social networks have contributed to the countless services that ease human interaction. But the veil of anonymity has become a resort to majority of cyber criminals who indulge in unethical cyber activities. The availability of Wi-Fi hotspots and smart phones has made tracking the individuals behind the activities, a daunting task. To curtail the worst impact of these activities, one can make use of identifying authors of text contents in online social media, the only readily available imprint of an individual. Here we propose a novel authorship analysis technique applied on twitter data using tone based, personality based and stylistic features. We propose an authorship attribution scheme by training author data using Convolutional Neural Network pretrained on personality data and combines the features obtained from this model with the features obtained from another CNN architecture for tone analysis proposed by us. These features are combined together with hand crafted features pertaining to the stylistic aspects of the author and an SVM is trained on these feature combination. To the best of our knowledge this is the first work employing tone based and personality based features for attributing authorship. The new approach paves way for a fool proof authorship analysis mechanism that can be employed to curb security issues like hacked account. This is because the features chosen for our attribution method are difficult to be imitated as well as consciously controlled.

**Keywords:** Authorship analysis · Personality · Stylistics  
Convolutional neural network · Tone analysis · Personality identification

## 1 Introduction

The recent advances in online media propelled an increase in popularity of online social networks. This is fuelled by the fact that the reach of information and the rate of propagation is more in this medium than any other forms of communication. But the other side of the coin points to the most disastrous aftermath of anonymity prevailing in online environment. There is a lack of healthy means to associate trust to an individual in online sphere due to the flexibility offered by anonymity. This motivated us into considering authorship analysis as a means to identify an individual in online domain.

Authorship analysis is the task of identifying the author associated with a document by considering the stylistic particulars associated with the document. It is also used to verify a claimed authorship of a document or to identify multiple authors associated with a multi authored document [1]. Authorship attribution refers to a sub problem of authorship analysis that focuses on identifying the author of a given document who might be one among the given list of candidate authors (closed problem) or might not be (open problem). Early studies of authorship focused on long text for analysis [2]. With the gained popularity of internet and emergence of social media, the authorship analysis techniques began to be applied on emails, blog corpuses and other digital corpuses [3, 4].

With increased influence of social media and populous online social media forums, the authorship attribution became more relevant in the applications related to identification of malicious contents, online terrorism and forensic linguistics. But investigations pertaining to authorship analysis are always limited by the important concerns namely the content available for analysis and the number of candidate authors available for attribution [5]. Another factor that deters the performance of analysis process are manipulation of writing style by consciously trying to hide ones identity, named as obfuscation attack and trying to imitate other's writing style named as imitation attack [6].

Recent trends in authorship attribution tasks have inclined the focus towards identification in micro-blogs like Twitter [7, 8]. The common methods employed for detecting the stylistic cues are lexical, syntactic, semantic and content based features [9]. These features can interfere with the performance of the attribution in scenarios where the number of candidate authors is very high. In micro-blogs the content length is limited by 149 characters which can be a crucial fact in mitigating the accuracy of attribution. The statistical feature learning strategy which includes word usage patterns and n-grams, do not work effectively in such situations. This is because; in a real time scenario of analyzing micro blogs where there is huge number of candidate authors, there is a chance of overlap in the traits identified. This becomes more predominant in situations where the users discuss topics that are similar. Similar is the problem faced by these features in dealing with manipulation attacks. The common features like word usage patterns, n-grams and phrases can be easily imitated by others as well as it can be consciously controlled by an author to hide his identity. Thus impersonation on behalf of innocent victim can be an unavoidable problem.

The above said facts point to the need of capturing traits that are manifested in limited length contents as well as those which are subtle so that they are difficult to be manipulated or hidden. The method should also be performing accurately with large number of candidate authors. Thus we base our attribution method on the assumption that, in micro blogs the authors can be identified using their consistency in tone usage and manifestations of personality traits in the content authored by them. These features are difficult to be imitated or hidden as they are indirect aspects unlike other customary features. This in fact motivated us into considering the tone usage pattern and personality trait manifestation in a person's writing as features for author identification. We propose a method that employs identification of tone of a user and personality trait apparent in the tweets of a user. The method also applies deep learning techniques to determine features relevant to these traits. This is used in unison with hand crafted



features, thereby resulting in better modeling of author traits pertaining to tweets. Our method faces a limitation in terms of requirement of number of tweets per author. But the approach performs better in comparison with exiting methods in attribution task associated with scenarios involving large number of candidate authors.

The contributions of our work can be highlighted as follows:

- To the best of our knowledge our work is the first authorship analysis task employing personality and tone features.
- We propose a transfer learning based CNN to identify tones of a tweet which is capable of identifying emotions, humor and sarcasm content using a single architecture.
- We employ a combination of handcrafted features and features obtained from CNN pretrained models thus exerting the advantage of transfer learning.
- The new approach makes authorship analysis task applicable in micro blogs by utilizing inherent useful details that have not yet been explored in this area thus making use of features that are difficult to be imitated or consciously controlled.
- The proposed approach performs better with increased number of candidate authors in comparison with existing methods.

The paper is further organized as follows. Section 2 elaborates on recent literature. Section 3 discusses the methodology, Sect. 4 discusses the experiments. Results are discussed in Sect. 5. The paper concludes in Sect. 6.

## 2 Literature Review

The authorship attribution techniques have shown a noteworthy transition from statistical methods [10] to machine learning based approaches [11].

The increase in the popularity of social media has paved way for the researchers to lay focus on authorship attribution in micro-blogs. Several recent works have been published regarding the application of authorship analysis in online tweets. In [12] Layton et al. discusses the effectiveness of using SCAP (Source code authorship profile) method for attributing authorship in micro blogs. It is inferred that there is no considerable improvement in accuracy even if the number of tweets is increased beyond the threshold value. The paper defines a threshold number of tweets to be considered for attribution task. In [7], Bhargava et al. discuss authorship attribution in tweets by employing lexical, syntactic, tweet specific and emoticon features as author style. Then the model is trained using different classifiers and performance has been compared.

In [13] Albadarneh et al. present a comprehensive means of authorship in large scale Arabic tweets by finding term frequency inverse document frequency weights associated with bag of words. It employs big data analytics techniques to deal with large scale data. In [15] Barbon et al. discuss about analyzing the authorship traits to confirm the legitimacy of twitter accounts. The lexical, syntactic, idiosyncratic and content specific features are utilized for the purpose. In [16], Macke and Hirshman applies deep learning techniques at the sentence level to identify traits in multi authored documents. The authors model the vocabulary and grammatical structure using

recurrent neural network model (RNN) which is observed to hold a lesser performance with increase in number of authors.

In [24] Rappoport et al. discuss about authorship attribution in tweets by considering unique signature associated with users. In [14] the authors compare the authorship attribution performance of different classifiers fed with traditional stylometric features in unison with social media specific features. It also analyses the performance based on voting algorithm that combines all classifiers. In [8] a comprehensive review of existing authorship analysis techniques in micro blogs is presented and the paper concludes with an inference that calls for the requirement of a plenary method that makes use of the data context and process it irrespective of its multimodality and a system that is tolerant to the lack of availability of all author data at the time of training has been proposed.

The methods mentioned in the literature focus on author identification using lexical, syntactic, semantic and content based features. These methods rely on large number of tweets collected per author, which might not be feasible for all cases as in the case where frequency of tweets by candidate authors is less. There is also drop in the performance level with the increase in candidate authors. The methods also rely on features that can be manipulated by traitors. Thus the gap in the recent literature calls for the need of a method resilient to the amount of content available for analysis, as well as number of candidate authors available for analysis and also features that are difficult to be manipulated. This motivates us into considering personality, tone and stylistic features for analyzing authorship.

We propose an authorship attribution method by modeling the tone and personality patterns associated with an author. This is obtained by using convolutional neural network trained on tone and personality data. The author data is then applied on these models. The final level features are combined with psycholinguistic features. These features are used to train a linear SVM, which predicts the author of an unknown tweet. We compare the result obtained with the state of the art techniques and observed an improved performance in comparison with benchmark. This is mainly due to the fact that we make use of personality and tone based features in combination with stylistic features to result in improved performance. We make use of deep learned features, which makes it even more suitable than other methods using handcrafted features in environment of large number of candidate authors.

### 3 Personality, Tone and Authorship

Personality is characterized by a person's pattern of thought, emotion and the way of action. There are five traits of personality also known as five factor model which can be listed as follows Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism [17].

The texts written by an author lays insight into the personality of the author. Campbell and Pennebaker explains the importance of words used in text in embodying a person's thought pattern [18]. The relative usage pattern of pronouns reflects the inherent security concerns of an author [19]. Thus, we can utilize the consistencies associated with the personality traits reflected through the text as an effective style

marker. Tracking the personality trait pattern can be an effective style emblem that is resistant towards an impersonation attack on behalf of innocent victim. Since this is an exquisite feature, imitating this can be arduous, thus making the authorship analysis task much more effective in comparison with existing methods. Our proposed approach gains motivation from this fact and this prompts us to use personality as an effective measure of style marker. This feature can be effectively captured from short texts and are difficult to be imitated and controlled. A person in an online sphere unconsciously leaves traces of personality aspects in the contents that are delivered by him. Our method captures these traits and the pattern is saved as a stylistic aspect associated with an author. This can also help to mitigate imitation of author traits. We use a dataset labeled on five personality traits to create a pretrained model on personality and uses this model to generate personality features of author data by feeding the text of authors.

Another factor which we are considering is tone associated with the author. Tone refers to the attitude of the authors to the subject of a document. Usually tone is plighted towards gaining attention of specific audiences. In many cases users follow particular tone of content delivery. For example some people often tweet in a sarcastic manner while there are others who always tweet seriously. These types of tone patterns can be utilized to model the stylistic aspect of a person. Tone basically embarks on a person's way of presentation. Just like the tone associated with speech, each person is unique in terms of the tone of text content authored by him. This consistency in tone of the content by a user can be captured to depict another authorship trait. Another striking fact is that, adopting other's tone pattern or hiding one's own innate tone pattern are difficult. Hence, these features can be effectively put into use for effective analysis to restrict illegal cyber activities like impersonation. We propose a tone analysis technique using transfer learning to predict tones of a tweet. The tones which we are considering for this particular work are anger, sad, disgust, fear, surprise, joy humor and sarcasm.

These are the tones considered because other tones like angry, disgust, love etc. can be subjective. The presence of these tones is influenced by the context under consideration and does not remain constant. But an author can often toggle between the above mentioned 5 tones. This pattern is captured to identify the user behavior, hence can be used as a style marker. There can be question of mood of a writer and the tone followed. But the inherent tendency of an author to follow a pattern in tone can be portrayed as his writing style. This then used in combination with stylistic features can be a useful direction towards an accurate author prediction scheme.

## 4 Methodology

Based on the extensive literature survey conducted, we inferred a need for an authorship analysis technique based on features that are evident in short text and proves effective even in case of increased number of candidate authors and those which are difficult to be manipulated. Thus we use a combination of customary features together with tone and personality features. The overall methodology can be depicted as in Fig. 1. The method can be explained as follows. The tweets are converted into its corresponding glove representation and then we use two pretrained models. One trained on personality data and other is the tone identification architecture. We feed the

author specific tweets represented using Glove based vectors, into personality model and tone identification CNN architecture. The features extracted from these models are combined together with handcrafted stylistic features and they are fed to the SVM for classification. The sections are explained below.

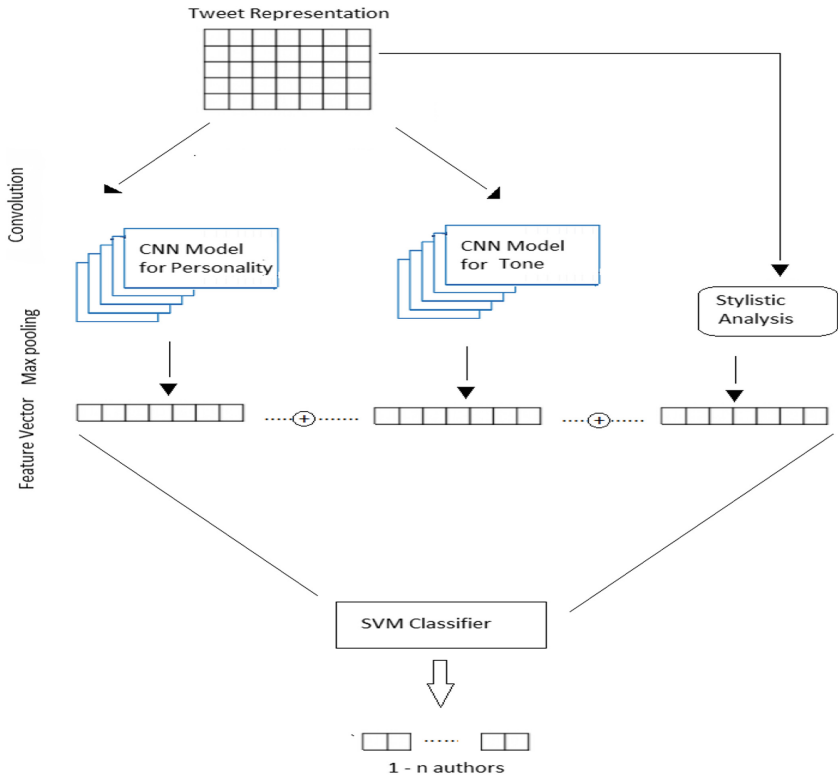


Fig. 1. Proposed authorship attribution technique

### 4.1 Dataset

The dataset for authorship analysis task carried out in this work is chosen to be tweets from Twitter. The process of data acquisition was challenged by the limitation of twitter server to handle multiple requests at the same time. There was limitation imposed on the number of tweets collected. The technical feasibility caused trouble in collecting tweets older than 10 days. Thus data collection task was a several days’ process. We identified suitable twitter users using suitable resources<sup>1</sup>. Then a corpus compilation tool provided in [21] was utilized for the purpose of collecting tweets from concerned users. The dataset comprises of 800 tweets from 200 users. The tweets were

<sup>1</sup> <http://cs229.stanford.edu/proj2012/CastroLindauerAuthorIdentificationOnTwitter.pdf>.

manually examined for the suitability in experimentations. Some tweets were too short of length to be subjected to analysis. Such tweets were combined together to form single tweet. Such modification interferes with the tweet specific stylistic features like words per tweet, punctuation per tweet etc. Such measures were obtained by averaging the values for the tweets combined, so as to maintain noise free dataset capable of modeling an author's writing style.

## 4.2 Preprocessing

The initial phase of the process is cleaning the tweets of irrelevant items and re-tweets. We tried to normalize the tweets among all authors by trying to include tweets of related topics. This is because tremendously varying topics among the authors can result in unreliable training samples that give way to deceptively accurate classification models. The major concern was to get rid of hyperlinks, and retweets and tweets containing quoted texts, as these tweets are non-representative of an author's writing style. Once the tweets have been preprocessed they are subjected to undergo the process of author trait modeling which is explained the next section. The preprocessing steps have been completed using in built functionalities of Natural Language Tool Kit for python [22].

## 4.3 Model Generation

Each tweet is represented by a matrix of size  $n$  by  $d$  when  $n$  refers to the number of words and  $d$  refers to the dimension of the word embedding used. We use publicly available Glove pretrained word vectors for twitter [25]. We fix the length of the tweet to be  $n$  which is chosen to be the maximum word count among all tweets. All the tweets are zero padded to obtain a fixed length. The matrix so formed can be subjected to convolution. The operation involves applying a filter ( $W \in \mathbb{R}^{hd}$ ) over the matrix. Here  $h$  denotes the number of consecutive words chosen to obtain the scalar  $c_i$ , also known as window size. The width of the filter is fixed as the dimension of the embedding. Scalar is defined as the value obtained from a nonlinear function of bias  $b$  applied over the filter.

$$c_i = f(w : x_{i+h-1} + b) \quad (1)$$

This value of  $i$  varies from 1 to  $n - h + 1$  resulting in the application of filter over all windows. Here  $w$  refers to the word vector. The values of scalar features produced out of this action are subjected to a max-over-time pooling operation, to get feature  $\hat{c}$  which possibly suggest the important feature.

$$\hat{c} = \max(c_1, c_2, c_3, \dots, c_{n-h+1}) \quad (2)$$

This is repeated for many filters thereby resulting in large feature vector which can be finally fed to the linear layer. The regularization is brought into avoid co-adaptation by means of dropout [20] at penultimate layer.

#### 4.4 Personality Prediction Model

We built a pretrained model on personality by training a convolutional neural network based on gold standard personality data that comprises 2400 essays labeled with personality traits [26]. The file was obtained from Mypersonality project site [27]. The training was carried out as mentioned in [28]. The pretrained model is fed with tweets and final layer is used to concatenate with other features to be fed to a final linear SVM classifier. The pretrained personality model when trained on author data, extracts relevant personality features specific to authors. The experiments have been conducted for several epochs to decide on the best one to attain perfect convergence. The pattern of personality pertaining to an author includes the unique manner in which each trait is displayed in his writing style. An author may have a pattern of consistency associated with the display of personality in his writing. For a person who exhibits blend of personality traits may have a pattern of weights associated with each of the personality models. This is identified at the final layer of pretrained personality model trained on author specific data. We capture these features in the final layer, which are the author specific personality features and this is fed to the SVM classifier to train on author personality pattern.

#### 4.5 Tone Prediction Model

We propose a Convolutional Neural Network architecture employing transfer learning for performing tone analysis. The tones to be identified are anger, fear, sad, disgust, surprise, joy, humor, sarcasm and neutral. The tones anger, fear, sad and disgust can be depicted as a negative valenced reaction to an event, while joy and surprise can be considered as positive valenced reaction to something. We extend two tones namely sarcasm and humor in addition to the emotions. Humor can be considered to be an incongruous positive and negative association. The presence of humor can be indicative of sarcasm in the text. Thus humorous features can be a contributing factor in the identification of sarcasm in a tweet. Thus the model pretrained on the humor can be utilized for the purpose of creating a pretrained model on sarcasm data. The model finally obtained has learned the traits of emotions, humor and sarcasm. Thus we create a model that is trained on tasks which are otherwise different from each other. The final model so obtained is used to train data that contains tweets belonging to all categories under consideration namely: anger, sad, fear, disgust, joy, surprise, humor, sarcasm and neutral. Thus the final model trained on entire data pertaining to different tones, predicts the tone categories of an unknown instance. The entire process can be depicted by the overall architecture given in Fig. 2. The whole process can be divided into two phases namely summarization and transfer learning. The following sections explain each step in detail.

##### Summarization

The process of tone analysis begins with pretraining a Convolutional Neural Network based on sentimental data. While feeding the data to the network we need to ensure that the length of the tweets is same. Since we are performing the experiments on varied length tweets, the first step is to ensure uniform input dimension. This can be attained

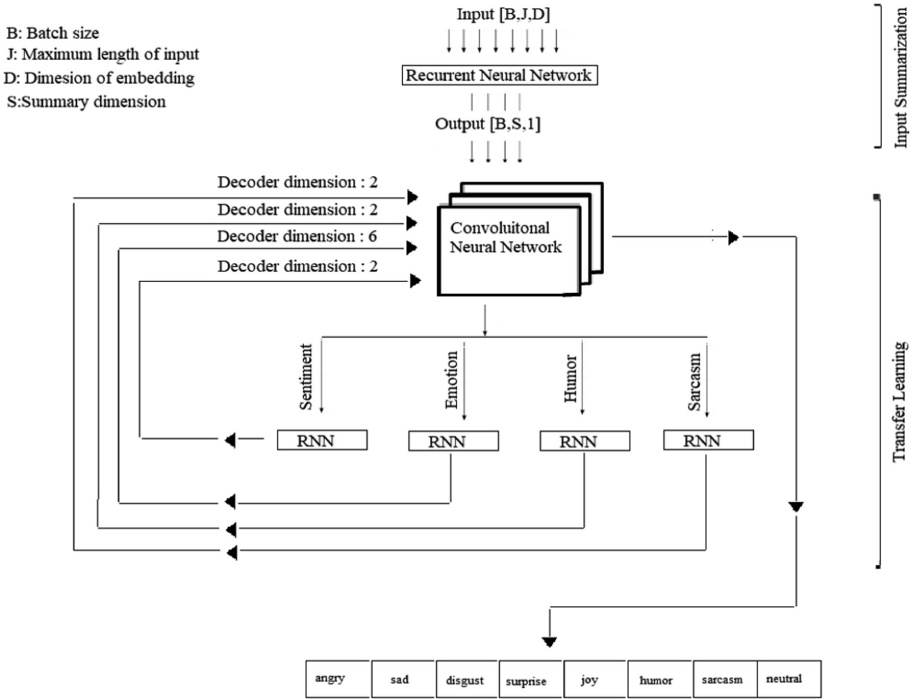


Fig. 2. The convolutional network architecture for tone analysis.

by performing an abstractive summarization of all tweets input, to form a summary of common dimension. We use a gated recurrent neural network to perform the task of summarization [29].

**Transfer Learning**

The proposed work employs the concept of transfer leaning to learn traits regarding multiple tasks that are inherently different from each other. Usually the task of emotion detection, sarcasm detection and humor detection are treated separately. Humor and sarcasm detection are to date treated as binary classification tasks. Here we create a single pretrained model capable of performing the task of identifying these tones (emotion, sarcasm and humor) as a single task. The overall process can be summarized as follows. We are applying transfer learning approach on CNN. We use a single Convolutional Neural Network that learns from different data. Initially we pretrain the CNN based on sentimental data. This pretrained model is used for initializing the next model to be trained on emotion data. Then the model is used to initialize another model which is to be trained on humor data. Then we use this model to initialize a model on sarcasm data. Finally we feed a combination of data comprising of labels angry, sad, disgust, fear, joy, surprise, humorous, sarcasm and neutral. Then the model is trained on this data. We keep all the layers of the network unfrozen during each phase, thus enabling the network to learn features pertaining to all models. In order to avoid

complex coadaptation each pretraining procedure involves the addition of a dropout layer. The subsequent sections explain different models in our tone analyzer.

### **Sentiment Model**

The sentiment model is obtained by pretraining the CNN using the data labeled with positive and negative sentiment. We obtained data from several open challenge task namely benchmark dataset in SemEval 2013 [30] and Semeval 2014 [31] twitter sentiment dataset for training our sentimental model. We combine some data from our emotion corpus by analyzing its sentiments by using Python toolkit TextBlob [32]. We also utilized the data from twitter sentiment corpus made available online for research purpose<sup>2</sup>. The combined dataset comprises of 4,98,586 data samples labeled with sentiments. The CNN is initially fed with sentiment specific data. The model is trained to learn the sentiment specific features pertaining to the data. The model converges to an optimal fine tuning level. This is used to initialize the next model based on emotion.

### **Emotion Model**

The pretrained model obtained using the sentiment data is used to train the emotion model. Here all the layers are left unfrozen, meaning that we are using the weights obtained from this unit as the weight for initializing the next model of emotion. There are 7 categories of emotion being addressed here. They are anger, fear, sad, disgust, surprise, joy and neutral. The dataset chosen for the task of identification have been compiled from three sources namely emotion dataset from crowd flower<sup>3</sup>, ISEAR Dataset<sup>4</sup> and emotion annotated corpus developed by Aman and Szpakowicz [33]. The final corpus compiled by us included 40,552 short sentences labeled with emotions pertaining to the categories anger, fear, sad, disgust, surprise, joy and neutral. The emotion annotated corpus contributed to the following labels in our task: angry, sad, surprise, fear, disgust, neutral and joy. We assumed the label happy given in the dataset to be the same as that of joy and have been included as the part of our corpus. The ISEAR data provided us with data for the labels, happy, sad, fear and angry. The crowd flower dataset contributed to sad, anger, joy, surprise and disgust categories. We assumed that the sentences belonging to happy to be of class joy and sentences belonging to hate in the dataset as class disgust in our corpus. The model is trained using this data and the final model obtained is used for initializing the next model for humor detection.

### **Humor Model**

We construct a model for humor by utilizing the weights obtained from our previous model on emotion. The humorous dataset we are considering comprises of funny jokes and humorous tweets. We have not included sarcastic data to our humor corpus. This is particularly due to the fact that, we have used sarcasm and humor as separated labels for tone analysis. A sarcastic note is often humorous while a humorous instance need not be sarcastic always. Thus our dataset for humor has been compiled cautiously to avoid sarcasms. This is to make distinction between sarcasm and humor. The dataset

---

<sup>2</sup> <http://www.sananalytics.com/lab/twitter-sentiment/>.

<sup>3</sup> [https://www.crowdfunder.com/wp-content/uploads/2016/07/text\\_emotion.csv](https://www.crowdfunder.com/wp-content/uploads/2016/07/text_emotion.csv).

<sup>4</sup> <http://emotion-research.net/toolbox/toolboxdatabase.2006-10-13.2581092615>.



has been collected using the scraping tool available online for creating humor dataset<sup>5</sup>. The inference based on the studies conducted in [34] and [35] it is clear that emotion based and sentiment based features play an important role in humor identification. Thus we use this model to train data specific to humor content. The model is used to initialize the next model based on sarcasm detection.

### **Sarcasm Model**

Sarcasm detection is a task that is highly dependent on the sentimental and emotional aspects of text content. Many previous work have laid focus on the importance of sentimental and emotional aspects in determining sarcasm [36, 37]. Thus we model our sarcasm detection model on top of the pretrained model obtained from above steps. We train the model using the corpus compiled by us which include sarcastic dataset published as a part of Pacific Asia Knowledge Discovery and Data mining Conference 2016<sup>6</sup>. The dataset comprises of public Reddit comments of sarcastic nature. We combined the dataset with sarcastic tweets mined by us using Twitter API to collect tweets with hash tag sarcasm. The tweets have been manually monitored to include genuine sarcastic tweets. Those tweets that require contextual knowledge beyond the text content have been removed. The non sarcastic tweets have been obtained by considering neutral tweets used for emotion, sentiment and humor. The final corpus comprised of about 9128 sarcastic data and 10,000 non sarcastic data. Training the model using this data completes the process of formulating our architecture.

### **Tone Identification**

The final phase of the tone analysis involves identifying the tone of a tweet. The final architecture obtained after performing the transfer learning is now fed with all the data used for training different models. This adapts the network to a generalized model capable of predicting tone specific to a tweet. This architecture is utilized for feeding the user tweet to identify tone of the tweet. The final layer of the architecture gives deep features specific to the tone of a user. These features are used in unison with personality features and stylistic features to predict the author traits.

## **4.6 Stylistic Features**

Using the personality and tone based features alone might affect the credibility of the attribution as there might be coincidental similarity of personality similarity. Similar is the case with tones. This can be overcome by considering the lexical and syntactic features prominent in person's writing style. We also consider psycholinguistic features associated with users. We use LIWC (Linguistic Inquiry Word Count) [38] to identify common lexical features and psycholinguistic features associated with author's tweet. The features included are depicted in Table 1. The choice of these features has been arrived at on the basis of the inferred results from existing literature [7, 8, 16, 39]. Based on this, we conducted experiments to arrive at conclusion regarding the most

<sup>5</sup> <https://github.com/CrowdTruth/Short-Text-Corpus-For-Humor-Detection>.

<sup>6</sup> <http://www.parrotanalytics.com/pacific-asia-knowledge-discovery-and-data-mining-conference-2016-contest/>.

suitable feature set for our dataset. The feature dimensionality reduction was performed using Principle Component Analysis [23]. Based on the cross validation result obtained we infer that the combination of the following features can result in better accuracy of authorship attribution in our dataset. Words per tweet, Punctuation per tweet, Words per sentence, 2–6 lettered words, Ratio of alphanumeric characters over normal words, Informal Language usage, Special character count, Uppercase letters count, Character n-grams, Dictionary word count, Capitalization of sentence beginning, POS n-grams.

The features generated in the manner described in the previous sections are concatenated and fed to the Support Vector Machine classifier. The choice of SVM was made due to the fact that the classifier is based on the concept of generating maximum margin hyper plane. This enables better generalization and accuracy is improved. The performance measure is computed by conducting a 10-fold cross validation on the dataset being considered.

## 5 Experimental Results and Discussion

The experiments have been conducted to prove the effectiveness of the novel means of authorship attribution using personality and tone features used in combination with stylistic features. The experiments were carried out in different steps. Authorship analysis task was performed using personality based, tone based features and stylistic features and it was compared against stylistic features used in recent literature. The pretrained models have been used to extract tone related features and personality based features. The features so obtained are combined with stylistic features to obtain linear features which are fed to an SVM that eventually gets trained on author characters based on these features. The experiments have been performed with stylistic features alone and then in combination with tone based and personality based features.

The accuracy in the attribution has been calculated in terms of precision, recall and F-score measure. The importance of personality based and tone based features in providing an added advantage to the existing stylometric means can be established by comparing the performance of authorship attribution with and without including these features. The result obtained for such an experiment conducted in an optimum set up (800 tweets per user) is as shown in Table 1.

The illustration in Table 1 clearly shows that the precision, recall and F-score measure associated with the attribution improves with the inclusion of tone based and personality based features. At the same time, tone based and personality based features used alone cannot contribute to the attribution. This is because, there can be tone and personality overlap among individuals. This worsens even more when the candidate authors increase. But combination of all features yields a better result. We have not included the attribution result obtained for tone based feature and personality based feature used alone. This is because the accuracy is very less in that scenario. Similar is the case with other combinations of features for example combination of stylistic features and tone based. The results are lesser in comparison with the combination of all three. Thus to illustrate the importance of tone based and personality based features. We have included only the results that are worth to be noticed, considering the page limitation.

**Table 1.** Evaluating the importance of including tone, personality based and stylistic features

Users	Precision	Recall	F-score	Tone	Personality	Stylometry
15	<b>0.85</b>	<b>0.79</b>	<b>0.81</b>	✓	✓	✓
	0.78	0.67	0.72			✓
	0.49	0.51	0.50	✓	✓	
50	<b>0.77</b>	<b>0.69</b>	<b>0.70</b>	✓	✓	✓
	0.69	0.60	0.64			✓
	0.39	0.45	0.41	✓	✓	
100	<b>0.66</b>	<b>0.52</b>	<b>0.58</b>	✓	✓	✓
	0.49	0.50	0.50			✓
	0.36	0.39	0.41	✓	✓	
200	<b>0.59</b>	<b>0.63</b>	<b>0.54</b>	✓	✓	✓
	0.46	0.46	0.47			✓
	0.43	0.33	0.34	✓	✓	

The performance variation of the proposed system under different number of tweets available per user is as shown in Table 2. The table shows the precision, recall, F-score and accuracy associated with the attribution using all three features for different number of users and with varied number of tweets per author. The results illustrated by Table 2 shows that the accuracy increases in case where the number of tweets per user are more. The accuracy also increases with less number of candidate authors. Thus we have an optimum performance in case where number of authors is 15 and tweets per author is 800. This can be attributed to the fact that the pretrained models on personality and tone gets trained inappropriately when the training data available for authors are less. This lessens the accuracy. But it can be seen from the results that the system attains considerable accuracy even when the number candidate authors increases. This shows the effectiveness of the personality based and tone based features in attributing authorship.

**Table 2.** Experimental results

Users	Tweets per user	Precision	Recall	F-score	Accuracy
15	250	0.55	0.50	0.52	51%
	800	0.85	0.79	0.81	80%
50	250	0.50	0.48	0.49	50%
	800	0.77	0.69	0.70	71%
100	250	0.44	0.49	0.46	47%
	800	0.66	0.52	0.58	58%
200	250	0.39	0.31	0.34	38%
	800	0.59	0.63	0.54	53%

To evaluate the performance of our proposed approach in comparison with the state of the art techniques, we consider three benchmark methods. Their experimental set up and results are as shown in the Table 3. From Table 3 we get to know the accuracy associated with the three baseline methods considered. The table shows the result obtained for the scenarios mentioned in the literature. Comparing the accuracy in similar set up, we can see that our method lacks in accuracy. But when we simulated the methods for our dataset using the optimal condition, our method performs better in comparison with existing methods. For this we simulated the experimental setup using our optimal condition. The following set of baseline methods was created.

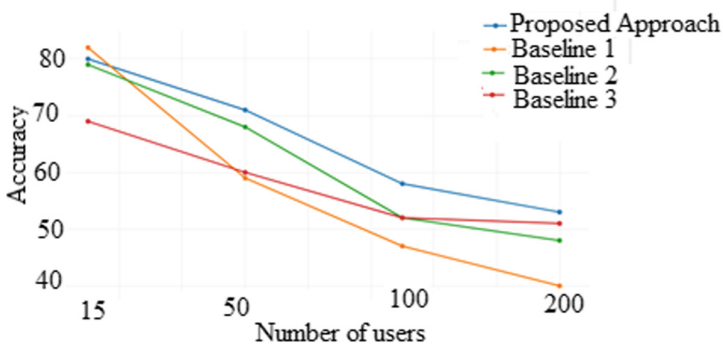
Baseline 1: We obtained lexical syntactic and tweet specific features mentioned in [7] and classified them using SVM classifier.

Baseline 2: We extracted word based and character based features specified in [24] and classified them using an SVM classifier.

Baseline 3: We utilized the features mentioned in [8] and employed PMSVM to classify the resulting data. The result obtained in this case by considering 800 tweets per user is as given in Fig. 3.

**Table 3.** The performance details of state of the art techniques

Method	No of users	Tweets per user	Features	Classifier	Accuracy
Bhargava et al. [7]	20	300	Lexical, Syntactic, Tweet specific	SVM	64.54%
Rappaport et al. [24]	50	300	Unique signature based	SVM	64%
Rocha et al. [8]	50	1000	Word and char n-grams	PMSVM	65%



**Fig. 3.** The comparison of proposed method with existing methods

From the result illustrated in Fig. 3 it can be inferred that our proposed approach works better in comparison with exiting methods in the scenario of large number of candidate authors. But our approach lacks behind in accuracy associated with

attribution scenario where number of tweets per user is less. This is due to the fact of enough content to train using a convolutional pretrained model. Since the attribution accuracy is seen to be better using our proposed approach, we can infer the effectiveness of personality based and tone based features. The demerit associated with the scheme being lack of enough data to train pretrained convolutional network. A possible solution to this could be using hand crafted features related to personality and tone based aspects of a user. The utilization of tone based and personality based features can be a suitable choice in an environment to detect compromised accounts. This is because as mentioned in introduction section these are features that cannot be imitated or consciously hidden by an individual.

## 6 Conclusion

The paper presents a novel attribution method using tone based, personality based and customary stylometric features. The features prove to be an effective means to attribute authorship in such a way to resist imitation and conscious hiding of the writing style. The method proved to be accurate and performed well in comparison with existing methods in case of large number of candidate authors. But the accuracy lessens when the amount of content available for each author decreases. This can be overcome by using hand crafted features related to personality and tone associated with a person's writing style.

## References

1. Juola, P.: Authorship attribution. *Found. Trends® Inf. Retr.* **1**(3), 233–334 (2008)
2. Mosteller, F., Wallace, D.L.: *Inference and Disputed Authorship: The Federalist*. Addison-Wesley, Boston (1964)
3. Luyckx, K., Daelemans, W.: Authorship attribution and verification with many authors and limited data. In: *Proceedings of the 22nd International Conference on Computational Linguistics*, vol. 1, pp. 513–520. Association for Computational Linguistics, August 2008
4. Jockers, M.L., Witten, D.M.: A comparative study of machine learning methods for authorship attribution. *Lit. Linguist. Comput.* **25**(2), 215–223 (2010)
5. Luyckx, K., Daelemans, W.: The effect of author set size and data size in authorship attribution. *Lit. Linguist. Comput.* **26**(1), 35–55 (2010)
6. Brennan, M.R., Greenstadt, R.: Practical attacks against authorship recognition techniques. In: *IAAI*, 14 July 2009
7. Bhargava, M., Mehndiratta, P., Asawa, K.: Stylometric analysis for authorship attribution on Twitter. In: Bhatnagar, V., Srinivasa, S. (eds.) *BDA 2013*. LNCS, vol. 8302, pp. 37–47. Springer, Cham (2013). [https://doi.org/10.1007/978-3-319-03689-2\\_3](https://doi.org/10.1007/978-3-319-03689-2_3)
8. Rocha, A., Scheirer, W.J., Forstall, C.W., Cavalcante, T., Theophilo, A., Shen, B., Stamatatos, E.: Authorship attribution for social media forensics. *IEEE Trans. Inf. Forensics Secur.* **12**(1), 5–33 (2017). IEEE
9. Stamatatos, E.: A survey of modern authorship attribution methods. *J. Assoc. Inf. Sci. Technol.* **60**(3), 538–556 (2009). Wiley Online Library
10. Burrows, J.F.: Word-patterns and story-shapes: the statistical analysis of narrative style. *Lit. Linguist. Comput.* **2**(2), 61–70 (1987)

11. Koppel, M., Schler, J., Argamon, S.: Computational methods in authorship attribution. *J. Assoc. Inf. Sci. Technol.* **60**(1), 9–26 (2009)
12. Layton, R., Watters, P., Dazeley, R.: Authorship attribution for Twitter in 140 characters or less. In: 2010 Second Cybercrime and Trustworthy Computing Workshop (CTC), pp. 1–8. IEEE, July 2010
13. Albadarneh, J., Talafha, B., Al-Ayyoub, M., Zaqabeh, B., Al-Smadi, M., Jararweh, Y., Benkhelifa, E.: Using big data analytics for authorship authentication of arabic tweets. In: IEEE/ACM International Conference on Utility and Cloud Computing, pp. 448–452. IEEE (2015)
14. Li, J.S., Chen, L.C., Monaco, J.V., Singh, P., Tappert, C.C.: A comparison of classifiers and features for authorship authentication of social networking messages. *Concurr. Comput.: Pract. Exp.*, **29**(14) (2017)
15. Barbon, S., Igawa, R.A., Bogaz Zarpelão, B.: Authorship verification applied to detection of compromised accounts on online social networks. *Multimed. Tools Appl.* **76**(3), 3213–3233 (2017)
16. Macke, S., Hirshman, J.: Deep Sentence-Level Authorship Attribution (2015). CS224
17. Digman, J.: Personality structure: emergence of the five-factor model. *Ann. Rev. Psychol.* **41**, 417–440 (1990)
18. Campbell, R.S., Pennebaker, J.W.: The secret life of pronouns: flexibility in writing style and physical health. *Psychol. Sci.* **14**(1), 60–65 (2003)
19. Pennebaker, J.W., Chung, C.K.: Computerized text analysis of Al-Qaeda transcripts. In: A Content Analysis Reader, pp. 453–465 (2008)
20. Hinton, G.E., et al.: Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint [arXiv:1207.0580](https://arxiv.org/abs/1207.0580) (2012)
21. Twitter corpus (2015). [https://github.com/bwbaugh/twitter-corpus/blob/master/twitter\\_corpus.py](https://github.com/bwbaugh/twitter-corpus/blob/master/twitter_corpus.py). Accessed 28 July 2017
22. Bird, S., Klein, E., Loper, E.: *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O'Reilly Media Inc., Sebastopol (2009)
23. Wold, S., Esbensen, K., Geladi, P.: Principal component analysis. *Chemometr. Intell. Lab. Syst.* **2**(1–3), 37–52 (1987)
24. Rappoport, A., Schwartz, R., Tsur, O., Koppel, M.: Authorship attribution of micro-messages, July 2013. [http://u.cs.biu.ac.il/~koppel/papers/twitter\\_authorship\\_emnlp.pdf](http://u.cs.biu.ac.il/~koppel/papers/twitter_authorship_emnlp.pdf)
25. Pennington, J., Socher, R., Manning, C.D.: GloVe: Global Vectors for Word Representation (2014)
26. Pennebaker, J.W., King, L.A.: Linguistic styles: language use as an individual difference. *J. Pers. Soc. Psychol.* **77**(6), 1296–1312 (1999)
27. Celli, F., Pianesi, F., Stillwell, D., Kosinski, M.: Workshop on computational personality recognition (shared task). In: Proceedings of WCPRI3, in Conjunction with ICWSM 2013 (2013)
28. Majumder, N., Poria, S., Gelbukh, A., Cambria, E.: Deep learning-based document modeling for personality detection from text. *IEEE Intell. Syst.* **32**(2), 74–79 (2017)
29. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint [arXiv:1406.1078](https://arxiv.org/abs/1406.1078) (2014)
30. Nakov, P., Rosenthal, S., Kozareva, Z., Stoyanov, V., Ritter, A., Wilson, T.: Semeval-2013 task 2: sentiment analysis in Twitter. In: Proceedings of the International Workshop on Semantic Evaluation, vol. 13 (2013)

31. Rosenthal, S., Ritter, A., Nakov, P., Stoyanov, V.: Semeval-2014 task9: sentiment analysis in Twitter. In: International Workshop on Semantic Evaluation (SemEval 2014), pp. 73–80 (2014)
32. Loria, S., Keen, P., Honnibal, M., Yankovsky, R., Karesh, D., Dempsey, E.: TextBlob: simplified text processing. Secondary TextBlob: Simplified Text Processing (2014)
33. Aman, S., Szpakowicz, S.: Using roget's thesaurus for fine-grained emotion recognition. In: IJCNLP, pp. 312–318 (2008)
34. Chen, L., Lee, C.M.: Convolutional neural network for humor recognition. arXiv preprint [arXiv:1702.02584](https://arxiv.org/abs/1702.02584) (2017)
35. Bertero, D., Fung, P.: A long short-term memory framework for predicting humor in dialogues. In: HLT-NAACL, pp. 130–135 (2016)
36. Poria, S., Cambria, E., Hazarika, D., Vij, P.: A deeper look into sarcastic tweets using deep convolutional neural networks. In: COLING, Osaka, pp. 1601–1612 (2016)
37. Ghosh, A., Veale, T.: Fracking sarcasm using neural network. In: WASSA@ NAACL-HLT, pp. 161–169 (2016)
38. Tausczik, Y.R., Pennebaker, J.W.: The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* **29**(1), 24–54 (2010)
39. Keretna, S., Hossny, A., Creighton, D.: Recognizing user identity in Twitter social networks via text mining. In: IEEE International Conference on Systems, Man, and Cybernetics, pp. 3079–3082. IEEE (2013)

# Privacy-Preserving Handover Authentication Protocol from Lightweight Identity-Based Signature for Wireless Networks

Changji Wang<sup>1(✉)</sup>, Shengyi Jiang<sup>1</sup>, and Yuan Yuan<sup>2</sup>

<sup>1</sup> School of Information Science and Technology,  
Collaborative Innovation Center for 21st-Century Maritime Silk Road Studies,  
Guangdong University of Foreign Studies, Guangzhou 510006, China  
wchangji@gmail.com

<sup>2</sup> School of Finance, Guangdong University of Foreign Studies,  
Guangzhou 510006, China

**Abstract.** Many handover authentication protocols for wireless networks have been proposed in recent years. However, most of them are either inefficient or insecure, or lack sufficient concern for user privacy. In this paper, we propose a new design idea for handover authentication protocol by using pseudonym-based cryptography and a variant of the SIGMA (Sign-and-MAC) approach, and present a new handover authentication protocol for wireless networks from elliptic curve cryptography. The proposed protocol satisfies all desirable basic security and privacy requirements, including perfect forward secrecy, user anonymity and untraceability, conditional privacy preservation and so on. The performance analysis shows that our proposed protocol is more efficient than previous handover authentication protocols in terms of computation and communication.

**Keywords:** Wireless networks · Handover authentication  
Identity-based signature · Authenticated key establishment  
Elliptic curve cryptography

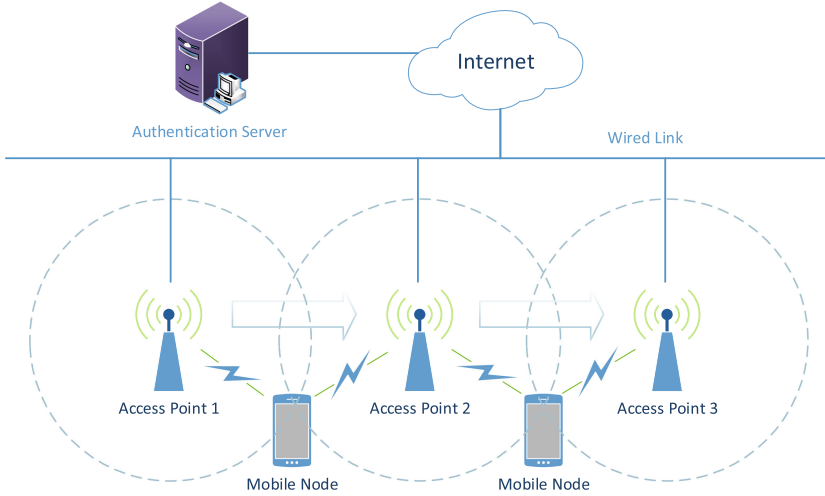
## 1 Introduction

With the rapid development of wireless internet access techniques, more and more mobile services appeared, which provide a more convenient life to people [1–4]. For example, wireless body area networks can alter the future of health-care services by enabling ubiquitous monitoring of patients. To overcome the geographical coverage limit of each access point, seamless handover over multiple access points is highly desirable to mobile nodes [5].

A typical handover authentication scenario involves three actors: an authentication server (AS), mobile nodes (MNs) and access points (APs). An MN registers to the AS, and then connects to any AP to access its subscription services.



An AP acts like a guarantor for vouching for an MN as a legitimate subscriber. When an MN moves from the current AP (e.g.,  $AP_1$ ) into a new AP (e.g.,  $AP_2$ ), it will trigger the execution of handover authentication at  $AP_2$ . Then  $AP_2$  verifies whether the MN is authorized user or not via handover authentication protocol (HAP). If the MN is an unauthorized user,  $AP_2$  will reject the MN's access request. If the MN is an authorized user, a session key will be established simultaneously for protecting data traffic between the MN and  $AP_2$ . Figure 1 illustrates a typical handover authentication scenario.



**Fig. 1.** A typical handover authentication scenario

In recent years, many HAPs based on identity-based cryptography have been presented [6–11, 13–16].

He et al. [6] proposed an HAP named PairHand from prime order bilinear pairings. To protect MN's privacy, they used a pool of shorter-lived pseudonyms. Unfortunately, PairHand is vulnerable to the private key compromised attack, where an adversary can recover any MN's private key. To withstand this attack, He et al. [7] proposed an improved PairHand by replacing the prime  $q$  order bilinear group with a composite  $n$  order bilinear group. However, Yeo et al. [9] showed that He et al.'s improved PairHand is still vulnerable to private key compromised attack. Even worse, an adversary can compute the master key when prime factors of  $n$  are all relatively small. Tsai et al. [10] presented a provably secure HAP from prime order bilinear pairings. Wang and Hu [11] proposed an improved HAP from prime order bilinear pairings. However, both Tsai et al.'s protocol [10] and Wang et al.'s protocol [11] can not achieve perfect forward secrecy.

Almost all the existing HAPs so far are constructed from bilinear pairings, which are regarded as a complex and expensive operation in modern

cryptography [12]. Recently, several pairing-free HAPs have been proposed by using elliptic curve cryptography. Cao et al. [13] proposed an identity-based HAP without bilinear pairing operation to decrease computation cost for a mobile device, and claimed that the scheme achieves user anonymity. However, Li et al. [14] showed Cao et al.'s protocol cannot achieve real user anonymity and Un-traceability, and proposed an efficient privacy-aware HAP. Unfortunately, Chaudhry et al. [15] and Xie et al. [16] showed that Li et al.'s protocol is vulnerable to access point impersonation attack and proposed two improved HAPs separately to overcome the security weakness of Li et al.'s protocol. However, both Chaudhry et al.'s protocol and Xie et al.'s protocol cannot provide key confirmation, key escrow-free and batch verification.

In summary, existing HAPs are either inefficient or insecure, or lack sufficient concern for user privacy. In this paper, we propose a design idea for HAP by using pseudonym-based cryptography and a variant of the SIGMA (Sign-and-MAC) approach [17], present a new HAP by adopting a lightweight identity-based signature scheme [18] and a new pseudonym-based private key issuing protocol, and show that our proposed HAP satisfies all basic desirable security and privacy requirements, including perfect forward secrecy, user anonymity and un-traceability, key confirmation and so on. Finally, we compare the performance of our proposed HAP with existing HAPs, and show that our proposed HAP is more efficient than existing HAPs in terms of computation and communication. Furthermore, batch verification for handover authentication is also achieved.

This paper is organized as follows. We introduce some necessary preliminary work in Sect. 2. We describe our proposed HAP in Sect. 3, and present security and efficiency analysis of our proposed HAP in Sect. 4. Finally, we conclude our work in Sect. 5.

## 2 Preliminaries

To facilitate further description, we introduce notations in Table 1.

### 2.1 Elliptic Curve and Complexity Assumptions

Let  $E_p(a, b)$  be a set of elliptic curve points over a large prime finite field  $\mathbb{F}_p$  ( $p > 3$ ), defined by the following non-singular elliptic curve equation:

$$y^2 = x^3 + ax + b \pmod{p}$$

with  $a, b \in \mathbb{F}_p$  and  $4a^3 + 27b^2 \pmod{p} \neq 0$ . The set of the solutions of the elliptic curve equation together with a point at infinity  $\mathcal{O}$  under the group operation of elliptic curve form an additive group, and  $\mathcal{O}$  serves as identity element for the group.

Let  $P \in E_p(a, b)$  be a point with a prime order  $q$  in the elliptic curve, and  $\mathbb{G}$  be a subgroup generated by the base point  $P$ , i.e.,  $\mathbb{G} \stackrel{\text{def}}{=} \langle P \rangle$ . We define  $[x]P = P + P + \dots + P$  ( $x$  times) as scalar multiplication.

**Table 1.** Notations

Symbol	Description
$\lambda$	Security parameter
$x \xleftarrow{\$} \mathbf{S}$	Pick an element $x$ uniformly at random from the set $\mathbf{S}$
$\ell_{\text{id}}$	The bit length of an identity or a pseudo-identity
$\ell_{\text{ts}}$	The bit length of a time-stamp
$\mathbf{ID}$	The identity space $\mathbf{ID} = \{0, 1\}^{\ell_{\text{id}}}$
$\mathbf{TS}$	The time-stamp space $\mathbf{TS} = \{0, 1\}^{\ell_{\text{ts}}}$
$H_1$	A secure hash function $H_1 : \{0, 1\}^{\ell_{\text{id}}} \times \mathbb{G} \rightarrow \mathbb{Z}_q^*$
$H_2$	A secure hash function $H_2 : \{0, 1\}^* \times \mathbb{G} \times \mathbb{Z}_q^* \rightarrow \mathbb{Z}_q^*$
$H_3$	A secure hash function $H_3 : \mathbb{Z}_q^* \times \mathbb{G} \rightarrow \{0, 1\}^{\ell_{\text{id}}}$
MAC	A secure message authentication code $\text{MAC} : \{0, 1\}^\lambda \times \{0, 1\}^{2\ell_{\text{id}}} \rightarrow \{0, 1\}^\lambda$
KDF	A secure session key derivation function $\text{KDF} : \mathbb{G} \rightarrow \{0, 1\}^\lambda$
$\oplus$	Exclusive-OR operation
$\parallel$	Concatenation operation

**Definition 1 (ECDLP).** Given  $Q \in \mathbb{G}$ , the elliptic curve discrete logarithm problem (ECDLP) is to find the integer  $x$ ,  $1 \leq x \leq q$ , such that  $Q = [x]P$ .

The advantage of an adversary  $\mathcal{A}$  in breaking the ECDLP is defined by

$$\text{Adv}_{\mathcal{A}}^{\text{ECDLP}}(1^\lambda) = \Pr[\mathcal{A}(P, Q = [x]P) = x \mid x \xleftarrow{\$} \mathbb{Z}_q^*].$$

We say that the elliptic curve discrete logarithm assumption (ECDLA) holds for the group  $\mathbb{G}$ , if for any probabilistic polynomial time (PPT) adversary  $\mathcal{A}$ , the above advantage is a negligible function in the security parameter  $\lambda$ .

**Definition 2 (ECCDHP).** Given  $(P, [a]P, [b]P) \in \mathbb{G}^{(3)}$  where  $a, b \xleftarrow{\$} \mathbb{Z}_q^*$ , the elliptic curve computational Diffie-Hellman problem (ECCDHP) for the group  $\mathbb{G}$  is to compute  $[ab]P$ .

The advantage of an adversary  $\mathcal{A}$  in breaking ECCDHP is defined by

$$\text{Adv}_{\mathcal{A}}^{\text{ECCDHP}}(1^\lambda) = \Pr[\mathcal{A}(P, [a]P, [b]P) = [ab]P \mid a, b \xleftarrow{\$} \mathbb{Z}_q^*].$$

We say that the elliptic curve computational Diffie-Hellman assumption (ECCDHA) holds for the group  $\mathbb{G}$ , if for any PPT adversary  $\mathcal{A}$ , the above advantage is a negligible function in the security parameter  $\lambda$ .

## 2.2 Galindo and Garcia Identity-Based Signature Scheme

Galindo and Garcia [18] proposed a lightweight identity-based signature scheme named GG-IBS in Africacrypt 2009. It is recognized as one of the most efficient

identity-based signature schemes till now, because no complicated bilinear pairings is required. The GG-IBS scheme is described as follows.

- **Setup:** The trusted private key generator (PKG) first generates an elliptic curve group  $\mathbb{G}$  of prime order  $q$  with a generator  $P$ . Then, the PKG chooses  $s \xleftarrow{\$} \mathbb{Z}_q^*$  and computes  $P_{\text{pub}} = [s]P$ . Finally, the PKG sets the master secret key  $msk = s$  and publishes the master public key  $mpk = \langle \mathbb{G}_1, q, P, P_{\text{pub}}, H_1, H_2 \rangle$ .
- **Extract:** A user submits a private key request with his/her identity information  $\text{id} \in \mathbf{ID}$  to the PKG. Upon receiving the request, the PKG chooses  $r_{\text{id}} \xleftarrow{\$} \mathbb{Z}_q^*$ , computes  $R_{\text{id}} = [r_{\text{id}}]P$ ,  $c = H_1(\text{id}, R_{\text{id}})$  and  $sk_{\text{id}} = r_{\text{id}} + cs \pmod q$ . Finally, the PKG sends  $(sk_{\text{id}}, R_{\text{id}})$  to the user via a secure channel. Upon receiving the response message, the user computes  $c = H_1(\text{id}, R_{\text{id}})$  and checks the following equation:

$$[sk_{\text{id}}]P \stackrel{?}{=} R_{\text{id}} + [c]P_{\text{pub}}$$

If it holds, the user stores  $R_{\text{id}}$  and sets  $sk_{\text{id}}$  as his/her identity-based signing private key. Anyone can get the user's public key from  $(\text{id}, R_{\text{id}})$  by computing

$$Q_{\text{id}} = R_{\text{id}} + H_1(\text{id}, R_{\text{id}})P_{\text{pub}}.$$

- **Sign:** To sign a message  $m \in \{0, 1\}^*$ , the signer with identity  $\text{id}$  and signing private key  $sk_{\text{id}}$  chooses  $a \xleftarrow{\$} \mathbb{Z}_q^*$ , computes  $c = H_1(\text{id}, R_{\text{id}})$ ,  $A = [a]P$ ,  $d = H_2(m, A, c)$  and  $b = a + sk_{\text{id}}d \pmod q$ . Finally, the signer sets  $\sigma = (b, R_{\text{id}}, A)$  as his/her signature on  $m$ .
- **Verify:** Given the signer's identity  $\text{id}$ , a pair of message  $m$  and signature  $\sigma = (b, R_{\text{id}}, A)$ , any verifier can first compute  $c = H_1(\text{id}, R_{\text{id}})$  and  $d = H_2(m, A, c)$ . Then, the verifier checks the following equation holds or not.

$$[b]P \stackrel{?}{=} A + [cd]P_{\text{pub}} + [d]R_{\text{id}}.$$

If it holds, the verifier accepts the signature and outputs true. Otherwise, outputs  $\perp$ .

The GG-IBS scheme is proved to be existential unforgeability under adaptively chosen identity and message attacks (EUF-ID-CMA) in the random oracle model under the ECDLA [19]. We will use the GG-IBS scheme in our HAP to provide mutual authentication between the roaming MN and the target AP.

### 3 Our Handover Authentication Protocol

Our proposed HAP also consists of four phases: system initialization phase, handover authentication phase, batch verification phase and DoS attack resistance phase. In order to defend against DoS attack, the method in [6] can be adopted in our proposed HAP. Therefore, we only briefly review the first three phases.

### 3.1 System Initialization

The AS first generates an elliptic curve subgroup  $\mathbb{G}$  of prime order  $q$  with a generator  $P$ . Then, the AS chooses  $s \xleftarrow{\$} \mathbb{Z}_q^*$  and computes  $P_{\text{pub}} = [s]P$ . Finally, the AS sets the master secret key  $msk = s$ , and publishes the master public key  $mpk = \langle \mathbb{G}, q, P, P_{\text{pub}}, H_1, H_2, H_3, \text{MAC}, \text{KDF} \rangle$ .

- When an AP, say  $j$ , registers to the AS, they perform the AP registration protocol illustrated in Fig. 2.
- When an MN, say  $i$ , registers to the AS with identity  $\text{id}_{\text{MN}_i} \in \mathbf{ID}$ , they perform the MN registration protocol illustrated in Fig. 3.

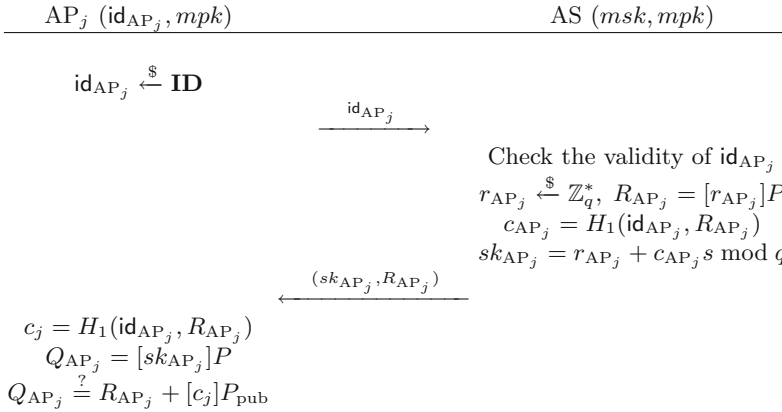


Fig. 2. AP registration protocol

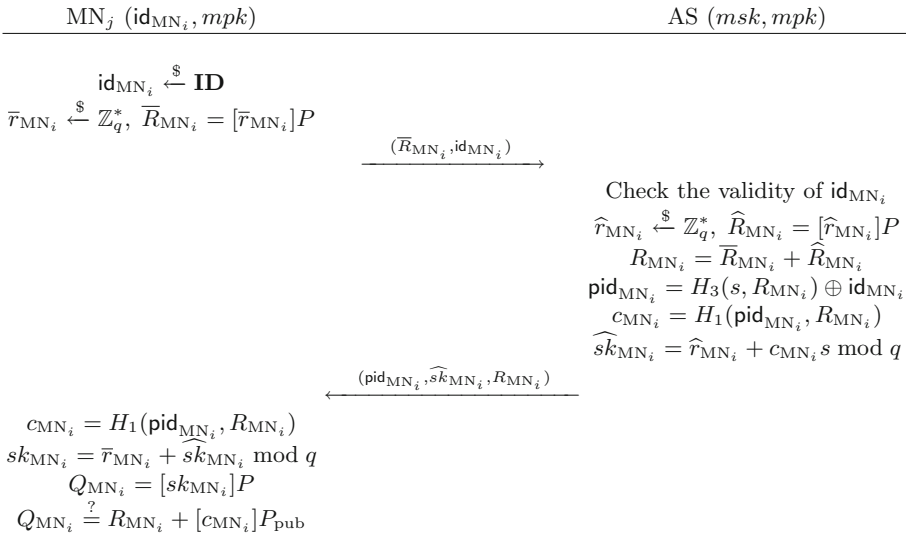


Fig. 3. MN registration protocol

### 3.2 Handover Authentication

When the MN  $i$  moves out of the coverage of current associated AP, it should handover to a new AP. Assume each AP  $j$  periodically broadcasts a beacon message, which includes the items  $\text{id}_{\text{AP}_j}$ ,  $R_{\text{AP}_j}$  and other regular information of the network. If the MN  $i$  chooses a target AP  $j$  and extracts  $\text{id}_{\text{AP}_j}$  and  $R_{\text{AP}_j}$  from the beacon message, then the MN  $i$  enters into the handover authentication phase. Figure 4 illustrates the steps of the handover authentication process.

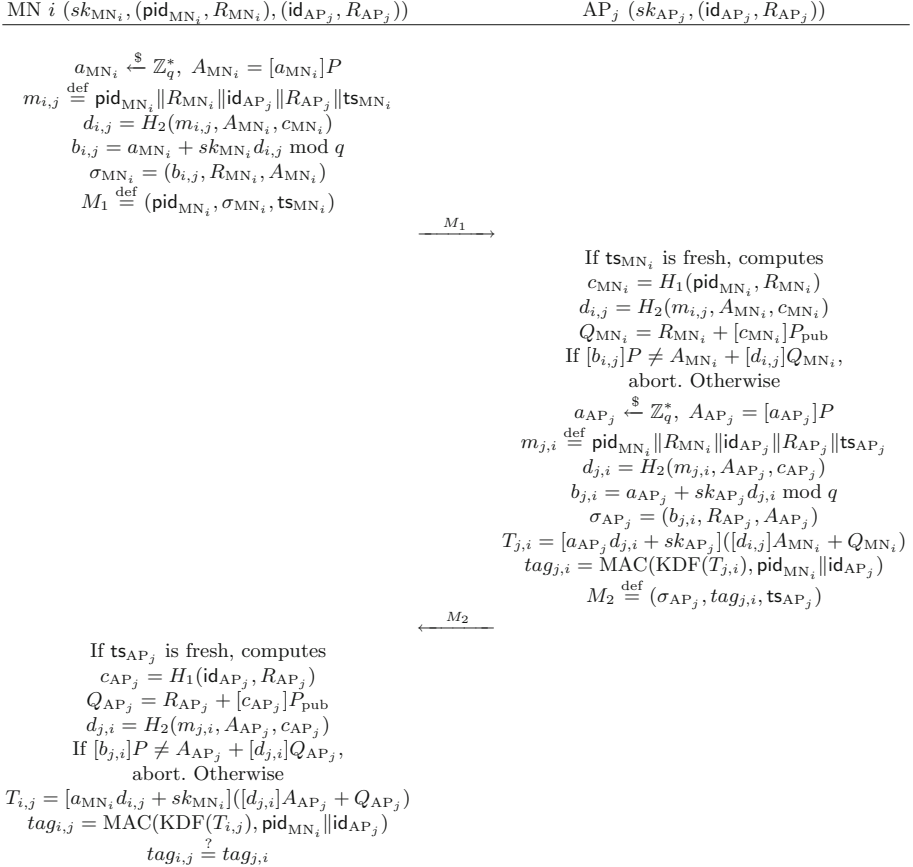


Fig. 4. Handover authentication phase

### 3.3 Batch Verification

Assume that an AP receives  $n$  distinct handover authentication request from  $n$  distinct MNs which denoted as

$$\langle \text{pid}_{\text{MN}_1}, \sigma_{\text{MN}_1}, \text{ts}_{\text{MN}_1} \rangle, \langle \text{pid}_{\text{MN}_2}, \sigma_{\text{MN}_2}, \text{ts}_{\text{MN}_2} \rangle, \dots, \langle \text{pid}_{\text{MN}_n}, \sigma_{\text{MN}_n}, \text{ts}_{\text{MN}_n} \rangle,$$

respectively. Instead of verifying each individual signature separately, the AP can verify these  $n$  signatures simultaneously by checking the following batch verification criterion.

$$\sum_{i=1}^n [b_{i,j}]P = - \sum_{i=1}^n [c_{MN_i} d_{i,j}]P_{\text{pub}} + \sum_{i=1}^n (A_{MN_i} + [d_{i,j}]R_{MN_i})$$

It is obvious that to verify these  $n$  signatures according to the batch verification criterion, the AP requires  $n + 2$  scalar multiplication over elliptic curve group  $\mathbb{G}$ . However, if the AP verifies each individual signature separately, it requires  $3n$  scalar multiplication over elliptic curve group  $\mathbb{G}$ .

## 4 Security and Efficiency Analysis

The proposed HAP has the following security and privacy properties.

- *Mutual authentication*: Our proposed protocol realizes the mutual authentication between the MN  $i$  and the AP  $j$ . In the handover authentication phase, the access request sent by MN  $i$  to AP  $j$  is actually a signature that generated by MN  $i$  with its signing private key on the message  $m_{i,j} = \text{pid}_{MN_i} \| R_{MN_i} \| b_{i,j} \| A_{MN_i} \| \text{ts}_{MN_i}$ , which is used to prove to AP  $j$  that MN  $i$  is the private key holder corresponding to the pseudonym  $\text{pid}_{MN_i}$ . Similarly, the authentication response sent by AP  $j$  includes a signature that generated by AP  $j$  with its signing private key on the message  $m_{j,i} = \text{pid}_{MN_i} \| R_{MN_i} \| \text{id}_{AP_j} \| R_{AP_j} \| \text{ts}_{AP_j}$ , which is used to prove to MN  $i$  that AP  $j$  is the private key holder corresponding to the identity  $\text{id}_{AP_j}$ .
- *Session key establishment*: If the roaming MN  $i$  and the target AP  $j$  successfully complete matching sessions, they are able to calculate the same session key as follows.

$$\begin{aligned} T_{j,i} &= [a_{AP_j} d_{j,i} + sk_{AP_j}]([d_{i,j}]A_{MN_i} + Q_{MN_i}) \\ &= [a_{AP_j} d_{j,i} + sk_{AP_j}][d_{i,j}a_{MN_i} + sk_{MN_i}]P \\ &= [d_{i,j}a_{MN_i} + sk_{MN_i}][a_{AP_j} d_{j,i} + sk_{AP_j}]P \\ &= [d_{i,j}a_{MN_i} + sk_{MN_i}][d_{j,i}]A_{AP_j} + Q_{AP_j} \\ &= T_{i,j} \\ &\Rightarrow K_{i,j} (= \text{KDF}(T_{i,j})) \\ &= K_{j,i} (= \text{KDF}(T_{j,i})) \end{aligned}$$

- *Key confirmation*: During the handover authentication phase, the roaming MN  $i$  calculates the message authentication code  $tag_{i,j}$ , and compares  $tag_{i,j}$  with the message authentication code  $tag_{j,i}$  received from the target AP  $j$ . If  $tag_{i,j} = tag_{j,i}$ , this implies that both MN  $i$  and AP  $j$  have computed the same shared session key. Thus, MN  $i$  can be confirmed that AP  $j$  has actually possesses the session key and a secure channel is established between AP  $j$

and MN  $i$  using the shared session key  $\text{KDF}(T_{i,j}) = \text{KDF}(T_{j,i})$ . After that, MN  $i$  sends a ciphertext that encrypted under the session key  $\text{KDF}(T_{i,j})$  to AP  $j$ . If AP  $j$  successfully decrypts the ciphertext with the session key  $\text{KDF}(T_{j,i})$ , this gives an assurance to AP  $j$  that MN  $i$  actually possesses the session key during real-time communication.

- *MN anonymity*: MN  $i$ 's pseudo identity  $\text{pid}_{\text{MN}_i} = H_3(s, R_{\text{MN}_i}) \oplus \text{id}_{\text{MN}_i}$  and  $R_{\text{MN}_i}$  are transmitted to AP  $j$  in the authentication request packet  $M_1 \stackrel{\text{def}}{=} (\text{pid}_{\text{MN}_i}, \sigma_{\text{MN}_i}, \text{ts}_{\text{MN}_i})$ . However, AP  $j$  cannot extract  $\text{id}_{\text{MN}_i}$  from  $M_1$  because AP  $j$  have no idea of the system master key  $s$ .
- *MN un-traceability*: Although the handover authentication request messages must include a pseudo identity of the roaming MN, however, there is no linkage between these pseudo identities. An adversary, including the malicious AP cannot link two sessions initiated by the same MN, because the roaming MN can constantly change its pseudo identity.
- *Conditional privacy preservation*: To guarantee the privacy for mobile node, the AS chooses a family of pseudo identities and generates associated private keys for each MN in our proposed protocol. Undoubtedly, the AS knows the relationship between a pseudo identity and the real identity. In fact, the AS can extract MN  $i$ 's real identity by computing  $\text{id}_{\text{MN}_i} = \text{pid}_{\text{MN}_i} \oplus H_3(s, R_{\text{MN}_i})$ . Thus, the AS can extract the real identity of the roaming MN through the intercepted messages in an emergency case.
- *Private key compromised security*: In the handover authentication phase, the access request sent by an MN  $i$  to an AP  $j$  is actually a signature that generated by the MN  $i$  with his/her signing private key on the message  $m_{i,j}$ , which is used to prove to the AP  $j$  that the MN  $i$  is the private key holder corresponding to the pseudonym  $\text{pid}_{\text{MN}_i}$ . Here, we use the GG-IBS scheme, one reason for this is its efficiency and simplicity, another more important reason is that it has been proved to be EUF-ID-CMA secure in the random oracle model under the ECDLA [19]. Even if an adversary gets polynomial messages and their corresponding signatures generated by the same MN  $i$ , he can not forge a valid signature of the MN  $i$ , let alone get the MN  $i$ 's private key.
- *No key control*: As the ephemeral parameters  $a_{\text{MN}_i}$  and  $a_{\text{AP}_j}$  are determined by the roaming MN  $i$  and the target AP  $j$ , respectively, no one can influence the result of the session keys, or enforce a session key to a preselected value.
- *Perfect forward secrecy*: The session key  $K_{i,j} = \text{KDF}(T_{i,j}) = \text{KDF}([d_{i,j}a_{\text{MN}_i} + sk_{\text{MN}_i}][[d_{j,i}]A_{\text{AP}_j} + Q_{\text{AP}_j}])$  calculated by MN  $i$  is equal to the session key  $K_{j,i} = \text{KDF}(T_{j,i}) = \text{KDF}([a_{\text{AP}_j}d_{j,i} + sk_{\text{AP}_j}][[d_{i,j}]A_{\text{MN}_i} + Q_{\text{MN}_i}])$  calculated by the AP  $j$ . According to the ECCDHA, there is no polynomial time adversary can compute the session key without MN  $i$ 's private key or AP  $j$ 's private key.
- *Known session key security*: Intuitively, in our proposed protocol, each session depends on two elements  $a_{\text{MN}_i}$  and  $a_{\text{AP}_j}$  that are uniformly chosen from  $\mathbb{Z}_q^*$  at random that makes the session key of each session computationally independent. Thus, from the knowledge of one session key nothing can be implied about the value of the other session keys.



We present security comparisons between our proposed HAP and existing HAPs [10,11,14–16] in Table 2. Only our proposed HAP can satisfy all security and privacy requirements.

**Table 2.** Security comparison with existing HAPs

	[10]	[11]	[14]	[15]	[16]	Ours
Mutual authentication	Yes	Yes	No	Yes	Yes	Yes
MN anonymity	Yes	Yes	Yes	Yes	Yes	Yes
MN un-traceability	Yes	Yes	Yes	Yes	Yes	Yes
Conditional privacy preservation	Yes	Yes	Yes	Yes	Yes	Yes
Private key compromised security	Yes	Yes	No	Yes	Yes	Yes
Known session key security	No	No	Yes	Yes	Yes	Yes
Perfect forward secrecy	No	No	Yes	Yes	Yes	Yes
Key confirmation	No	No	No	No	No	Yes
No key control	No	No	Yes	Yes	Yes	Yes
Key escrow-free	No	No	No	No	No	Yes

Next, we compare the performance of our proposed HAP with that of existing HAPs [10,11,14–16] in Table 3. For computational cost, we focus on the time spent on the high cost operations, such as the time ( $t_{bp}$ ) spent on executing a bilinear paring operation, the time ( $t_{mtp}$ ) spent on executing a map-to-point hash operation, and the time ( $t_{sm}$ ) spent on executing a scalar multiplications over the group  $\mathbb{G}_1$  or  $\mathbb{G}$ , while the time spent on highly efficient operations, such as the general hash function and key derivation function, is neglected. To evaluate the length of the messages transmitted in the protocol execution, we assume that the size of a timestamp, the size of the general hash function’s output, the length of the pseudo identity are 32 bits, 160 bits, and 128 bits, respectively.

Tsai et al.’s HAP [10] and Wang et al.’s HAP [11] are based on bilinear pairing. The group  $\mathbb{G}_1$  in the bilinear pairing  $\hat{e} : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_T$  is generated on

**Table 3.** Performance comparison with existing HAPs

	Computational cost for MN	Computational cost for AP	Communication cost (bits)	Batch verification	Bilinear pairings
[10]	$t_{bp} + t_{mtp} + 2t_{sm}$	$3t_{bp} + t_{mtp} + t_{sm}$	1728	Yes	Yes
[11]	$t_{bp} + t_{mtp} + 3t_{sm}$	$3t_{bp} + t_{mtp} + t_{sm}$	1728	Yes	Yes
[14]	$4t_{sm}$	$5t_{sm}$	1248	No	No
[15]	$4t_{sm}$	$6t_{sm}$	1088	No	No
[16]	$4t_{sm}$	$6t_{sm}$	1408	No	No
Ours	$6t_{sm}$	$6t_{sm}$	1312	Yes	No

the super singular elliptic curve  $y^2 = x^3 + x + 1 \pmod{p}$ . To provide 80-bit security,  $p$  is a 512-bit prime number and the order  $q$  of  $\mathbb{G}_1$  is a 160-bit prime number. Li et al.'s HAP [14], Chaudhry et al.'s HAP [15], Xie et al.'s HAP [16] and our proposed HAP are based on elliptic curve cryptography, group  $\mathbb{G}$  is generated on a non-singular elliptic curve  $y^2 = x^3 + ax + b \pmod{p}$  with  $4a^3 + 27b^2 \pmod{p} \neq 0$ . To achieve 80-bit security, both  $p$  and  $q$  are 160-bit prime numbers.

## 5 Conclusions

In this paper, we show none of existing handover authentication protocols employed for wireless networks meets the desirable security and privacy requirements. Then, we present a new handover authentication protocol based on a lightweight identity-based signature scheme and a pseudonym-based private key issuing protocol. Compared with existing handover authentication protocols, our proposed protocol is more efficient in terms of computation and communication.

**Acknowledgments.** This research is jointly funded by Science and Technology Program of Guangzhou (Grant No. 201707010358), and the Opening Project of Shanghai Key Laboratory of Integrated Administration Technologies for Information Security (Grant No. AGK201707).

## References

1. Park, S.H., Ganz, A., Ganz, Z.: Security protocol for IEEE 802.11 wireless local area network. *Mob. Netw. Appl.* **3**(3), 237–246 (1998)
2. Zekri, D., Defude, B., Delot, T.: Building, sharing and exploiting spatio-temporal aggregates in vehicular networks. *Mob. Inf. Syst.* **10**(3), 259–285 (2014)
3. Oliveira, L., et al.: Ubiquitous monitoring solution for wireless sensor networks with push notifications and end-to-end connectivity. *Mob. Inf. Syst.* **10**(1), 19–35 (2014)
4. Mohammad, G., et al.: A survey on wireless body area networks for eHealthcare systems in residential environments. *Sensors* **16**(6), 831 (2016). <https://doi.org/10.3390/s16060831>
5. He, D.J., et al.: A strong user authentication scheme with smart cards for wireless communications. *Comput. Commun.* **34**(3), 367–374 (2011)
6. He, D.J., et al.: Secure and efficient handover authentication based on bilinear pairing functions. *IEEE Trans. Wirel. Commun.* **11**(1), 48–53 (2012)
7. He, D.J., Chen, C., Chan, S., Bu, J.J.: Analysis and improvement of a secure and efficient handover authentication for wireless networks. *IEEE Commun. Lett.* **16**(8), 1270–1273 (2012)
8. He, D.J., Bu, J.J., Chan, S., Chen, C.: Handauth: efficient handover authentication with conditional privacy for wireless networks. *IEEE Trans. Comput.* **62**(3), 616–622 (2013)
9. Yeo, S.L., et al.: Comments on “analysis and improvement of a secure and efficient handover authentication based on bilinear pairing functions”. *IEEE Commun. Lett.* **17**(8), 1521–1523 (2013)

10. Tsai, J.L., Lo, N.W., Wu, T.C.: Secure handover authentication protocol based on bilinear pairings. *Wirel. Pers. Commun.* **73**(3), 1037–1047 (2013)
11. Wang, W.J., Hu, L.: A Secure and efficient handover authentication protocol for wireless networks. *Sensors* **14**(7), 11379–11394 (2014)
12. Cao, X., Kou, W.: A pairing-free identity-based authenticated key agreement protocol with minimal message exchanges. *Inf. Sci.* **180**(15), 2895–2903 (2011)
13. Cao, J., Ma, M., Li, H.: An uniform handover authentication between E-UTRAN and non-3GPP access networks. *IEEE Trans. Wirel. Commun.* **11**(10), 3644–3650 (2012)
14. Li, G.S., et al.: A new privacy-aware handover authentication scheme for wireless networks. *Wirel. Pers. Commun.* **80**(2), 581–589 (2015)
15. Chaudhry, S.A., Farash, M.S., Naqvi, H., et al.: A robust and efficient privacy aware handover authentication scheme for wireless networks. *Wirel. Pers. Commun.* **93**(2), 311–335 (2017)
16. Xie, Y., Wu, L.B., Kumar, N., Shen, J.: Analysis and improvement of a privacy-aware handover authentication scheme for wireless network. *Wirel. Pers. Commun.* **93**(2), 523–541 (2017)
17. Krawczyk, H.: SIGMA: the ‘SIGn-and-MAC’ approach to authenticated diffie-Hellman and its use in the IKE protocols. In: Boneh, D. (ed.) *CRYPTO 2003*. LNCS, vol. 2729, pp. 400–425. Springer, Heidelberg (2003). [https://doi.org/10.1007/978-3-540-45146-4\\_24](https://doi.org/10.1007/978-3-540-45146-4_24)
18. Galindo, D., Garcia, F.D.: A schnorr-like lightweight identity-based signature scheme. In: Preneel, B. (ed.) *AFRICACRYPT 2009*. LNCS, vol. 5580, pp. 135–148. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-02384-2\\_9](https://doi.org/10.1007/978-3-642-02384-2_9)
19. Chatterjee, S., Kamath, C., Kumar, V.: Galindo-Garcia identity-based signature revisited. In: Kwon, T., Lee, M.-K., Kwon, D. (eds.) *ICISC 2012*. LNCS, vol. 7839, pp. 456–471. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-37682-5\\_32](https://doi.org/10.1007/978-3-642-37682-5_32)
20. Schnorr, C.P.: Efficient signature generation by smart cards. *J. Cryptol.* **4**(3), 161–174 (1991)
21. Pointcheval, D., Stern, J.: Provably secure blind signature schemes. In: Kim, K., Matsumoto, T. (eds.) *ASIACRYPT 1996*. LNCS, vol. 1163, pp. 252–265. Springer, Heidelberg (1996). <https://doi.org/10.1007/BFb0034852>

# Spatial Outlier Information Hiding Algorithm Based on Complex Transformation

Zhaoyu Shou<sup>1</sup>, Akang Liu<sup>2(✉)</sup>, Simin Li<sup>1</sup>, and Xiawei Cheng<sup>2</sup>

<sup>1</sup> Key Laboratory of Cognitive Radio and Information Processing Ministry of Education, Guilin University of Electronic Technology, Guilin 541004, China

{guilinshou, siminli}@guet.edu.cn

<sup>2</sup> School of Information and Communication,

Guilin University of Electronic Technology, Guilin 541004, China

kangluu@foxmail.com, chengxiawei@gmail.com

**Abstract.** The anomaly data is easily disturbed by malicious third party in the information sharing or transmission process. To guarantee the safety and integrity of outlier information, a novel method of outlier information privacy preserving is proposed, namely, spatial outlier information hiding algorithm based on complex transformation. Firstly, the anomaly dataset is obtained by outlier detection algorithm. Then the two-dimensional feature data of anomaly objects is selected to construct the complex data and complex factors. Finally, the outlier information is hidden by complex transformation. The receiver receives the hidden dataset and the complex factor set, in which the hidden data can be effectively restored. The feasibility and validity of this algorithm are verified by simulation and contrast experiment.

**Keywords:** Data perturbation · Complex transformation · Outlier detection  
Outlier information hiding · Privacy preserving

## 1 Introduction

The anomaly data obtained by outlier detection algorithm is vulnerable to be disturbed and acquired by malicious third party in the process of information sharing or transmission. In order to guarantee the security and integrity of outlier information, the anomaly data must be effectively hidden. Data privacy preserving technology [1] are divided into two kinds: one kind is anonymity [2–5], namely deletion or encryption identifier (such as name, ID card number, social insurance code, etc.) which can directly identify personal identity. Because of the ability to prevent sensitive data from being leaked in the data publishing environment and to ensure the authenticity of the published data, the anonymity technique has received extensive attention in the practical application. Another kind is data perturbation [6–10], namely distortion, confusion and random change initial data. This method adds “noise” to initial data to prevent discovery of the real values. Since the data no longer reflect real-world values, they can’t be misused to violate privacy information.

Focusing on data perturbation technique to hide anomaly data, a spatial outlier information hiding algorithm based on complex transformation is proposed. It can

ensure data sender to send the least amount of data information and parameters. While the safety and integrity of anomaly data being guaranteed in the process of information sharing and transmission, the hidden data can be accurately restored in the receiver.

The remainder of this paper is organized as follows: Sect. 2 is the related work. Section 3 presents the basic concepts. Section 4 explains the proposed methodologies. Section 5 describes about experiments and analysis while providing the restoration of hidden data. The final section summarizes the whole paper.

## 2 Related Work

At present, data perturbation techniques are applied to these aspects of data clustering and privacy preserving. Oliveira et al. [6] proposes a method of privacy preserving clustering by data transformation, which translates the initial data by geometric transformations such as translation, scaling and simple rotation. This method is simple but easy to lead to data error, and the method also has some restrictions for the dimension of data. Liu and Xu [8] also adopt the geometric transformation with random response technology to achieve data perturbation. This method increases the randomness and unpredictability of data transformation, to achieve a good safety protection.

As one of data perturbation techniques, rotation transformation [7, 10, 12, 13] is usually used for privacy preserving. This technique is applied to global data, but there is no the restoration processing of data. Li and Zhang [7] find some important issues from an existing data mining technique such as balancing privacy and accuracy. The author proposes a hybrid data transformation method (HDT) which is built upon the application of Double-Reflecting Data Perturbation (DRDP) and Rotation based Translation (RBT), in order to provide secrecy of data confidential numerical attribute without losing accuracy. By using HDT, data owners can share their data with data miners to find accurate clusters without any concern about violating data privacy. However, this technique is aimed mainly to handle centralized data. At the same time, the restoration of data is infeasible. Ahmed and Rauf [10] propose a fuzzy hybrid data transformation method which is a combination of fuzzy data modification (FDM) with various membership function and random rotation perturbation method (RRP) to anonymize original data. The hybrid method achieves good privacy and utility. But this method works only for numerical dataset.

Data perturbation has high time complexity in dealing with global data. In this paper, the anomaly data is chosen for hidden processing, which can greatly reduce the computational complexity. The anomaly data itself is treated as sensitive information hidden by complex transformation.

Complex transformation proposed in this paper belongs to one of data perturbation techniques. The theory of complex transformation is as follows. As shown in Fig. 1, the complex data  $a + jb$  ( $j^2 = -1$ ), can be expressed in the form of rotation radius and angle. Assuming that there are two complex data:  $a + jb$  and  $c + jd$ , they can be transformed as:  $a + jb = r_1 * (\cos \alpha + j \sin \alpha)$ ,  $c + jd = r_2 * (\cos \beta + j \sin \beta)$ .

$$\begin{aligned}
 (a + jb)(c + jd) &= r_1 * (\cos \alpha + j \sin \alpha) * r_2 * (\cos \beta + j \sin \beta) \\
 &= r_1 * r_2 * (\cos \alpha \cos \beta - \sin \alpha \sin \beta + j(\cos \alpha \sin \beta + \sin \alpha \cos \beta)) \\
 &= r_1 * r_2 * (\cos(\alpha + \beta) + j \sin(\alpha + \beta))
 \end{aligned}$$

Therefore, the multiplication of two complex data, is equal to the rotation radius being multiplied, and rotation angle being added.

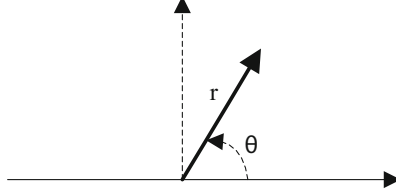


Fig. 1. The form of complex data

### 3 Formal Definitions

In order to better describe this algorithm, some definitions are defined below.

**Definition 1 (Local Density).** The symbol  $\rho_i$  [15] represents the local density of the  $i$ -th object. The calculation is shown as Formula (1):

$$\rho_i = \sum_{p \in N_{k\text{-distance}(i)}(i) \cap p \neq i} e^{-\left(\frac{d(i,p)}{k\text{-distance}(i)}\right)^2} \quad (1)$$

The  $k$ -distance of an object is calculated by standard Euclidean distance. For any positive integer  $k$ , the  $k$ -distance of object  $i$ , denoted as  $k\text{-distance}(i)$ , is defined as the distance  $d(i, o) (i \in D)$  [16] between  $i$  and an object  $o$ , such that:

- (i) for at least  $k$  objects  $o' \in D \setminus \{i\}$  it holds that  $d(i, o') \leq d(i, o)$ ,
- (ii) for at most  $k-1$  objects  $o' \in D \setminus \{i\}$  it holds that  $d(i, o') < d(i, o)$ .

As shown in Fig. 2, the local density of an object  $i$  is related to the number of objects in its  $k$  distance neighbor, while ignoring those objects far away from  $i$ . From Formula (1), as the data objects far away from  $i$  show a tiny effect on the local density of  $i$ .

**Definition 2 (Distance Dissimilarity).** Distance dissimilarity indicates the dissimilarity between data objects. The symbol  $\delta_i$  represents the distance dissimilarity of the  $i$ -th object. The calculation is shown as Formula (2):

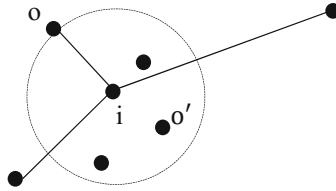


Fig. 2. Local density,  $k = 4$

$$\delta_i = \begin{cases} \min_{q_j}(d(q_i, q_j)), & i \geq 2, j < i \\ \max(\delta_{q_j}), & i = 1, j \geq 2 \end{cases} \quad (2)$$

Assuming that  $\{q_i\}_{i=1}^N$  is the descending order of subscript of  $\{\rho_i\}_{i=1}^N$ , sort  $\rho_i$  to  $\rho_{q_1} \geq \rho_{q_2} \geq \rho_{q_3} \geq \dots \geq \rho_{q_N}$ . From Formula (2), providing that the data object  $o$  has the highest local density, the dissimilarity is the largest  $d_{max}(o, o'), o' \in D$ . On the contrary, the smallest dissimilarity is  $d_{min}(o, o'), o' \in D$  in all data objects whose local density is greater than  $o$ .

As shown in Fig. 3, (a) contains 28 data objects, and (b) shows the result about the local density and distance dissimilarity of each data object after calculating with Definitions (1) and (2). From Fig. 3(a), the marks 16 and 17 can be judged as anomaly objects. According to Fig. 3(b), the local density of the data objects is small, and their distance dissimilarity is large.

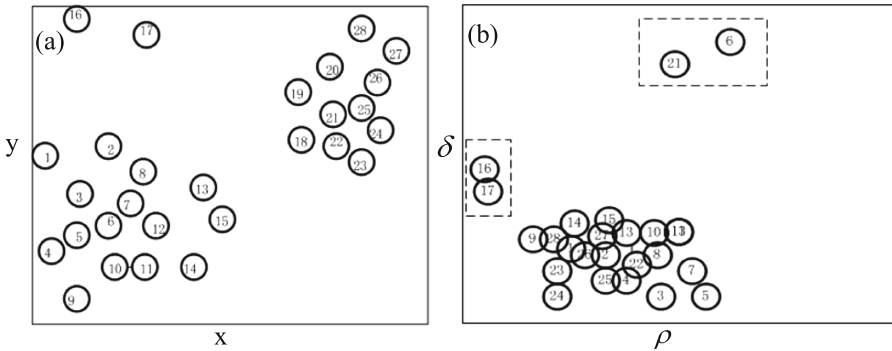


Fig. 3. (a) 28 data objects; (b) local density and distance dissimilarity

**Definition 3 (Outlier Degree Coefficient, ODC).** The outlier degree coefficient is determined by the local density and distance dissimilarity of the data object. The symbol  $ODC(i)$  represents the outlier degree coefficient of the  $i$ -th object ( $i \in D$ ). The calculation is shown as Formula (3):

$$ODC(i) = e^{-\left(\frac{\delta_i}{\rho_i}\right)} \quad (3)$$

ODC(i) is normalized to the range of [0, 1]. When  $\rho_i$  is smaller and  $\delta_i$  is bigger, the coefficient of ODC(i) is smaller, indicating the greater degree of outlier.

**Definition 4 (Complex Factor).** *The complex factor is constructed by the local density and distance dissimilarity. The calculation is shown as Formula (4):*

$$C(i) = \rho_i + je^{-\delta_i} (j^2 = -1) \quad (4)$$

From Formula (4), when  $\delta_i$  is bigger, the factor of  $e^{-\delta_i}$  is smaller. Within the range of (0, 1], the complex modulus of C(i) is very small. What's more, the factor can be as one of scale transformation for the original dataset to achieve complex transformation accordingly.

**Definition 5 (Complex Transformation).** *The complex transformation is expressed as:  $T = D \times C$ .  $D = [A_p, A_q]^T$  represents for a  $2 \times n$  data matrix about the  $p$ -th and  $q$ -th feature objects of original datasets, in which  $2 \times n$  is the number of a data object in feature objects, and  $m$  is the number of feature objects ( $1 \leq p, q \leq m$ ).  $A_p, A_q$  respectively corresponds to a  $1 \times n$  vector of the  $p$ -th and  $q$ -th feature objects.  $T$  stands for a  $2 \times n$  hidden data matrix followed by the complex transformation.  $C$  represents the corresponding complex factor.*

**Definition 6 (Inverse Transformation Factor).** *Data restoration method is inseparable from the process of complex transformation. Inverse transformation is to restore the hidden data from complex transformation. Inverse transformation is expressed as:  $D' = T \times C'$ , and the calculation of inverse transformation factor is shown in Formula (5):*

$$C'(i) = \frac{\rho_i - je^{-\delta_i}}{\rho_i^2 + e^{-2\delta_i}} (j^2 = -1) \quad (5)$$

Local density  $\rho_i$  and distance dissimilarity  $\delta_i$  need one-to-one correspondence with the outlier numbered information.  $(\rho_i^2 + e^{-2\delta_i})$  is the square of complex modulus. The premise of this structure is that the following holds:

$$D' = T \times C'(i) = D \times C(i) \times C'(i) = D \times (\rho_i + je^{-\delta_i}) \times \frac{\rho_i - je^{-\delta_i}}{\rho_i^2 + e^{-2\delta_i}} = D$$

The hidden data information can be restored accurately.



## 4 Spatial Outlier Information Hiding Algorithm Based on Complex Transformation (SOIH-CT)

A spatial outlier information hiding algorithm based on complex transformation is proposed in this paper. The anomaly data is detected to better hide the outlier information. Specifically, given a metric space  $(D, m)$  ( $D$  is a dataset and  $m$  represents the number of dimensions), each object  $x \in D$  is assigned a degree of outlier. Then the candidates selected by the top  $N$  of outlier degree (Top- $N$ ) are used to determine the final anomaly objects. And the complex data and corresponding complex factors are constructed by the selected anomaly objects. Thus the process of complex transformation is operated to hide outlier information.

The selected Top- $N$  is based on the overall distribution of ODC value. The anomaly data objects should contain some normal data objects, to achieve a higher level of outlier information hiding security.

### 4.1 ODBMCLD Outlier Detection

The spatial outlier information hiding algorithm is based on ODBMCLD outlier detection [15, 17]. Firstly, the original spatial data is processed by the data sender with calculating the distance measure between two data objects. Then, the  $\rho_i$ ,  $\delta_i$  and ODC are calculated by the proposed formula for each data object. Finally, the candidates selected by the Top- $N$  are used to determine the final anomaly objects.

#### Algorithm1: ODBMCLD

**Input:**  $D$ , the set of data objects

$k$ ,  $k$ -distance

Top- $N$

begin:

1: calculate the Euclidean distance between two data object

2: according to Eq. (1), calculate  $\rho_i$  of every data object

3: sort  $\rho_i$  to  $\rho_{q_1} \geq \rho_{q_2} \geq \rho_{q_3} \geq \dots \geq \rho_{q_N}$

4: according to Eq.(2), calculate  $\delta_i$

5: according to Eq.(3), calculate ODC(i) , ascending order of ODC(i)

6: select Top- $N$  objects as the final anomaly objects

**Output:** anomaly objects dataset

### 4.2 SOIH-CT Algorithm

The data objects for spatial outlier detection are multidimensional. Based on ODBMCLD, the complex transformation hiding process is performed in the data sender. Firstly, the two-dimensional feature data of anomaly objects are selected to construct the complex data and corresponding complex factor. Then, the process of complex transformation is operated to hide outlier information. Finally, the hidden

dataset and complex factor set are saved after the complex transformation. The two datasets can be independently sent to ensure the safety of data information. The receiver can receive all the data information to restore the anomaly data.

This algorithm also involves the construction of complex data for spatial outlier data objects. The concept of the complex itself is defined in the two-dimensional data. Now this problem is extended to the multidimensional spatial dataset based on complex transformation.  $m$  represents the number of multidimensional attributes. According to the following rules:

- (i) when  $m$  is even,  $k' = m/2$
- (ii) when  $m$  is odd,  $k' = (m + 1)/2$ .

The characteristics are transformed into  $k'$  groups, to compose  $k'$  subsets containing two-dimensional feature data. Then, these subsets are implemented separately by the complex transformation hiding process. Finally the hidden dataset is obtained by merging.

The hidden dataset is formed by complex transformation hiding process, which can be used in public transmission. Complex factors set includes the anomaly objects number, local density and distance dissimilarity, directly sent to the receiver.

#### **Algorithm2: SOIH-CT**

**Input:** anomaly objects dataset

begin:

1: if  $m$  is even, then  $k' = m/2$

    else  $k' = (m + 1)/2$

2: select  $k'$  pairs of attributes randomly  $P_{k'}: (A_i, A_j), (1 \leq i, j \leq m \text{ and } i \neq j)$

3: for each paired attribute  $P_{k'}$  do

4: select the two-dimensional feature data of anomaly objects to construct corresponding complex data

5: according to Eq.5, construct the complex factor  $C(i)$

6: perform the complex transformation

**Output:** the hidden dataset and complex factor set

This algorithm involves two important parameters:  $\rho_i$  and  $\delta_i$ . According to the characteristics of outlier detection itself, the ODC is constructed by these two parameters. Also, the complex factor and inverse transformation factor are related to this parameter in the subsequent process.

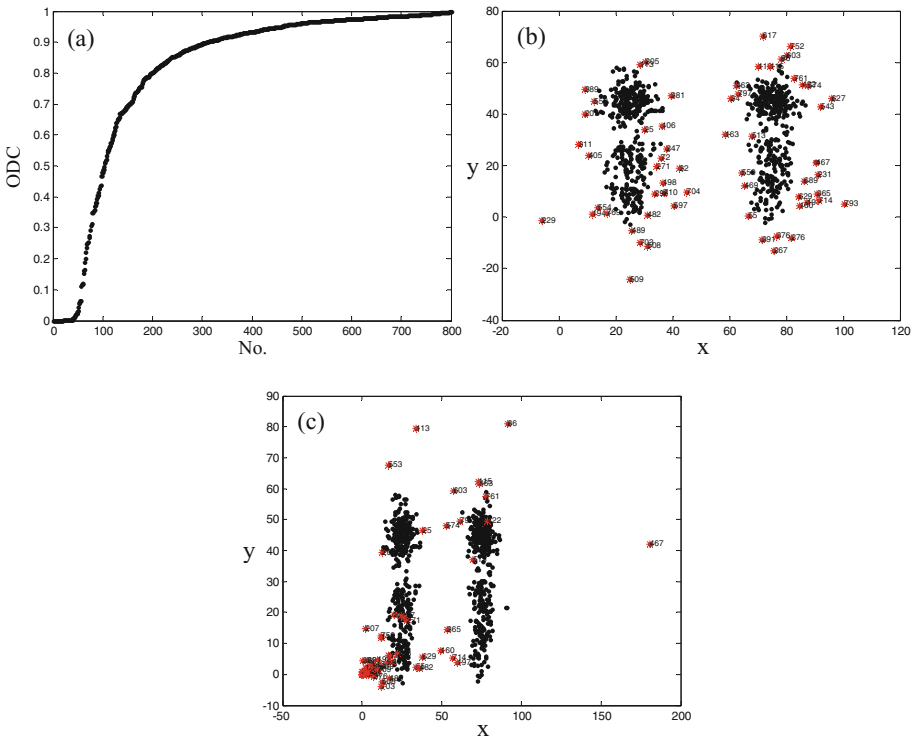
This algorithm is designed to guarantee the safety and integrity of outlier information. The computational complexity is not required, but it is obvious that the computational complexity of SOIH-CT is higher than that of ODBMCLD. ODBMCLD outperforms other existing approaches [15], due to the consideration with both  $\rho_i$  and  $\delta_i$  in the stage of choosing candidate outliers and deciding the cluster to which every object belongs.

## 5 Experiment and Analysis

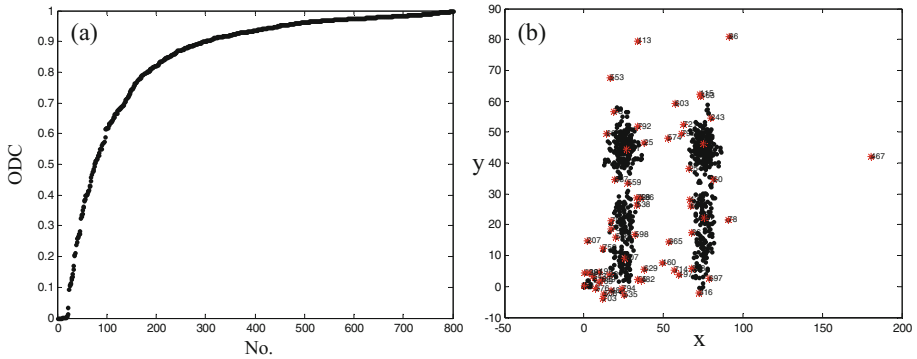
In this section, experiments are conducted to verify the effectiveness of the proposed algorithm on both synthetic and real datasets. All the experiments are performed in MATLAB R2014a on Intel Pentium-B960, 2.2 GHz, with 4 GB memory running on Windows7 64 bit. In particular, the first part validates the availability of the SOIH-CT on synthetic datasets. The second part provides the restoration of hidden data. The third part verifies the outperformance of SOIH-CT on real datasets.

### 5.1 Experiments on Synthetic Datasets

In this section, to intuitively show the performance of the proposed algorithm, the two-dimensional dataset is generated. Some experiments to verify the validity of SOIH-CT algorithm are conducted. This synthetic dataset contains 800 data objects, a random dataset that contains 4 clusters in a normal distribution. The results of SOIH-CT are shown in Figs. 4 and 5.



**Fig. 4.** Diagram of SOIH-CT: (a)  $ODC(i)$  in ODBMCLD, (b) final anomaly objects (Top-N = 60), (c) the hidden dataset



**Fig. 5.** Diagram of ODBMCLD in the hidden dataset: (a)  $ODC(i)$ , (b) final anomaly objects (Top- $N = 60$ )

Figure 4 shows the results of simulation by Algorithm 1 and the results of the hidden dataset after SOIH-CT. In Fig. 4(c), outlier information still are as those in (b). Compared with Fig. 4(b) and (c), it can be shown that some of the detected anomaly objects are hidden in the normal dataset, and the other outlier information has also changed, so that some implied correlation between the data has changed. Compared with Figs. 4 and 5, we can see that the ODBMCLD method is used to detect the hidden dataset. The anomaly results are obviously different. Although there is still some anomaly data to be detected, the outlier information has been changed greatly. And the sorting and the value of ODC have changed greatly. Thus, this SOIH-CT algorithm is effective and feasible, to achieve an expected effect.

## 5.2 The Restoration of Hidden Data

After the hidden processing performed, anomaly data needs to be restored accurately in the receiver. In particular, this algorithm relates to the spatial outlier information method based on complex transformation, so as to achieve the integrity process with the security of outlier information and the availability of data restoration.

The receiver receives the hidden dataset and a complex factor set. According to the anomaly numbered information in the complex factor dataset, the two-dimensional feature data is selected to construct the complex data. The selected two-dimensional outlier information comes from not only the hidden dataset but also the complex factor set. The inverse transformation factor is then constructed by the proposed formula. Local density and distance dissimilarity need one-to-one correspondence with the outlier numbered information. Namely, the hidden anomaly data can be restored accurately.

**Algorithm3: the restoration of hidden data**

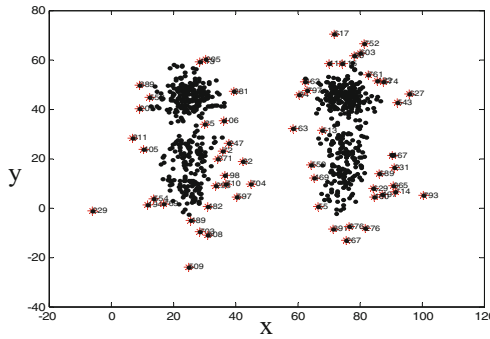
**Input:** the hidden dataset,  
 complex factor set

begin:

- 1: According to the anomaly numbered information in the complex factor dataset, select the two-dimensional feature data
- 2: construct local complex data and corresponding inverse transformation factor  $C'(i)$
- 3: as described Definition (6), perform inverse transformation to restore anomaly data

**Output:** the restored dataset

As shown in Fig. 6, the hidden synthetic data is restored accurately.



**Fig. 6.** Diagram of hidden data restoration

**5.3 Experiments on Real Datasets**

The effectiveness of the proposed algorithm is verified by experiments on synthetic random datasets. In this section, contrast experiments are conducted on real datasets taken from UCI machine learning repository [18] to further validate the priority over other algorithms. Seven real datasets are chosen to carry out performance comparison and analysis, with different scales and dimensions. The brief information of chosen datasets is described in Table 1.

*A. Detection rate and Repetition rate* Detection rate [15] is defined as the ratio between the number of outliers detected by the system to the total number of outliers presented in the dataset, i.e.  $DR = (\text{No. of detected outlier}) / (\text{No. of outliers})$ . The greater the value is, the more the number of outliers is detected.

In view of the outlier objects detected in the original dataset, whether these objects are detected as anomaly objects after the hidden processing is measured by Repetition rate. It is shown as  $RR = (\text{No. of outlier again}) / (\text{No. of detected outlier})$ . And its value is smaller, explaining that the outlier information hiding is better.

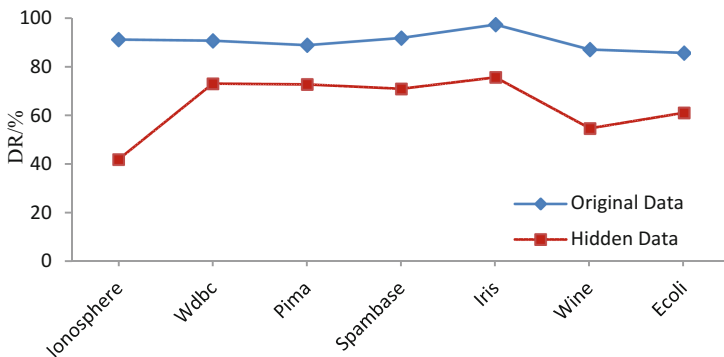
**Table 1.** Properties of datasets

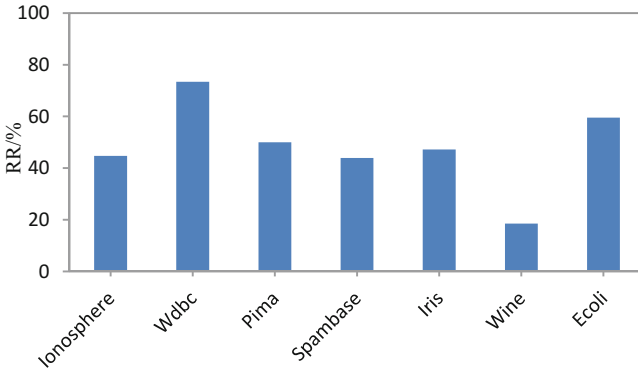
Dataset	No. of records	No. of attributes	No. of outlier
Ionosphere	351	34	126
Wdbc	569	30	212
Pima	768	8	268
Spambase	4601	57	1813
Iris	150	4	37
Wine	178	13	31
Ecoli	336	7	49

**Table 2.** Contrast results of the original data and hidden data

Dataset	No. of outlier	DR of original data/%	No. of outlier for hidden data	DR of hidden data/%	RR/%
Ionosphere	114	91.2	53	42.1	44.7
Wdbc	192	90.7	155	73.1	73.4
Pima	238	88.9	195	72.8	50.0
Spambase	1664	91.8	1287	70.99	43.9
Iris	36	97.3	28	75.7	47.2
Wine	27	87.1	17	54.8	18.5
Ecoli	42	85.7	30	61.2	59.5

As Fig. 7 shows, the detection rate is greatly reduced in results of outlier detection for the original data and hidden data. This indicates that the probability that the hidden data is detected as outliers will be greatly reduced. Except for the Wdbc in Fig. 8, the repetition rate of other datasets is low, demonstrating that the data hiding processing is more effective.

**Fig. 7.** DR results of the original data and hidden data



**Fig. 8.** RR results

*B. Degree of privacy* Traditionally, the privacy provided by a perturbation technique has been measured by the variance between the original and the perturbed values [7]. This measure is given by  $\text{Var}(X - Y)$  where  $X$  represents a single original attribute and  $Y$  the distorted attribute. Privacy level can be specified by the metric as  $\text{Sec} = \text{Var}(X - Y)/\text{Var}(X)$ , the higher  $\text{Sec}$  shows the higher protection level. In [7], Iris, Wine and Ecoli are selected for experimental evaluation, so only these datasets are selected in privacy evaluation. Table 3 shows the degree of privacy provided by these methods with RBT, HDT [7] and SOIH-CT.

**Table 3.** Sec for hidden datasets

Dataset	RBT	HDT	SOIH-CT
Iris	1.54	1.57	2.17
Wine	1.45	1.63	2.95
Ecoli	1.26	1.41	4.18

From Fig. 9, in terms of privacy, we can see that SOIH-CT is better than RBT and HDT, which indicates that SOIH-CT is more secure and can preserve privacy better by applying RBT and HDT.

*C. Hiding Failure* Hiding failure [12, 19] is the portion of sensitive information that is not hidden by the application of a privacy preservation technique. The percentage of sensitive information that is still discovered after the data has been sanitized gives an estimate of the hiding failure parameter. Most of the developed privacy preserving algorithms are designed with the goal of obtaining zero hiding failure.

In the SOIH-CT algorithm, all of the anomaly objects can be disturbed. The value of every anomaly object in the dataset is modified. Although the repetition rate exists in Table 2, all of outlier information has changed, and some normal data objects are introduced to interfere. Hence, the hidden failure rate for the proposed technique is 0.

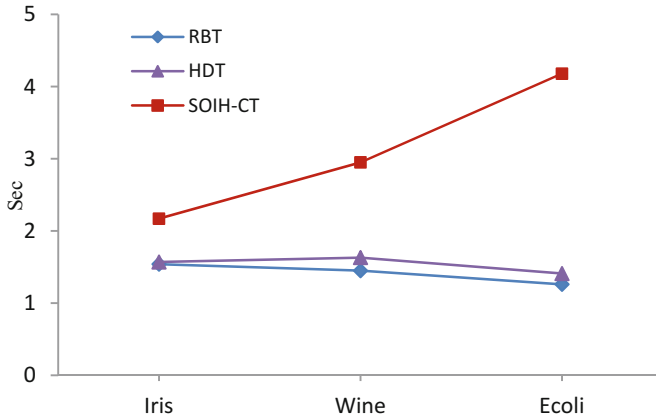


Fig. 9. Sec of RBT, HDT and SOIH-CT

#### 5.4 Adversary Attack on Data Privacy

Existing work [19] on analysis of privacy preserving perturbation techniques primarily considers an adversary that correlates publicly available data to reconstruct sensitive information. The worst case scenario would be an attacker that has prior information about the perturbation technique as well as some of the original data objects. The attacker could then proceed to reconstruct the entire original dataset. The effectiveness of a perturbation technique, therefore, would be the ability of the perturbed dataset to withstand such an attack.

The proposed technique considers the selected Top-N and  $k'$  subsets containing two-dimensional feature data. Different combinations of Top-N and  $k'$  subsets result in a wide various types of outlier information protection. Furthermore, the formula of complex factor is constructed specifically by the local density and distance dissimilarity for spatial outlier data objects. Complex factor set can be independently sent to the receiver to ensure the safety of data information.

From the adversary point of view, it would almost be impossible to use brute force to reverse this technique. An attacker wouldn't break the anomaly data by repeating this algorithm several times on the same data and learning from the weakness of this algorithm.

## 6 Conclusions

The spatial outlier information hiding algorithm proposed in this paper, is an innovation method of outlier information privacy preserving. Not only can this algorithm guarantee the safety and information integrity of anomaly objects in the process of information sharing and transmission, but also the hidden data can be restored accurately.



The proposed algorithm doesn't need to judge the deviation degree of every point. By the Top-N, computational cost is cut down. The complex transformation is performed by the characteristic of Top-N, to simplify the outlier detection processing of the whole algorithm. It also reduces the amount of data processing, and ensure data sender to send the least amount of data information and parameters. In the future, we will further explore other algorithms based on complex transformation to hide the outlier information. And we also research other efficient outlier detection algorithms for outlier information privacy preserving.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61662013, 61362021), Natural Science Foundation of Guangxi province (2016GXNSFAA380149), the Key Laboratory of Cognitive Radio and Information Processing, Ministry of Education (2011KF11), Key Lab of Trusted Software (kx201511), Innovation Project of GUET Graduate Education (2016YJ CXB02, 2017YJ CX34).

## References

1. Zhang, G.R.: Privacy data preserving method based on fuzzy discretization. In: Eighth International Conference on Fuzzy Systems and Knowledge Discovery, vol. 2, pp. 1201–1205. IEEE (2011)
2. Shinde, A., et al.: Privacy prevention of sensitive rules and values using perturbation technique. In: 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 577–581 (2016)
3. Ghate, R.B., Ingle, R.: Clustering based anonymization for privacy preservation. In: International Conference on Pervasive Computing, pp. 1–3. IEEE (2015)
4. Rajalakshmi, V., Mala, G.S.A.: Data anonymization using augmented rotation of sub-clusters for privacy preservation in data mining. In: 2013 Fifth International Conference on Advanced Computing (ICoAC), pp. 22–26. IEEE (2013)
5. Rajalakshmi, V., Mala, G.S.A.: Anonymization by data relocation using sub-clustering for privacy preserving data mining. *Indian J. Sci. Technol.* **7**(7), 975–980 (2014)
6. Oliveira, S.R.M., Zaane, O.R., Agropecuaria, E.I.: Privacy preserving clustering by data transformation. In: Proceedings of Brazilian Symposium on Databases, vol. 1, pp. 37–52 (2003)
7. Li, L., Zhang, Q.: A privacy preserving clustering technique using hybrid data transformation method. In: IEEE International Conference on Grey Systems and Intelligent Services, pp. 1502–1506. IEEE (2009)
8. Liu, J., Xu, Y.: Privacy preserving clustering by random response method of geometric transformation. In: Fourth International Conference on Internet Computing for Science and Engineering, pp. 181–188. IEEE (2010)
9. Gokulnath, C., Priyan, M.K., Balan, E.V., Prabha, K.P.R., Jeyanthi, R.: Preservation of privacy in data mining by using PCA based perturbation technique. In: International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials, pp. 202–206. IEEE Computer Society (2015)
10. Ahmed, K.A., Rauf, H.A.: Privacy preserving data using fuzzy hybrid data transformation technique. *Indian J. Sci. Technol.* **10**(24) (2017)
11. Vashkevich, A.V., Zhukov, V.G.: Privacy-preserving clustering using C-means. In: International Siberian Conference on Control and Communications, pp. 1–4. IEEE (2015)

12. Dhiraj, S.S.S., Khan, A.M.A., Khan, W., Challagalla, A.: Privacy preservation in k-means clustering by cluster rotation. In: TENCON 2009 - 2009 IEEE Region 10 Conference, pp. 1–7. IEEE (2009)
13. Lin, Z., Wang, J., Liu, L., Zhang, J.: Generalized random rotation perturbation for vertically partitioned data sets. In: IEEE Symposium on Computational Intelligence and Data Mining, pp. 159–162. IEEE (2009)
14. Leontiadis, I., Molva, R., Chorley, M.J., Colombo, G.B.: Privacy preserving similarity detection for data analysis. In: International Conference on Cloud & Green Computing, vol. 10, pp. 547–552. IEEE Computer Society (2013)
15. Shou, Z.-Y., Li, M.-Y., Li, S.-M.: Outlier detection based on multi-dimensional clustering and local density. *J. Cent. South Univ.* **24**(6), 1299–1306 (2017)
16. Breunig, M.M., Kriegel, H.P., Ng, R.T., Sander, J.: LOF: identifying density-based local outliers, vol. 29(2), pp. 93–104 (2000)
17. Rodriguez, A., Laio, A.: Clustering by fast search and find of density peaks. *Science* **344** (6191), 1492–1496 (2014)
18. Lichman, M.: UCI Machine Learning Repository. University of California, School of Information and Computer Science, Irvine, CA. <http://archive.ics.uci.edu/ml>
19. Challagalla, A., Dhiraj, S.S.S., Somayajulu, D.V.L.N., Mathew, T.S., Tiwari, S., Ahmad, S. S.: Privacy preserving outlier detection using hierarchical clustering methods. In: Computer Software and Applications Conference Workshops, pp. 152–157. IEEE (2010)
20. Zimek, A., Campello, R.J.G.B.: Data perturbation for outlier detection ensembles. In: International Conference on Scientific and Statistical Database Management, pp. 1–12. ACM (2014)

# A Reputation Model Considering Repurchase Behavior and Mechanism Design to Promote Repurchase

Yuan Liu<sup>(✉)</sup>, Jin Bai, Guibing Guo, Xingwei Wang<sup>(✉)</sup>, and Zhenhua Tan

Northeastern University, Shenyang, China  
liuyuan@swc.neu.edu.cn, wangxw@mail.neu.edu.cn

**Abstract.** In electronic commerce environment, reputation systems have been widely investigated for several decades towards building secure online market platforms. In this paper, we propose a new reputation model considering the repurchase behavior of buying agents (buyers). The buyer repurchase behavior is described by three factors: recency, frequency, and monetary. Since the repurchase behavior is essential for the survival of vendors in e-commerce in the long run, we further design a price premium based mechanism to encourage customers to conduct repeat transactions with their satisfactory selling agents (sellers). Theoretical analysis and simulation based experiments are conducted to evaluate the proposed system. The results show that there exists a unique pure strategy Nash equilibrium where buyers always repurchase from satisfactory sellers and sellers behave honestly.

**Keywords:** Reputation model · Repurchase · Repurchase strategy  
Dishonesty behavior · Electronic marketplace

## 1 Introduction

An electronic marketplace provides opportunities for distributed buying agents (buyers or customers) and selling agents (sellers or vendors) to exchange products and services through the Internet based platforms, such as, [Amazon](#) and [Taobao.com](#) [17]. E-marketplaces have been widely recognized to bear the information asymmetry problem [11] where buyers and sellers hold asymmetric information about their common transactions, which results in ‘Moral Hazard’ phenomena in economics [6]. Reputation systems have been widely employed to mitigate this problem. A reputation system collects ratings from buyers about the performance of respective sellers in their past transactions, and aggregates the ratings as the seller reputation score which is provided for other buyers in their subsequent transactions so as to choose satisfactory transaction partners (sellers) [15]. In this way, the buyers are able to know more information about the sellers before conducting transactions via the reputation systems and the information asymmetry problem can be effectively mitigated.

Meanwhile, from the perspective of a seller, the buyers are generally classified into two types: potential buyers and return buyers, depending on whether the buyers ever make transactions with the seller. The buyers in different types possess different amount of information and also use different strategies in making purchase decisions. The profit created by a return buyer is almost 4 times over a potential buyer, and the cost of maintaining a return buyer to purchase is much lower than attracting a potential buyer to purchase [12]. Thus, for the survival and long-term profitability of sellers, it is essential to attract buyers to return [2]. There are plenty of studies on investigating the repurchase intention through the Internet, based on expectation confirmation theory, the theory of acceptance model etc. The repurchase behavior of a buyer is actually not only a reflection of the buyer's intention in conducting repeat transactions, but also a cogent evidence that the buyer trusts the respective seller and feels relatively safe and reliable in their continuous transactions. Thus, the ratings provided by return buyers should be considered differently. However, the existing reputation systems rarely consider this factor in constructing seller reputation. This paper is the first attempt to formally consider repurchase behavior by proposing a novel reputation model, where the capability of buyers in choosing reputable sellers is also considered. Furthermore, based on the proposed reputation model, we also design a price premium based mechanism to promote repeat sales.

Therefore this work contributes the literature in the following three aspects:

1. A novel reputation model is proposed considering repurchase behavior, where the capability of buyers in choosing reputable sellers is also modeled.
2. Based on the proposed reputation model, we design a mechanism to promote repurchase behavior. In this way, the e-marketplace becomes a positive feedback closed-loop safe e-commerce system.
3. In the proposed system, there exists a unique pure strategy Nash equilibrium where buyers repeatedly conduct transactions with satisfactory sellers and sellers behave honestly.

## 2 Related Work

### 2.1 Reputation Computational Models

The reputation systems have been widely implemented in e-commerce applications, where the reputation score of an entity refers to the beliefs or opinions that are generally held by the reviewers regarding the history behavior of the particular entity. The reputation scores can be calculated in various ways, which are basically classified into three main types of computational models in the literature. The first type is Dempster-Shafer evidence theory based models, such as [4, 5, 7, 19–22]. The distribution of the collected ratings is interpreted as a basic probability assignment over a frame of discernment, and the Dempster-Shafer theory provides rules to combining and propagate the ratings, resulting in the reputation score of respective entities. These models have been investigated in vast p2p applications such as web service selection, file-sharing network, etc.

The second type is probability theory based reputation computational models, and various classical distribution models have been applied, such as normal distribution [1], Beta distribution [3, 10, 23], Dirichlet distribution [16], Bayesian theory [18], and Markov chain theory [13]. The reputation score is interpreted as the probability of an entity to behave properly.

The third type is direct aggregation or weighted aggregate the collective ratings, which is easily understandable and also a mostly utilized computational model in many e-commerce platforms such as eBay, Amazon, Taobao.com, JD.com. Due to the popularity and practicality of this model, we choose this type of reputation computational model and propose our reputation model considering repurchase behavior of buyers.

## 2.2 Repurchase Study in E-Commerce

In the domain of marketing research, buyer or customer repurchase intention and behavior is an important issue for both traditional markets and online marketplaces, since the return buyers are valuable and essential for the survival of vendors in the long run [2]. It has been shown that the satisfaction ratings are positively related with the repurchase behavior [14]. Thus, the repurchase behavior should also impact the rating aggregation in some way, which inspires us to propose a novel reputation model considering repurchase behavior in this paper.

In order to understand and quantify buyer repurchase behavior, RFM (recency, frequency and monetary) analysis, as a simple and powerful method, is popularly applied [8, 9]. It is based on the assumption in customer relation management (CRM) that buyers who have purchased more recently, more frequently, or give more monetary value to a seller, are much more likely to become a profitable buyer in the future. In this work, we borrow the RFM idea to quantitatively model buyer repurchase behavior.

## 3 The Proposed Reputation Model

In the system, the buyers investigate available sellers and make decisions in purchasing. After each transaction between a buyer and a seller, the buyer has a chance to provide a rating in the range  $[0, 1]$ . A high rating from a buyer indicates that the buyer feels satisfied during the transaction, and a low rating reflects dissatisfaction experience. The ratings provided by buyers are aiming to calculate the reputation of sellers so as to qualify the honesty of the seller in providing satisfactory products or services. The credibility of a buyer is then constructed based on the reputation values of the interacted sellers. A high credibility value of a buyer indicates that the buyer is capable of choosing reputable transaction partners. Furthermore, the system is designed with a price premium based mechanism to encourage buyers to repurchase from sellers. The main procedure of the proposed system is shown as follows.

---

**Procedure 1.** The Procedure of the proposed reputation system with A Mechanism to Promote Repurchase

---

**Input:** Seller  $s_i$ , Buyer  $b_j$ , Regular Price  $p_i$

**Output:** Final Price  $p_i^j$ ,  $R_i$ ,  $C_j$

1.  $b_j$  are shown with the product details with price  $p_i$ ;
  2. Both  $s_i$  and  $b_j$  are agree to conduct a transaction and process to payment step;
  3. Calculate final  $p_i^j$  by deducting the price premium;
  4. The system allow  $s_i$  charge  $b_i$  for  $p_i^j$ ;
  5. Product delivery;
  6.  $b_j$  provides rating  $r_{ij}$ ;
  7.  $s_i$ 's reputation is updated;
  8.  $b_j$ 's credibility is updated;
- 

**3.1 Repurchase Behavior Model**

The repeat purchase behavior of a buyer  $b_j$  in purchasing products from a seller  $s_i$  is described by the following three factors: *Recency*  $E_{ij}$  which is the time since buyer  $b_j$ 's last purchase with seller  $s_i$ ; *Frequency*  $F_{ij}$  which is the average number of purchases happening in a predefined period  $T$  between  $s_i$  and  $b_j$ ; *Monetary*  $M_{ij}$  which is the total money buyer  $b_j$  ever paid to  $s_i$  where average cost of the product is denoted as  $\bar{c}$ . For a buyer who never conduct transactions with a seller,  $E_{ij}$  is initially set to be an infinite large number, and  $F_{ij} = 0$ ,  $M_{ij} = 0$ .

In order to convert the three factors in the same range, each of the values  $E_{ij}$ ,  $F_{ij}$ , and  $M_{ij}$  is assigned with a score in the set  $\{1, 2, 3, 4, 5\}$ , as specified in Table 1. The scores are denoted as  $e_{ij}$ ,  $f_{ij}$ , and  $m_{ij}$ , respectively.

**Table 1.** Score assignment for  $E_{ij}$ ,  $F_{ij}$ ,  $M_{ij}$  factors

Scores	$E_{ij}$	$F_{ij}$	$M_{ij}$
5	$E_{ij} \leq \frac{T}{4}$	$F_{ij} \geq 4$	$M_{ij} \geq 4\bar{c}$
4	$\frac{T}{4} \leq E_{ij} < \frac{T}{2}$	$3 \leq F_{ij} < 4$	$3\bar{c} \leq M_{ij} < 4\bar{c}$
3	$\frac{T}{2} \leq E_{ij} < T$	$2 \leq F_{ij} < 3$	$2\bar{c} \leq M_{ij} < 3\bar{c}$
2	$T \leq E_{ij} < 2T$	$1 \leq F_{ij} < 2$	$\bar{c} \leq M_{ij} < 2\bar{c}$
1	$E_{ij} > 2T$	$F_{ij} < 1$	$M_{ij} < \bar{c}$

The repurchase behavior is then characterized by the parameter  $W_{ij}$  which is calculated based on the scores of the above three factors, by weighted averaging the scores  $e_{ij}$ ,  $f_{ij}$ , and  $m_{ij}$ , as shown in Eq. (1).

$$W_{ij} = \alpha e_{ij} + \beta f_{ij} + \gamma m_{ij} \tag{1}$$

where  $\alpha, \beta, \gamma \geq 0$  and  $\alpha + \beta + \gamma = 1$ . A high value of  $W_{ij}$  indicates that the buyer  $b_j$  and seller  $s_i$  view each other as valuable partner. For a pair of buyer and seller

who never conduct transactions, the value of  $W_{ij}$  should be its minimum value 1. For a pair of long-term partners who very frequently conduct transactions with high monetary value, the value of  $W_{ij}$  should be its maximum value 5.

The settings of  $\alpha$ ,  $\beta$ , and  $\gamma$  are adaptable for products or services in different categories. For example, with respect to the products like clothes with relative short lifetime, a valuable buyer should frequently purchase products with high prices, as a result the weights  $\beta$  and  $\gamma$  should be greater than  $\alpha$ .

### 3.2 Seller Reputation Model

Regarding a seller  $s_i$ , the rating provided by a buyer  $b_j$  is denoted as  $r_{ij} \in [0, 1]$ , and the reputation of the seller is denoted as  $R_i$ .

$$R_i = \frac{\sum_{b_k \in B_i} W_{ik} \times r_{ik}}{\sum_{b_k \in B_i} W_{ik}} \quad (2)$$

where  $B_i$  is the set of buyers who ever conduct transactions with seller  $s_i$ , and  $W_{ik}$  is the result of Eq. (1). Therefore  $R_i \in [0, 1]$ .

In this reputation computation model, the ratings of buyers are treated differently according to the value of the buyers for the respective seller. In other words, for high valuable buyers, their ratings should be considered more, and vice versa.

### 3.3 Buyer Credibility Model

The credibility of a buyer  $b_j$  reflects the intention of the buyer in conducting transactions with reputable sellers, which is denoted as  $C_j$ . It is further calculated as in Eq. (3).

$$C_j = \frac{\sum_{s_k \in S_j} W_{kj} \times R_{k-j}}{\sum_{s_k \in S_j} W_{kj}} \quad (3)$$

where  $S_j$  is the set of sellers who ever provide products or service for buyer  $b_j$ , and  $R_{k-j}$  is the seller reputation excluding the ratings from the buyer  $b_j$ . Therefore  $C_j \in [0, 1]$ .

According to Eq. (3), a buyer with high credibility indicates that the buyer has the capability to choose reputable sellers in their transactions. Thus, the credible buyers should be valued by reputable sellers, which inspires us to design a mechanism by offering credibility proportional price-premium to return buyers.

## 4 A Mechanism to Promote Repurchase

Suppose the regular price of a product or service provided by seller  $s_i$  is  $p_i$  which satisfies  $p_i > c$  to guarantee the feasibility of the marketplace. Thus, in a transaction between seller  $s_i$  and buyer  $b_j$  for product whose regular price is  $p_i$ , the price premium is defined as

$$\delta_i^j = C_j(p_i - c) \quad (4)$$

where  $c$  is the estimated cost of producing the product or providing the service. The value  $c$  is not necessary real cost, but it is required that  $c$  is not less than real cost. Therefore, the final price for the buyer  $b_j$  is obtained as follows.

$$p_i^j = p_i - \delta_i^j \tag{5}$$

It is valuable to emphasize that this price premium based mechanism can only be triggered when buyers repurchase from sellers. According to Eq. (4), for buyers who always conduct transactions with high reputation sellers, they are offered with high price premium when they return to purchase from the sellers again.

## 5 Analysis

In this section, we analyze the basic properties and the equilibrium of the proposed system.

### 5.1 Basic Properties

As a system which can sustain in the long run, it should attract the users (both buyer and sellers in the discussed scenario) are willing to participate with non-negative profit. This property is also called *individual rationality*.

**Proposition 1.** *In the proposed system, both buyers and sellers are individual rational (IR) in the transactions.*

*Proof.* In order to proof that buyers and sellers are individual rational, we need to show that their utility is always nonnegative. As the buyers and sellers have agreed to conduct transaction before the price premium is exposed. Thus, before applying the price premium, the utility of a seller  $s_i$   $U_i = p_i - c$  should greater or equal to 0, and the utility of a buyer  $b_j$   $U_j = v_j^i - p_i$  is also greater or equal to 0 where  $v_j^i$  is the valuation of  $b_j$  regarding the product provided by  $s_j$ .

The price premium is  $\delta_i^j = C_j(p_i - c)$ , which makes the utility of the buyer increases by  $\delta_i^j$  and meanwhile makes the seller utility decreases by  $\delta_i^j$ . Since the buyer utility  $U_j \geq 0$ ,  $U_j + \delta_i^j$  is naturally nonnegative. For the seller  $s_i$ , the utility becomes  $U_i - \delta_i^j = (1 - C_j)U_i$  which is nonnegative given that  $C_j \in [0, 1]$ .

The second property we are going to analyze is whether the reputable sellers in the proposed system are more profitable than dishonest sellers. This property guarantee that the sellers are individually motivated to improve their honesty, which is a necessary condition for a system to grow as a healthy and long term ecosystem.

**Proposition 2.** *Buyers have the incentive to repurchase from reputable sellers, and the sellers have the incentive to improve their reputation to attract buyers repurchasing.*



*Proof.* Based on Eqs. (3) and (4), we can obtain that

$$\frac{d\delta_i^j}{dR_{i-j}} = (p_i - c) \frac{W_{ij}}{\sum_{s_k \in S_j} W_{kj}} > 0 \quad (6)$$

given that  $p_i > c$ .

Meanwhile, a reputable seller has a high reputation which results in high price premium attractive for the buyers who intend to repurchase. The repurchase behavior brings the seller additional profit. Therefore, the sellers are willing to achieve the repurchase profit through sustaining high reputation.

The third property is whether the proposed system is robust against untruthful information input. Since the whole system is constructed based on the ratings provided by the buyers, it is essential to disappoint the buyers in providing untruthful ratings.

**Proposition 3.** *The buyers have no incentives to provide untruthful ratings.*

*Proof.* According to Eq. (3), a buyer's credibility is closely related with the reputation of the sellers with whom the buyer ever conducted transactions. However, the reputation value is calculated by excluding the ratings of the respective buyer. In other words, it is impossible to manipulate a buyer's credibility by providing untruthful ratings by the buyer self. Therefore, the buyers have no incentives to provide untruthful ratings in the proposed system.

## 5.2 Equilibrium Analysis

For buyers, we consider three possible purchase strategies:

1. Strategy 1 [*PriceFirst*]: select sellers offering with the lowest price;
2. Strategy 2 [*ReputationFirst*]: select sellers with the highest reputation;
3. Strategy 3 [*SuccessRepurchase*]: repeatedly select a seller having satisfactory experience until the seller behavior dishonestly, and then change to another seller.

Regarding sellers, for the reason of simplicity, we consider two strategies:

1. Strategy 1 [*BehaveHonestly*]: provide truthful product details with acceptable price and deliver products truthfully;
2. Strategy 2 [*BehaveDishonestly*]: provide misleading product details with low price and deliver unsatisfactory products.

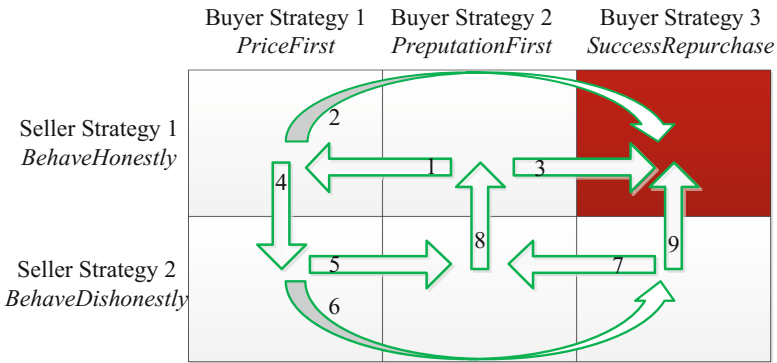
**Proposition 4.** *There exists a unique pure strategy Nash equilibrium (NE) where buyers take SuccessRepurchase strategy and sellers take BehaveHonestly strategy.*

*Proof.* The interaction between a pair of buyer and seller is a one-shot two-player game, where the buyer strategy set is  $\{PriceFirst, ReputationFirst, SuccessRepurchase\}$  and the seller strategy set is  $\{BehaveHonestly, BehaveDishonestly\}$ , as shown in Fig. 1. This proof is based on the game theoretic concept of best-response dynamics, where each player updates his strategy periodically by reacting optimally to other players' strategies. A fixed point of the system of all players' best-response is a Nash equilibrium of the game.

First of all, we set the initial strategy of a buyer is *ReputationFirst*, given a seller takes *BehaveHonestly* strategy, the buyer has two better options.

1. The buyer can either switch his strategy to *PriceFirst* so as to achieve higher utility by choosing a seller offering lower price (as shown in arrow 1 in Fig. 1). The buyer then will further switch his strategy to *SuccessRepurchase* so as to achieve the price premium from the seller offering low price (arrow 2).
2. The buyer can also update his strategy to *SuccessRepurchase* so as to achieve consistently satisfactory transactions with price premium (arrow 3).

After several rounds of dynamics, the buyer will stabilize his strategy at *SuccessRepurchase* strategy, and the seller has no better choose by taking *BehaveHonestly* strategy.



**Fig. 1.** The demonstration of best response dynamics in the game between buyers and sellers.

When the buyer takes *PriceFirst* strategy, the seller will prefer to switch his strategy to *BehaveDishonestly* so as to offer a even lower price to attract the buyer (arrow 4). Then the buyer will change his strategy to *ReputationFirst* so as to avoid conducting transactions with the dishonest seller (arrow 5), or *SuccessRepurchase* to enjoy the price premium (arrow 6) and then switch to *ReputationFirst* to avoid the seller (arrow 7). The seller then will prefer to change his strategy to so as to improve his reputation and be chosen by the buyer (arrow 8). Thus the buyer and seller come back their initial strategy, and they will further stabilize at strategy *SuccessRepurchase* and *BehaveHonestly* respectively.

When the buyer takes *SuccessRepurchase* strategy, the seller takes *BehaveDishonestly*. The seller will prefer to change his strategy to *BehaveHonestly* so as to be chosen continually (arrow 9). Then both the buyer and the seller will not change their strategy any more.

In a word, based on best-response dynamics analysis, the two players will stable their strategy at  $\{SuccessRepurchase, BehaveHonestly\}$ , making it the unique pure strategy Nash equilibrium.

## 6 Experimental Evaluations

To evaluate the proposed reputation model and price premium based mechanism, a set of simulation based experiments are conducted.

In the experiments, there are 100 sellers and 200 buyers conducting transactions for 100 days. The honesty of sellers has ten levels  $\{0.1, 0.2, \dots, 1\}$  and each level contains 10 sellers. When a seller's honesty is 0.1, the seller will satisfy the buyers at probability 10%. The cost in producing the product is set to be 6, and the valuation of buyers is no less than 10. The price offered by sellers is in proportional to their honesty following the pricing rule in realistic marketplaces, i.e. the price of seller with honesty 0.1 is  $6 + 0.1 \times (10 - 6)$ . The set of buyers also are divided into three types according to their strategy in selecting sellers. There are 1/3 buyers take *PriceFirst* strategy where the probability of a seller being chosen as transaction partner is in proportional to the price the seller offers. There are another 1/3 buyers take *ReputationFirst* strategy where the probability of a seller being chosen is in proportional to the reputation the seller achieves. The last 1/3 buyers take *SuccessRepurchase* strategy where they repeatedly conduct transactions with a satisfactory seller until the seller behaves dishonestly then change to another seller.

In each day, each buyer needs to choose a seller from whom the buyer purchase a product. The chosen seller delivers the product according to his honesty. After the end of the day, the buyer provides a rating for each transaction.

### 6.1 Validation Results

The results in validating the proposed system are shown in Figs. 2, 3, 4, 5 and 6. After 100 simulation days, the reputation of sellers is presented in Fig. 2. We can observe that the modeled reputation in the proposed system well reflects the honesty of sellers. According to Fig. 3, the buyers taking strategy *SuccessRepurchase* can achieve significantly higher credibility value than buyers taking strategy *ReputationFirst* whose credibility value is higher than those taking strategy *PriceFirst*. This results shows that the buyers taking strategy *SuccessRepurchase* strategy has strong capability to select honest sellers. The transaction prices of buyers taking the three strategies are shown in Fig. 4. We can observe that the buyers taking strategy *SuccessRepurchase* can always achieve lower price than the buyers taking any of the other two strategies. According

to Fig. 5, the buyers taking strategy *SuccessRepurchase* achieve higher satisfactory transaction rate than the buyers taking the other two strategies. Comparing Figs. 3 and 5, the credibility value of buyers closely reflects the satisfactory transaction rate. It is because the credibility computation model aggregates the reputation of their history transaction partners, and high credibility indicates that they always choose sellers with high reputation. Finally, according to Fig. 6, the sellers with high honesty will attract significantly more buyers taking strategy *SuccessRepurchase* than those taking the other two strategies. Thus, in the proposed system, it is beneficial for buyers who take strategy *SuccessRepurchase* and sellers who behave honestly, contributing an stable equilibrium as analyzed in Section SectionNEAnalysis.

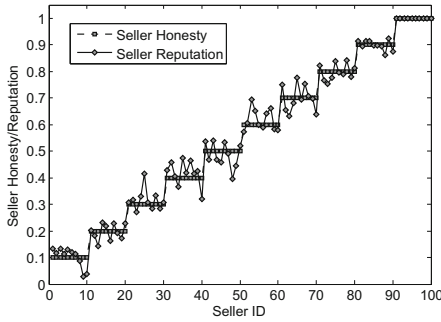


Fig. 2. Seller reputation vs Seller honesty

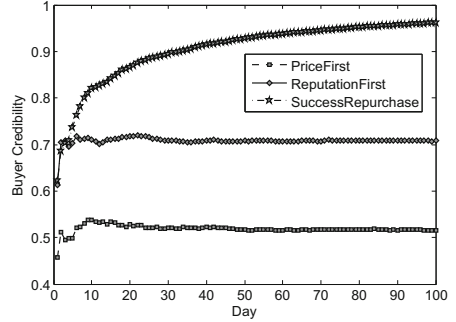


Fig. 3. Buyer credibility

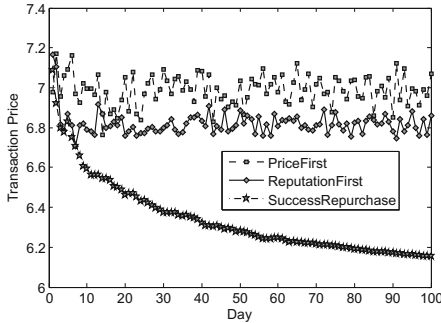


Fig. 4. The transaction price of buyers

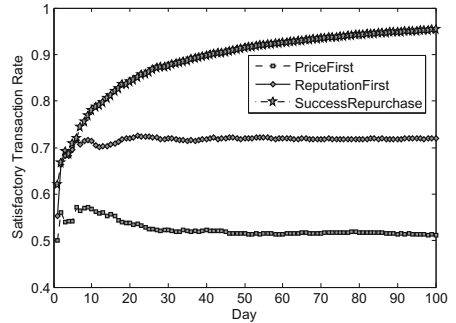
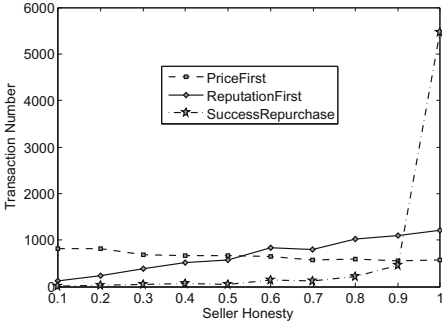


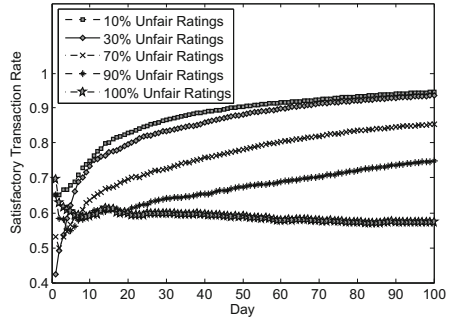
Fig. 5. Satisfactory transaction rate of buyers

### 6.2 Robustness Against Unfair Ratings

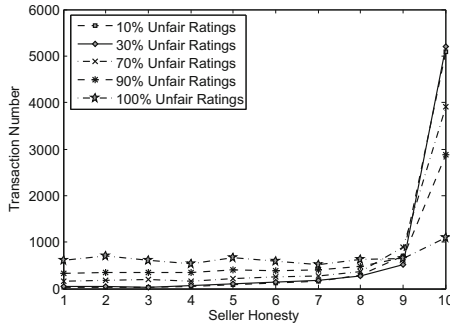
We also conduct experiments to evaluate the performance of the proposed system when the irrational buyers provide unfair (untruthful) ratings. We run the



**Fig. 6.** Transaction numbers of sellers



**Fig. 7.** Satisfactory transaction rate of buyers taking strategy *SuccessRepurchase* having unfair ratings



**Fig. 8.** Transaction numbers of sellers with buyers taking strategy *SuccessRepurchase* having unfair ratings

same set of experiments as Sect. 6.1 when the unfair ratings takes 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100%. We observe that the buyers taking strategy *PriceFirst* and *ReputationFirst* show reverse trend when the unfair ratings increase to 50%, where the buyers conduct more transactions with sellers with lower honesty. It means that the effectiveness of two strategies still rely on the majority truthful ratings. Due to the space limitation, the results are not presented in this paper. When the buyers take strategy *SuccessRepurchase*, their satisfactory transaction rate and the transaction numbers of sellers with different honesty levels are presented as in Figs. 7 and 8. We can observe that the buyers' satisfactory transaction rate still increases with time (days), however the increase rate grows slowly as the unfair ratings increases. Even in the worst case where all the ratings are unfair, the satisfactory transaction rate is not sacrificed too much. The reason behind this phenomenon is that the buyers make more mistakes before finding an honest seller when there are more unfair ratings. From Fig. 8, the sellers who behave honestly will always conduct more transactions with the buyers than the seller behave dishonestly to a certain extent. It

indicates that the sellers behaving fully honestly will attract more buyers in the proposed system, making the strategy of *BehaveHonestly* always dominant to *BehaveDishonestly*.

## 7 Conclusions

In this paper, we have proposed a novel reputation computation model of sellers considering buyer repurchase behavior, and construct the credibility of buyers. The credibility of buyers reflects the capability of buyers in selecting reputable sellers. Furthermore, to encourage buyers to repurchase, a price premium based mechanism is proposed, where the credible buyers are offered with the price premium when they purchase again from sellers. The theoretical analysis shows that there exists a unique pure strategy Nash equilibrium where buyers repurchase from satisfactory sellers and sellers behave honestly. The experimental results confirm the theoretical analysis and also show that the Nash equilibrium in the proposed system can fully robust against unfair ratings.

**Acknowledgments.** This research is partially supported by National Natural Science Foundation of China under Grant Nos. 61572123, 61402097 and 61602102; the National Science Foundation for Distinguished Young Scholars of China under No. 71325002; the Natural Science Foundation of Liaoning Province of China under Grant Nos. 20170540319 and 201602261; and the Fundamental Research Funds for the Central Universities under Grant Nos. N162410002, N161704001, N151708005, N161704004, N151704002.

## References

1. Abdel-Hafez, A., Xu, Y., Jøsang, A.: A normal-distribution based reputation model. In: Eckert, C., Katsikas, S.K., Pernul, G. (eds.) TrustBus 2014. LNCS, vol. 8647, pp. 144–155. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-09770-1\\_13](https://doi.org/10.1007/978-3-319-09770-1_13)
2. Atchariyachanvanich, K., Okada, H., Sonehara, N.: What keeps online customers repurchasing through the internet? ACM SIGecom Exchanges **6**(2), 47–57 (2007)
3. Bidgoly, A.J., Ladani, B.T.: Quantitative verification of beta reputation system using prism probabilistic model checker. In: Proceedings of International ISC Conference on Information Security and Cryptology, pp. 1–6 (2013)
4. Chen, X., Ming, F.U., Wang, X.Q.: Reputation evaluation model in grid-supported based on d-s evidence theory. J. Chin. Comput. Syst. **29**(3), 411–416 (2008)
5. Cvrček, D., Matyáš Jr., V., Patel, A.: Evidence processing and privacy issues in evidence-based reputation systems. Comput. Stand. Interfaces **27**(5), 533–545 (2005)
6. Dowd, K.: Moral hazard and the financial crisis. Cato J. **29**(1), 141–166 (2009)
7. Feng, R., Che, S., Wang, X., Yu, N.: Trust management scheme based on d-s evidence theory for wireless sensor networks. Int. J. Distrib. Sens. Netw. **2013**(4), 130–142 (2013)
8. Gupta, S., Lehmann, D.R.: Models of customer value. In: Wierenga, B. (ed.) Handbook of Marketing Decision Models. ISOR, vol. 121, pp. 255–290. Springer, New York (2007). [https://doi.org/10.1007/978-0-387-78213-3\\_8](https://doi.org/10.1007/978-0-387-78213-3_8)

9. Jašek, P.: Analyzing user activity based on RFM models complemented with website visits and social network interactions. *Network. Soc.-Cooperat. Conflict* **43**, 181–189 (2014)
10. Jøsang, A., Ismail, R.: The beta reputation system. In: *Proceedings of the 15th Bled Electronic Commerce Conference*, pp. 324–337 (2002)
11. Jurca, R.: Truthful reputation mechanisms for online systems. Ph.D. Thesis, EPFL (2007)
12. Kim, H., Gupta, S.: A comparison of purchase decision calculus between potential and repeat customers of online store. *Decis. Support Syst.* **47**(4), 477–487 (2009)
13. Ma, X.X., Liu, Y., Zhang, F.L., Qin, Z.G.: The markov-based evaluation on trust and reputation in peer-to-peer. In: *Proceedings of International Conference on Communications Circuits and Systems*, pp. 1552–1556 (2006)
14. Mittal, V., Kamakura, W.A.: Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics. *J. Mark. Res.* **38**(1), 131–142 (2001)
15. Noorian, Z., Marsh, S., Fleming, M.: zTrust: adaptive decentralized trust model for quality of service selection in electronic marketplaces. *Comput. Intell.* **32**(1), 127–164 (2016)
16. Qi, Y., Shen, Y., Lou, F.: The study on trust management model based on probability in distributed e-commerce system. In: *Second International Conference on E-Learning, E-Business, Enterprise Information Systems, and E-Government*, pp. 312–320 (2010)
17. Standing, S., Standing, C., Love, P.: A review of reserch on e-marketplaces 1997–2008. *Decis. Support Syst.* **49**(1), 41–51 (2010)
18. Teacy, W.T.L., Luck, M., Rogers, A., Jennings, N.R.: An efficient and versatile approach to trust and reputation using hierarchical Bayesian modelling. *Artif. Intell.* **193**(6), 149–185 (2012)
19. Tian, C., Yang, B.: A d-s evidence theory based fuzzy trust model in file-sharing p2p networks. *Peer-to-Peer Network. Appl.* **7**(4), 332–345 (2014)
20. Wang, P., Chao, K.M., Lo, C.C., Farmer, R.: An evidence-based scheme for web service selection. *Inf. Technol. Manage.* **12**(2), 161–172 (2011)
21. Wang, Y., Singh, M.P.: Evidence-based trust: a mathematical model geared for multiagent systems. *ACM Trans. Auton. Adapt. Syst.* **5**(4), 14 (2010)
22. Yu, B., Singh, M.P.: An evidential model of distributed reputation management. In: *Proceedings of International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 294–301 (2002)
23. Zhang, J.: Extensive experimental validation of a personalized approach for coping with unfair ratings in reputation systems. *J. Theoret. Appl. Electron. Commer. Res.* **6**(3), 43–64 (2011)

# Chinese Named Entity Recognition Based on B-LSTM Neural Network with Additional Features

Liubo Ouyang<sup>1</sup>, Yuan Tian<sup>1(✉)</sup>, Hui Tang<sup>1</sup>, and Boyun Zhang<sup>2</sup>

<sup>1</sup> School of Information Science and Engineering, Hunan University,  
Changsha 410082, China  
tianyuan.@hnu.edu.com

<sup>2</sup> Department of Information Technology, Hunan Police Academy,  
Changsha 410138, China

**Abstract.** Traditional methods for named entity recognition (NER) require heavy feature engineering to achieve high performance. We propose a novel neural network architecture for NER that detects word features automatically without feature engineering. Our approach uses word embedding as input, feeds them into a bidirectional long short-term memory (B-LSTM) for modeling the context within a sentence, and outputs the NER results. This study extends the neural network language model through B-LSTM, which outperforms other deep neural network models in NER tasks. Experimental results show that the B-LSTM with word embedding trained on a large corpus achieves the highest F-score of 0.9247, thus outperforming state-of-the-art methods that are based on feature engineering.

**Keywords:** Deep learning · Natural language processing  
Named entity recognition · Bidirectional long short-term memory  
Word embedding

## 1 Introduction

Recognition of named entities, such as names, locations, and organizations, is an important task in natural language processing (NLP) [1]. It extracts useful information from texts and facilitates downstream tasks, such as machine translation and question answering. Named entity recognition (NER) aims to localize and classify words in texts into predefined categories, such as person name (PER), organization (ORG), location (LOC), time expressions, quantities, and monetary values [2]. NER can be used as the basis of semantic analysis in the context of vector space model.

The mainstream methods for NER have transformed from rule-based methods [3] to machine learning [4]. However, most machine learning methods rely on large-scale labeled corpus and neglect the large amount of implicit information in unlabeled data, thereby performing poorly in low-frequency concept



extraction tasks. Since 2006, researchers have been combining various deep neural networks [5] and semi-supervised learning to solve NER and other chunking tasks [6]. Unlike traditional machine learning, deep learning can simultaneously classify patterns and learn representations, and greatly reducing the difficulty of NER tasks. Recurrent neural network (RNN) [7] is a special deep learning neural network. It consistently performs better than other approaches in English NER and many other sequence-labeling tasks. B-LSTM is based on RNN. In B-LSTM not only local context but also global context are considered.

We present a hybrid model of B-LSTM that learns word-level features and does not use language-specific features, gazetteers, or dictionaries. We use a small amount of supervised training data and unlabeled corpus for training word embedding and yet achieve accuracies that are on par with those of state-of-the-art models on labeled Peoples Daily in January 1998 dataset for Chinese F-score relative to the conditional random field (CRF) model, with improved values of 6.35% (PER), 7.94% (LOC), and 3.05% (ORG).

## 2 Related Work

There are many traditional machine learning methods, such as hidden Markov model [8], maximum entropy model [9], decision tree [10], and support vector machine [11]. Conditional random field (CRF) and so on. CRF have been widely adopted in NER [12]. In CRF, the characteristic functions is known, and the weights of each feature is the unique parameters trained by the model. So you configure the position relationship of features, in the training in accordance with the expected position, become a feature. According to the feature of the template, CRF get a lot of characteristic function from the training corpus, and then trained to get the weight of each characteristic function. These approaches for NER rely on limited annotated corpus or aligned parallel corpora and neglect the implicit information in large amounts of unlabeled data.

Traditional machine learning methods depend on feature engineering requires time-consuming and specialized domain knowledge. Therefore, many researchers have begun to pay attention to the semi-supervised learning method that is based on labeled and unlabeled data. Deep learning reduces the dependence on feature engineering. This techniques have successfully applied in NLP [13, 14]. Huang et al. presented a novel neural network [15] for learning word representations from the global context. RNN is a well-studied solution that allows a neural network to process variable-length input and have long-term memory [16]. This model has recently shown great success in diverse NLP tasks, such as machine translation [17]. Lample et al. proposed a RNN-based approach for NER [18]. It obtained the state-of-the-art results on English datasets. For NER, a B-LSTM model can theoretically consider an infinite amount of bidirectional context and eliminate the problem of context limited. The long-distance dependencies of the LSTM unit with the forget gate are highly nontrivial [19]. For sequence-labeling tasks, such as NER and speech recognition [20], a bidirectional LSTM (B-LSTM) model can consider an effectively infinite amount of bidirectional word sequence

and can eliminate the problem of limited context that applies to any feed-forward model. A bidirectional memory network is used to effectively consider numerous contextual links in the two directions of a word.

In addition to English NER, Chinese NER has been attracting attention. Chinese NER task is more complex than English NER because Chinese words do not have capitalization characteristics [21] and have unrestricted named entity lengths. Different entities have varying structures, no strict rules are followed (Numerous nests, aliases, abbreviations, and other issues exist.), and word formation, such as that of long names of ethnic minorities or translations of foreign names, is not uniform. Different areas include various scenarios, and the different extensions of named entities lead to fuzzy classification problems. This study uses the skip-n-gram method to obtain word embedding and then adds other feature vectors as inputs to the network. The optimal NER model is obtained after adjusting the B-LSTM network structure.

### 3 Model

Our network architecture is inspired by the work of Collobert et al. [22]. Instead of a feed-forward network, we use the B-LSTM network to acquire bidirectional word sequence for making predictions Fig. 1.

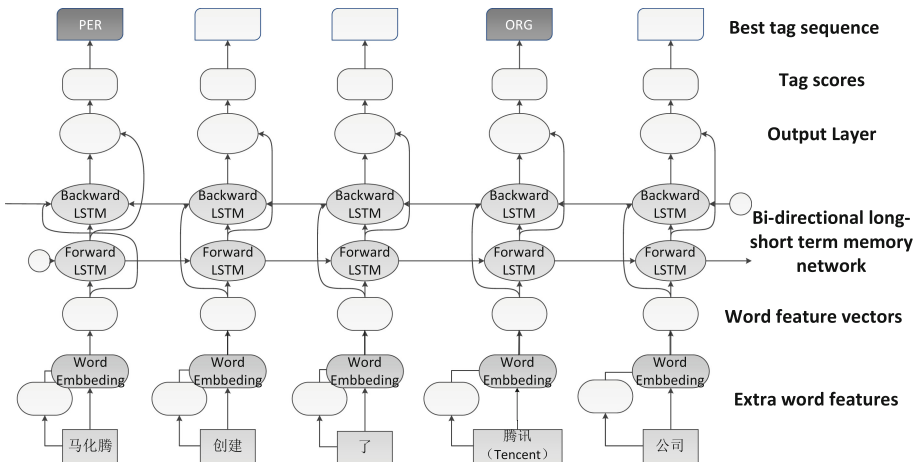


Fig. 1. Overview of proposed model framework.

Given an input sentence, we first compute the word embedding using Word2vec [23]. The word embedding is then concatenated and fed into the B-LSTMs neural network. Finally obtained the result from the output layer. In the model, the output of each layer is used as input to the next layer. The first step we get segmentation from a sentence without manual work. The words are then

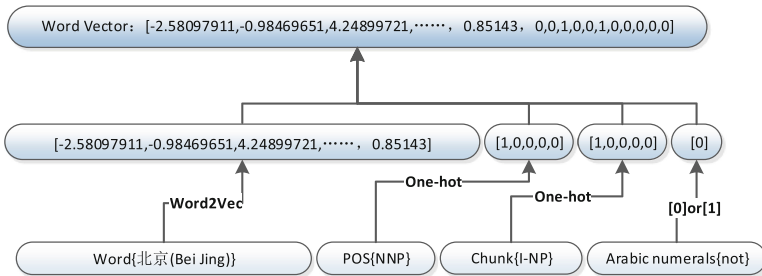
transformed into continuous vector representations at the second step. Finally every word embedding is fed into the B-LSTM network to obtain bidirectional word sequence for making predictions.

### 3.1 Data Preprocessing

The raw word and part-of-speech (POS) are extracted for the NER task in this study. We obtain the POS and chunking features, which help training the neural network model, using primitive annotations in corpus.

Jieba [24] is Chinese text segmentation system, is used for word segmentation and POS tagging. This system subdivides nouns into five categories, and these five tags are automatically transformed into one hot vector by our model. In addition to POS, chunking feature considerably influences the NER task. Identifying digital types in entity recognition is meaningful, but relevant classifications are rarely performed in deep learning. We add a bit for pattern recognition due to the complexity of Chinese numbers. This bit is 1 if the word contains Arabic numerals, otherwise the bit is 0.

We use Word2vec tool convert words into dense vectors that computers can process. Skip-gram is often insufficient in generating special word vectors only in the context of adjacent words. Therefore, the skip-n-gram method is adopted in this study. For completeness, we add our experimental corpus and Chinese Wikipedia raw corpus to the corpus. A 60D Word2vec model is obtained by conducting skip-5-gram unsupervised training for this corpus. We generate the word vector for each word by running the trained model and add a 5D POS tagging vector, a 5D chunking feature vector, and a 1D numeral tagging vector to form the 71D final word vector Fig. 2.



**Fig. 2.** Word distributed expression process. In our model, we use the previously mentioned N-Skip-gram method to obtain the better word vector.

### 3.2 B-LSTM Neural Network

In the node structure of LSTM (Fig. 3), a neuron (memory module) controlled by multiple control gates is designed for overcoming the phenomenon of vanishing gradient in RNNs. The gate can be controlled to disturb the information stored

in the neuron, thereby restricting new information. Therefore, the model can store and transfer information for a long time.

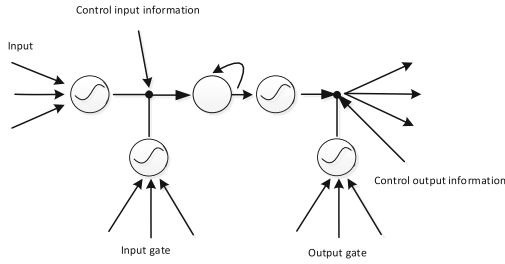


Fig. 3. A long and short memory neuron node.

We test the hidden layers from 2 to 4. We use an end-to-end training model using a minimum batch training method with a fixed learning rate. Every mini-batch contains the same number of sentences, and we use ADAM [25] for optimization. Each model layer is provided a fixed number of nodes to prevent overfitting, and we use dropout layers with a 0.5 dropout rate. We select cross-entropy function. Thus, the update speed of the weight is efficiently controlled. Contextual information is important in network training. The B-LSTM structure is presented in Fig. 4. LSTM is used instead of the traditional RNN to address the vanishing gradient problem and capture long-distance dependencies, which are solved by capturing information context. The forward propagation algorithm in the input sequence is arranged in the opposite direction in the hidden layer. The output layer is not updated until the hidden layers in the two directions process the input sequence. The output layer in the backpropagation algorithm passes the feedback information in the opposite direction throughout the period.

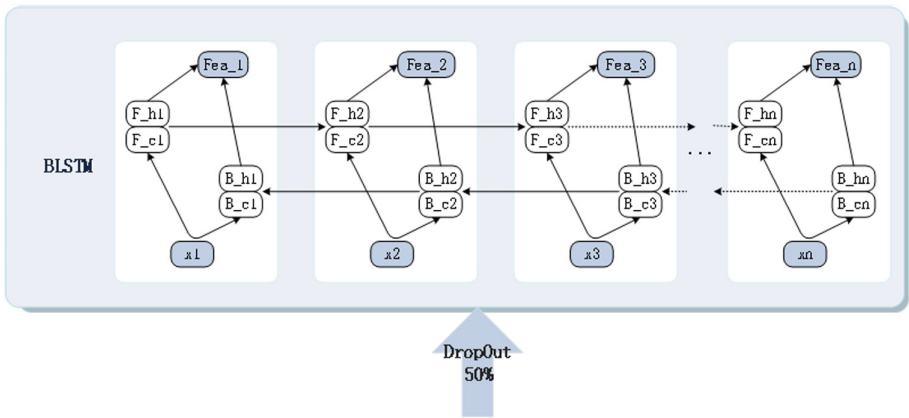


Fig. 4. Architect of the neural network hidden layer.

### 3.3 Softmax Layer

For the input sentence  $X = (x_1, x_2, \dots, x_n)$ ,  $P$  is defined as the output  $n \times k$  matrix for the score, where  $k$  is the number of entity tags (categories).  $P_{ij}$  corresponds to the score when the  $i$ -th word ( $x_i$ ) in the sentence categories  $j$ .  $y = (y_1, y_2, y_3, \dots, y_n)$  is the predicted sequence. The score is defined as (1).

$$s(X, y) = \sum_{i=1}^n P_{i, y_i} \quad (1)$$

The softmax classifier generates the probability of sequence  $y$  on the basis of all tag sequences as (2).

$$p(y|X) = \frac{e^{s(X, y)}}{\sum_{\tilde{y} \in Y_x} e^{s(X, \tilde{y})}} \quad (2)$$

According to the output after the neural network training, the maximum score of the Viterbi algorithm [26] is used to find the tag sequence as (3).

$$\log(p(y|X)) = s(X, y) - \log \underset{\tilde{y} \in Y_x}{\text{adds}}(X, \tilde{y}) \quad (3)$$

In the formula,  $Y_x$  represents all possible markup sequences for the input statement  $X$ . With the above approach, we can use our network structure to produce candidate sequences for the output tag, and the output sequence of the maximum score can be obtained by decoding as (4).

$$y^* = \arg \max_{\tilde{y} \in Y_x} s(X, \tilde{y}) \quad (4)$$

This objective function and its gradient be calculated using dynamic programming. In the reasoning stage given the neural network output  $P_{ij}$ . We use the maximum score of the Viterbi algorithm  $s(X, \tilde{y})$  to find the tag sequence, then obtain the final output by calculating likelihood probability of each tag on a softmax layer.

## 4 Experiment

### 4.1 Datasets

We use the labeled Peoples Daily in January 1998 dataset to compare our proposed model with other methods. The corpus consists of 357544 sentences (approximately 9200000 Chinese words), including 1044487 Chinese names (PER), 51708 foreign person names (translated PER), 218904 location names (LOC), and 87391 organization names (ORG). We split the corpus data using an 8:1:1 split for training, development, and testing. A fivefold cross-validation method is used in the experiment to generate significant objective statistical data in the evaluation.

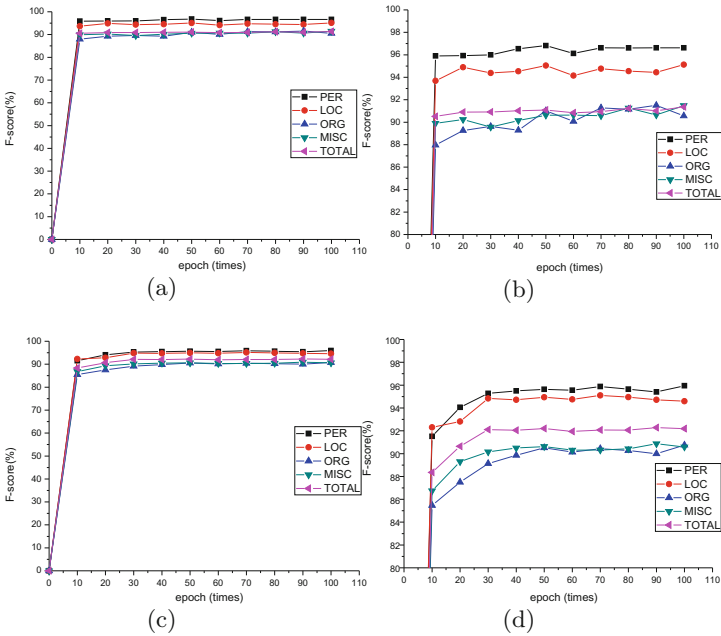
In the experiment, the entity is divided into four categories: person name (PER), place name (LOC), organization name (ORG), and mixed class (MISC). We use the rough segmentation provided by the labeled corpus as input and the unified annotation set as the system output tag. Three common entities categories are trained and labeled at the same time by our model, and the test results are counted.

### 4.2 Result

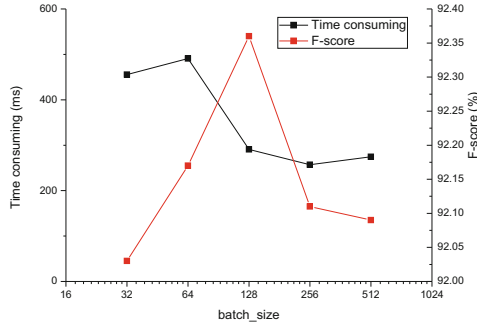
We implement all codes using the Tensorflow framework on the PowerEdge R530 server (16 GB memory/600 GB hard drive) in Python to train our models. Figure 5 shows the F-score of the model when the epoch is from 0 to 100 with different learning rates. Figures 5b and d magnify details from Figs. 5a and c. The variation range decreases after 50 iterations. Nearly no change occurs after the 70th round, and good convergence is observed.

We also consider the effect of batch\_size on the experimental results. Different time consuming and F-scores are yielded with the various values of batch\_size. See in Fig. 6.

The time consumption reduces with increasing batch\_size. When the batch\_size is 128, we obtain the maximum F-score and a relatively minimal time consumption. The accuracy rate is the highest, and the effect is the best.



**Fig. 5.** The changes of F-score for test sets A in the model with different parameter values



**Fig. 6.** Time consuming and F-scores with different batch\_size values after 50 iterations.

In view of the complex parameters of the model, numerous comparative tests are conducted to analyze the influence of each parameter on model performance. The experience result of the B-LSTM model in NER task is summarized (Table 1). Layers (L), Hidden nodes (H), Learning rate (R).

Comparison of the test results indicates that a high learning rate in the training does not necessarily correspond to good performance when the numbers of layers and nodes are fixed. The training model shows the best performance when the learning rate is 0.003, the model layer number is 2, and the number of hidden nodes is 256. Our model achieves an F-score of 92.47% on the Peoples Daily dataset without any manual labeling and gazetteer. We conduct a set of experiments and compare the results of our model with those of three other NER methods, namely, an NER method based on CRF [27], an NER method based

**Table 1.** Results of comparative tests.

Experiments	PER (%) F-score	LOC (%) F-score	ORG (%) F-score	MISC (%) F-score	Total F-score
L = 2 H = 128 R = 0.001	95.26	94.69	89.13	90.28	92.19
L = 2 H = 128 R = 0.003	95.56	96.12	92.23	90.62	92.27
L = 2 H = 128 R = 0.009	93.87	94.40	91.67	89.44	90.92
L = 2 H = 256 R = 0.001	95.82	95.93	90.86	90.29	92.09
L = 2 H = 256 R = 0.003	96.07	96.31	90.31	90.60	92.47*
L = 2 H = 256 R = 0.009	94.00	94.48	92.35	89.20	90.81
L = 4 H = 128 R = 0.001	93.57	95.54	88.89	90.21	91.59
L = 4 H = 128 R = 0.003	94.26	95.07	91.01	90.12	91.56
L = 4 H = 128 R = 0.009	93.56	93.90	85.79	89.33	90.65
L = 4 H = 256 R = 0.001	93.79	95.46	89.67	90.26	91.63
L = 4 H = 256 R = 0.003	94.56	96.01	92.63	90.67	92.15
L = 4 H = 256 R = 0.009	91.91	93.08	84.42	88.01	89.36

on deep neural network (DNN) [28], and a hybrid NER method based on deep belief network (DBN) [29], to verify the reliability and accuracy of our proposed NER model. B-LSTM with no chunking features is denoted as B-LSTM1, and B-LSTM that adds all additional features is denoted as B-LSTM2. The results are shown in Table 2.

**Table 2.** Comparison results of different named entity recognition method

Method	Entity type	Precision	Recall	F-score
CRF	PER	98.30%	66.70%	73.80%
	LOC	95.30%	64.20%	79.20%
	ORG	96.30%	76.20%	85.10%
DNN	PER	94.48%	89.62%	91.98%
	LOC	87.80%	80.05%	83.75%
	ORG	84.82%	75.63%	79.96%
DBN	PER	90.00%	90.96%	90.48%
	LOC	89.39%	89.93%	89.66%
	ORG	88.30%	88.93%	88.61%
B-LSTM1	PER	94.14%	94.18%	94.16%
	LOC	95.48%	93.66%	94.57%
	ORG	86.21%	90.91%	88.50%
B-LSTM2	PER	96.34%	95.80%	96.07%
	LOC	96.88%	95.75%	96.31%
	ORG	95.67%	85.51%	90.31%

The results imply that our model has a better labeling performance than the traditional CRF model or the other deep learning models. Our model considers two ways and the global context, thus achieving a higher F-score than the other methods. CRF considers only one way, thereby achieving worse results. The performances of DNN and DBN are also relatively poor because they consider only local context.

The F-scores of our method are 6.35% (PER), 7.94% (LOC), and 3.05% (ORG) higher than those of CRF; 4.09% (PER), 12.56% (LOC), and 10.35% (ORG) higher than those of DNN; and 5.59% (PER), 6.65% (LOC), and 1.70% (ORG) higher than those of DBN. The evaluation index of our proposed model is also significantly higher those of the other deep learning methods, thus indicating the effectiveness of our method. Furthermore, the proposed method has a good effect on the identification of mixed entities, such as numbers and time.

## 5 Conclusion

We propose an improved deep learning module that is superior to traditional and certain hybrid methods for NER. This module uses B-LSTM neural network and



skip-n-gram to learn word features, which provide wide contextual information and thus effectively solve the problem of different word order expressions. The word embedding in this model is obtained by unsupervised learning to minimize the need for manual annotation. The model adds the features of POS and chunks and can thus effectively improve the accuracy and recall in the NER task. Future research will focus on exploring additional DNN architectures using large datasets and reduced training time.

**Acknowledgments.** This work was supported by the project of National Natural Science Foundation of China (No. 61471169).

## References

1. Borthwick, A., Sterling, J., Agichtein, E., et al.: NYU: Description of the MENE Named Entity System as Used in MUC-7 (1998)
2. Chinchor, N.: MUC7 Named Entity Task Definition Message Understanding Conference (1997)
3. Abbas, A., Ekrem, V., Nazife, D.: ChemTok: a new rule based Tokenizer for chemical named entity recognition. *BioMed Res. Int.* **2016**(5), 1–9 (2016)
4. Zhu, J., Li, T., Liu, S.: Research on Tibetan name recognition technology under CRF. *J. Nanjing Univ.* **3494**, 234–250 (2016)
5. Collobert, R., Weston, J.: A unified architecture for natural language processing: deep neural networks with multitask learning. In: *International Conference*, pp. 160–167. *DBLP* (2008)
6. Mikolov, T.: *Statistical Language Models Based on Neural Networks* (2012)
7. Mikolov, T., Karafit, M., Burget, L., et al.: Recurrent neural network based language model. In: *INTERSPEECH Conference of the International Speech Communication Association*, Makuhari, Chiba, Japan, September 2010, pp. 1045–1048. *DBLP* (2010)
8. Ahmadi, F., Moradi, H.: A hybrid method for Persian named entity recognition. In: *Information and Knowledge Technology*, pp. 1–7. *IEEE* (2015)
9. Skenduli, M.P., Biba, M.: A named entity recognition approach for Albanian (2013)
10. Ekbal, A., Saha, S., Singh, D.: Ensemble based active annotation for named entity recognition. In: *International Conference on Advances in Computing, Communications and Informatics*, pp. 973–978. *IEEE* (2013)
11. Bam, S.B., Shahi, T.B.: Named entity recognition for Nepali text using support vector machines. *Intell. Inf. Manag.* **06**(2), 21–29 (2014)
12. Manamini, S.A.P.M., Ahamed, A.F., Rajapakshe, R.A.E.C., et al.: Ananya - a Named-Entity-Recognition (NER) system for Sinhala language. In: *Moratuwa Engineering Research Conference*, pp. 30–35. *IEEE* (2016)
13. Aryoyudanta, B., Adji, T.B., Hidayah, I.: Semi-supervised learning approach for Indonesian Named Entity Recognition (NER) using co-training algorithm. In: *International Seminar on Intelligent Technology and ITS Applications*, pp. 7–12. *IEEE* (2017)
14. He, H., Sun, X.: F-Score Driven Max Margin Neural Network for Named Entity Recognition in Chinese Social Media (2016)
15. Huang, E.H., Socher, R., Manning, C.D., et al.: Improving word representations via global context and multiple word prototypes. In: *Meeting of the Association for Computational Linguistics: Long Papers*. Association for Computational Linguistics, pp. 873–882 (2012)

16. Goller, C., Kuchler, A.: Learning task-dependent distributed representations by backpropagation through structure. In: IEEE International Conference on Neural Networks, vol. 1, pp. 347–352. IEEE (1996)
17. Wang, R., Panju, M., Gohari, M.: Classification-based RNN machine translation using GRUs (2017)
18. Lample, G., Ballesteros, M., Subramanian, S., et al.: Neural Architectures for Named Entity Recognition, pp. 260–270 (2016)
19. Gers, F.A., Schmidhuber, J., Cummins, F.: Learning to forget: continual prediction with LSTM. In: Ninth International Conference on Artificial Neural Networks, ICANN 1999, p. 2451 (2002)
20. Graves, A., Mohamed, A.R., Hinton, G.: Speech recognition with deep recurrent neural networks **38**(2003), 6645–6649 (2013)
21. Liu, Y., Burkhart, C., Hearne, J., et al.: Enhancing sumerian lemmatization by unsupervised named-entity recognition. In: Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 1446–1451 (2015)
22. Collobert, R., Weston, J., Karlen, M., et al.: Natural language processing (almost) from scratch. *J. Mach. Learn. Res.* **12**(1), 2493–2537 (2011)
23. Mikolov, T., Sutskever, I., Chen, K., et al.: Distributed representations of words and phrases and their compositionality. *Adv. Neural Inf. Process. Syst.* **26**, 3111–3119 (2013)
24. <https://github.com/fxsjy/jieba>
25. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. *Comput. Sci.* (2014)
26. Forney, G.D.J.: The viterbi algorithm. *Proc. IEEE* **61**(5), 268–278 (1973)
27. Jing, X.: Research on named entity recognition based on word vector and conditional random field. *Wirel. Internet Technol.* (1), 111–112 (2017)
28. Wang, G., Cai, Y., Ge, F.: Using hybrid neural network to address Chinese named entity recognition. In: IEEE, International Conference on Cloud Computing and Intelligence Systems, pp. 433–438. IEEE (2015)
29. Feng, Y.-T., Zhang, H.-J., Hao, W.-N., Chen, G.J.: Named entity recognition based on deep belief net. *Comput. Sci.* **43**(4), 224–230 (2016)

# Amplified Locality-Sensitive Hashing for Privacy-Preserving Distributed Service Recommendation

Lianyong Qi<sup>1,2(✉)</sup>, Wanchun Dou<sup>2</sup>, Xuyun Zhang<sup>3</sup>, and Shui Yu<sup>4</sup>

<sup>1</sup> School of Information Science and Engineering,  
Qufu Normal University, Rizhao 276826, China  
lianyongqi@gmail.com

<sup>2</sup> State Key Laboratory for Novel Software Technology,  
Department of Computer Science and Technology,  
Nanjing University, Nanjing 210023, China  
douwc@nju.edu.cn

<sup>3</sup> Department of Electrical and Computer Engineering,  
University of Auckland, Auckland 1023, New Zealand  
xuyun.zhang@auckland.ac.nz

<sup>4</sup> School of Information Technology, Deakin University,  
Melbourne 3125, Australia  
syu@deakin.edu.au

**Abstract.** With the ever-increasing volume of services registered in various web communities, service recommendation techniques, e.g., Collaborative Filtering (i.e., CF) have provided a promising way to alleviate the heavy burden on the service selection decisions of target users. However, traditional CF-based service recommendation approaches often assume that the recommendation bases, i.e., historical service quality data are centralized, without considering the distributed service recommendation scenarios as well as the resulted privacy leakage risks. In view of this shortcoming, Locality-Sensitive Hashing (LSH) technique is recruited in this paper to protect the private information of users when distributed service recommendations are made. Furthermore, LSH is essentially a probability-based search technique and hence may generate “False-positive” or “False-negative” recommended results; therefore, we amplify LSH by AND/OR operations to improve the recommendation accuracy. Finally, through a set of experiments deployed on a real distributed service quality dataset, i.e., *WS-DREAM*, we validate the feasibility of our proposed recommendation approach named *DistSR<sub>Amplify-LSH</sub>* in terms of recommendation accuracy and efficiency while guaranteeing privacy-preservation in the distributed environment.

**Keywords:** Distributed service recommendation · Collaborative Filtering  
Privacy-preservation · Recommendation accuracy  
Amplified Locality-Sensitive Hashing

## 1 Introduction

With the ever-increasing volume of services registered in various web communities (e.g., *Amazon* and *IBM*), it is becoming a nontrivial task to find the web services that a target user is really interested in from massive candidates [1–3]. In this situation,

various recommendation techniques, e.g., Collaborative Filtering (i.e., CF) are recruited to make service recommendation so as to reduce the heavy burden on the service selection decisions of target users [4–7].

However, traditional CF-based service recommendation approaches (e.g., user-based CF, item-based CF and hybrid CF) often assume that the service recommendation bases, i.e., historical service quality data are centralized, without considering the distributed service recommendation scenarios when historical service quality data are from multiple independent platforms (e.g., historical service quality data observed by user *A* and user *B* are recorded by *Amazon* and *IBM*, respectively). Generally, two challenges are present in the above distributed service recommendation problems. First, *IBM* is often not willing to share its data with *Amazon* due to privacy concerns, vice versa, which hampers the data collaboration between *Amazon* and *IBM* and consequently renders the distributed recommendation process infeasible. Second, the volume of service quality data recorded by *Amazon* and *IBM* may become increasingly huge with updates over time, which brings additional communication cost between these two platforms; as a consequence, the recommendation efficiency is reduced severely and cannot satisfy the quick response requirements from target users.

In view of these challenges, Locality-Sensitive Hashing (LSH) technique is recruited in this paper to achieve privacy-preserving and efficient service recommendation in the distributed environment. Furthermore, we amplify LSH so as to reduce the “False-positive” and “False-negative” recommended results. Finally, according to the amplified LSH, we propose a novel distributed service recommendation approach, i.e.,  $DistSR_{Amplify-LSH}$ . With the inherent nature of amplified LSH,  $DistSR_{Amplify-LSH}$  can achieve a good recommendation performance in terms of recommendation accuracy and efficiency while guaranteeing privacy-preservation.

In summary, the contributions of our paper are three-fold.

- (1) We introduce LSH technique into distributed service recommendation, so as to protect the private information of users and improve the recommendation efficiency. Meanwhile, we recognize the possible “False-positive” and “False-negative” recommended results produced by LSH-based recommendation approach.
- (2) We amplify LSH by integrating the AND/OR operations, to reduce the “False-positive” and “False-negative” recommended results and improve the recommendation accuracy.
- (3) Extensive experiments are conducted on a real distributed service quality dataset *WS-DREAM* to validate the feasibility of our proposal. Experiment results show that  $DistSR_{Amplify-LSH}$  outperforms other state-of-the-art approaches in terms of recommendation accuracy and efficiency while guaranteeing privacy-preservation.

The rest of paper is structured as follows. We motivate our paper in Sect. 2 and formulate the distributed service recommendation problems in Sect. 3. In Sect. 4, LSH technique is introduced briefly and afterwards, an amplified LSH-based recommendation approach, i.e.,  $DistSR_{Amplify-LSH}$  is proposed to solve the privacy-preserving distributed service recommendation problems. In Sect. 5, a set of experiments are conducted to validate the feasibility of our proposal. Related work and comparison analyses are presented in Sect. 6. And finally, in Sect. 7, we conclude the whole paper and point out the future research directions.

## 2 Motivation

An example is presented in Fig. 1 to demonstrate the motivation of our paper. In Fig. 1, there are two platforms, i.e., *Amazon* and *IBM* which record the historical quality data of services  $\{ws_1, \dots, ws_n\}$  observed by target user  $u_{target}$  and user  $u_1$ , respectively. In this situation, according to the traditional CF recommendation approaches [7], it is necessary to calculate the similarity between  $u_{target}$  and  $u_1$  (denoted by  $sim(u_{target}, u_1)$ ). However, as Fig. 1 shows, the calculation of  $sim(u_{target}, u_1)$  requires the collaboration between *Amazon* and *IBM* and often faces the following three challenges (see Fig. 1).

- (1) Due to privacy concerns, *IBM* is often not willing to share its data with *Amazon*, which hampers the cross-platform collaboration between *Amazon* and *IBM* and consequently renders the calculation of  $sim(u_{target}, u_1)$  infeasible.
- (2) For both *Amazon* and *IBM*, their volume of recorded service quality data may become increasingly huge with updates over time; in this situation, the collaboration efficiency and scalability between *Amazon* and *IBM* may be decreased significantly and cannot satisfy the quick response requirements of target users.
- (3) It is possible to generate the “False-positive” or “False-negative” recommended results; in this situation, user satisfaction is reduced significantly.

In view of these challenges, we amplify LSH and further propose an amplified LSH-based service recommendation approach, i.e.,  $DistSR_{Amplify-LSH}$ , so as to achieve privacy-preserving and efficient service recommendation in the distributed environment.

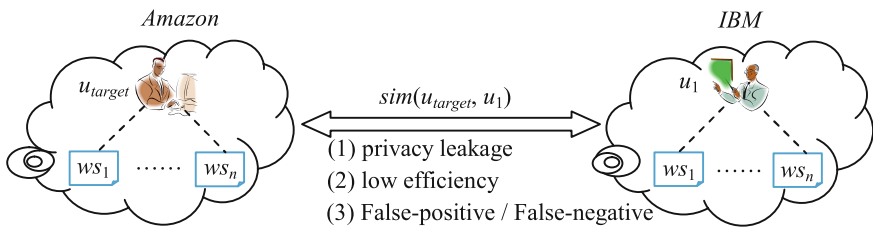


Fig. 1. Distributed service recommendation and its challenges: an example

## 3 Problem Formulation

Generally, the distributed service recommendation problems involving multiple platforms can be formulated as a five-tuple  $Dist\_Ser\_Rec(PF, U, u_{target}, WS, q)$ , where

- (1)  $PF = \{pf_1, \dots, pf_z\}$ : the set of platforms that record historical service quality data observed by users; e.g.,  $z = 2$  holds in Fig. 1.
- (2)  $U = \{u_1, \dots, u_m\}$ : the set of users. For each user, his/her observed service quality data are recorded by a platform in set  $PF$ .
- (3)  $u_{target}$ : a target user to whom a recommender system intends to recommend services. Here,  $u_{target} \in U$  holds.

- (4)  $WS = \{ws_1, \dots, ws_n\}$ : the set of candidate web services. For simplicity, we assume that the services held by different platforms  $pf_1, \dots, pf_z$  are the same. Specifically, if user  $u$  has never invoked web service  $ws$ , then the quality data of  $ws$  observed by  $u$  is denoted by 0.
- (5)  $q$  is a quality dimension of web services, e.g., *response time*. For simplicity, only a quality dimension is considered in the following discussions.

With the above formulation, we can specify the privacy-preserving distributed service recommendation problems more formally as below: according to the historical quality data (over dimension  $q$ ) of services ( $\in WS$ ) observed by users ( $\in U$ ) in multiple platforms ( $\in PF$ ), select an optimal service ( $\in WS$ , never invoked by  $u_{target}$ ) and recommend it to  $u_{target}$ , during which the private information of users (e.g., *the service quality data observed by a user, the service set invoked by a user*) should not be exposed to other users. To achieve this goal, a novel distributed service recommendation approach  $DistSR_{Amplify-LSH}$  is introduced in the next section.

## 4 A Privacy-Preserving Distributed Service Recommendation

### Approach: $DistSR_{Amplify-LSH}$

We firstly introduce the Locality-Sensitive Hashing technique briefly in Subsect. 4.1; afterwards, in Subsect. 4.2, we introduce the details of our proposed amplified LSH-based service recommendation approach  $DistSR_{Amplify-LSH}$ .

#### 4.1 Locality-Sensitive Hashing

The main idea behind LSH [8] is: select a specific hash function (or a hash function family) so that (1) for two neighboring points in original data space, they are still neighbors after hash with large probability (2) for two non-neighboring points in original data space, they are still non-neighbors after hash with large probability. More formally, a hash function  $h(\cdot)$  is called a LSH function iff conditions (1) and (2) hold, where  $x$  and  $y$  are two points in original data space,  $d(x, y)$  denotes the distance between points  $x$  and  $y$ ,  $h(x)$  is the hash value of point  $x$ ,  $P(A)$  represents the probability that event  $A$  holds,  $\{d_1, d_2, p_1, p_2\}$  are a set of thresholds.

$$\text{If } d(x, y) \leq d_1, \text{ then } P(h(x) = h(y)) \geq p_1 \quad (1)$$

$$\text{If } d(x, y) \geq d_2, \text{ then } P(h(x) = h(y)) \leq p_2 \quad (2)$$

The rationale of LSH-based neighbor search can be explained intuitively in Fig. 2, where  $h(\cdot)$  is a LSH function and  $\{A, B, C, D\}$  are four points in original data space. As points  $A$  and  $B$  are neighbors, they are projected into the same bucket of the LSH table, i.e.,  $h(A) = h(B)$  with large probability; while points  $C$  and  $D$  are non-neighbors, therefore, they are projected into different buckets, i.e.,  $h(C) \neq h(D)$  with large probability. Conversely, if  $h(A) = h(B)$  holds, then points  $A$  and  $B$  are neighbors in original data space with large probability; if  $h(C) \neq h(D)$  holds, then points  $C$  and  $D$  are non-neighbors in original data space with large probability.

Assume  $X$  is a new point and we hope to find its similar neighbors in a privacy-preserving way (i.e., without revealing the inner details of point  $X$ ), then we can first calculate point  $X$ 's hash value, i.e.,  $h(X)$ . Next, if  $h(X) = h(A) = h(B)$  holds, then points  $A$  and  $B$  are point  $X$ 's neighbors; while if  $h(X) = h(C)$  holds, then point  $C$  is the only neighbor of point  $X$ . This is the main rationale of LSH-based neighbor search. As the LSH table can be built offline, the neighbor search efficiency can be improved significantly. Besides, for a point (e.g., point  $A$  in Fig. 2), only its hash value with little privacy is used in LSH-based neighbor search; as a consequence, the private information of this point can be protected.

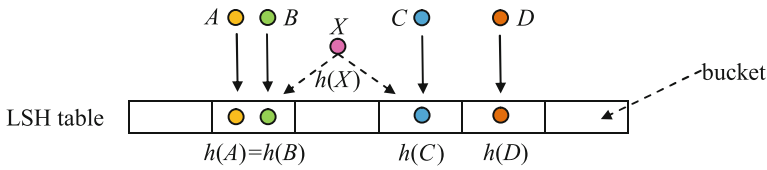


Fig. 2. Rationale of LSH-based neighbor search

#### 4.2 $DistSR_{Amplify-LSH}$ : Amplified LSH-Based Service Recommendation Approach

The amplified LSH-based service recommendation approach, i.e.,  $DistSR_{Amplify-LSH}$  is essentially a variant of user-based CF approach, which mainly consists of the four steps in Fig. 3.

- Step 1: Build user indexes offline based on LSH.** Choose a LSH function to project the users into corresponding buckets offline, based on the users' observed service quality data. Then the bucket No. can be regarded as the indexes for users.
- Step 2: Define "neighbor" relationship between users based on amplified LSH.** In order to reduce the "False-positive" and "False-negative" search results, we amplify LSH to define a novel "neighbor" relationship between two users based on the user indexes derived in Step 1.
- Step 3: Online neighbor finding for  $u_{target}$ .** According to the "neighbor" relationship defined in Step 2, find the similar neighbors of  $u_{target}$ .
- Step 4: Service recommendation.** According to the historical service quality data observed by neighbors (derived in Step 3) of  $u_{target}$ , predict the quality of services never invoked by  $u_{target}$ , and return the quality-optimal service to  $u_{target}$ .

Fig. 3. Four steps of amplified LSH-based service recommendation approach  $DistSR_{Amplify-LSH}$

**Step 1: Build User Indexes Offline Based on LSH**

In this step, we choose a LSH function  $h(\cdot)$  to build indexes for all the  $m$  users in distributed platforms. The selection of LSH function  $h(\cdot)$  depends on the adopted “distance” type of  $d(x, y)$  (see conditions (1)–(2) in Subsect. 4.1). As Pearson Correlation Coefficient (PCC) [9] is often taken as the similarity measurement or distance type in service recommendation, we adopt the LSH function  $h(\cdot)$  corresponding to PCC distance for user indexes building.

Concretely, for a user  $u$ , we can model her/his historical quality data over  $n$  services, i.e.,  $ws_1, \dots, ws_n$  with an  $n$ -dimensional vector  $\vec{u} = (ws_1.q, \dots, ws_n.q)$ , where  $q$  is a quality dimension of services and  $ws_j.q$  denotes service  $ws_j$ 's quality over dimension  $q$  observed by user  $u$ . Specifically,  $ws_j.q = 0$  if user  $u$  has never invoked web service  $ws_j$  before. Then for vector  $\vec{u}$ , its LSH function  $h(\vec{u})$  can be represented by (3) [10]. Here,  $\vec{v}$  is an  $n$ -dimensional vector  $(v_1, \dots, v_n)$  where  $v_j (1 \leq j \leq n)$  is a random value in range  $[-1, 1]$ ; symbol “ $\circ$ ” denotes the dot product between two vectors. To ease the understanding of readers, we explain the physical meaning of (3) as follows: take vector  $\vec{v}$  as a hyper plane for space partition; if two vectors  $\vec{u}_a$  and  $\vec{u}_b (1 \leq a, b \leq m \text{ and } a \neq b)$  are located on the same side of  $\vec{v}$  (i.e., both  $\vec{u}_a \circ \vec{v} > 0$  and  $\vec{u}_b \circ \vec{v} > 0$  hold, or, both  $\vec{u}_a \circ \vec{v} \leq 0$  and  $\vec{u}_b \circ \vec{v} \leq 0$  hold), then  $\vec{u}_a$  and  $\vec{u}_b$  can be regarded as similar (with high probability).

$$h(\vec{u}) = \begin{cases} 1 & \text{if } \vec{u} \circ \vec{v} > 0 \\ 0 & \text{if } \vec{u} \circ \vec{v} \leq 0 \end{cases} \tag{3}$$

Thus through the LSH function  $h(\cdot)$  in (3), each user  $u (u \in U)$  is converted to a binary value  $h(\vec{u}) (\in \{0, 1\})$ . However, a single LSH function often falls short in profiling the service quality data observed by a user. Therefore, we choose  $r$  LSH functions  $h_1(\cdot), \dots, h_r(\cdot)$  to generate an  $r$ -dimensional vector  $H(\vec{u}) = (h_1(\vec{u}), \dots, h_r(\vec{u}))$  offline for user  $u$ . Then  $H(\vec{u})$  can be regarded as the index for user  $u$ .

**Step 2: Defining “Neighbor” Relationship Between Users Based on Amplified LSH**

According to the LSH theory, two users  $u_a$  and  $u_b$  are similar neighbors iff they fall into the same bucket, i.e.,  $H(\vec{u}_a) = H(\vec{u}_b)$ , vice versa. However, as formulas (1) and (2) indicate (see Subsect. 4.1), LSH is essentially a probability-based search technique; therefore, it is inevitable to produce unsatisfactory search results. In other words, LSH may generate “False-positive” (i.e., dissimilar users of the target user are regarded as similar) or “False-negative” (i.e., similar neighbors of the target user are regarded as dissimilar) neighbor search results, which reduces the service recommendation accuracy severely.

To overcome this shortcoming, we amplify LSH by adopting the AND/OR operations over multiple LSH functions or LSH tables. Concretely, the following two strategies are taken to amplify LSH.

**Strategy-1:** OR operation over  $r$  LSH functions so as to reduce the “False-negative” search results.



In Step 1,  $r$  LSH functions  $h_1(\cdot), \dots, h_r(\cdot)$  are recruited to build index for user  $u$ , i.e.,  $H(\vec{u}) = (h_1(\vec{u}), \dots, h_r(\vec{u}))$ . Here, in order to reduce the “False-negative” search results, OR operation is taken over these  $r$  LSH functions. More concretely, two users  $u_a$  and  $u_b$  are regarded as similar iff the condition in (4) holds. Here, we utilize equation  $H(\vec{u}_a) \stackrel{OR}{=} H(\vec{u}_b)$  to represent the “similarity” relationship defined in condition (4).

$$\exists j, \text{ satisfy } h_j(\vec{u}_a) = h_j(\vec{u}_b) (1 \leq j \leq r) \tag{4}$$

**Strategy-2:** AND operation over  $T$  LSH tables so as to reduce the “False-positive” search results.

In order to reduce the “False-positive” search results, we repeat Step 1  $T$  times to generate  $T$  hash tables. Next, AND operation is taken over these  $T$  LSH tables. More concretely, two users  $u_a$  and  $u_b$  are regarded as similar neighbors iff the condition in (5) holds. Here,  $H_x(\cdot)$  denotes the LSH function family (see Strategy-1) recruited in  $x$ -th LSH table; we utilize  $u_a \stackrel{sim}{\leftrightarrow} u_b$  to represent the “similarity” relationship defined in condition (5).

$$\forall x \in \{1, \dots, T\}, \text{ satisfy } H_x(\vec{u}_a) \stackrel{OR}{=} H_x(\vec{u}_b) \tag{5}$$

Thus through Strategy-1 and Strategy-2, we amplify LSH and define a novel “neighbor” relationship between two users  $u_a$  and  $u_b$ , i.e.,  $u_a \stackrel{sim}{\leftrightarrow} u_b$ , so as to reduce the “False-negative” and “False-positive” search results and improve the recommendation accuracy.

**Step 3: Online Neighbor Finding for  $u_{target}$**

The index for  $u_{target}$ , i.e.,  $H(\vec{u}_{target})$  can be calculated based on the LSH function family  $\{h_1(\cdot), \dots, h_r(\cdot)\}$  chosen in Step 1. Afterwards, if  $u_{target} \stackrel{sim}{\leftrightarrow} u_a$  holds according to (4) and (5), then  $u_a$  can be regarded as a similar neighbor of  $u_{target}$  and is put into set  $NB\_Set$ , where  $NB\_Set$  denotes the neighbor set of  $u_{target}$ .

**Step 4: Service Recommendation**

If  $NB\_Set = \emptyset$ , then no similar neighbors of  $u_{target}$  are returned and the subsequent service recommendation fails accordingly. Otherwise, we utilize the similar neighbors in  $NB\_Set$  to make a service recommendation to  $u_{target}$ . Concretely, for each service  $ws$  ( $\in WS$ ) never invoked by  $u_{target}$ , we can predict its quality over dimension  $q$  observed by  $u_{target}$ , denoted by  $ws.q_{target}$ , based on the equation in (6). Here,  $ws.q_a$  denotes  $ws$ ’ quality over dimension  $q$  observed by user  $u_a$ ;  $NB\_Set^\#$  represents the set of  $u_{target}$ ’s neighbors who have invoked service  $ws$  before, i.e.,  $NB\_Set^\# = \{u_a \mid u_a \in NB\_Set, ws.q_a \neq 0\}$ . Finally, we select a service with the optimal quality predicted by (6) and recommend it to  $u_{target}$ .

$$ws.q_{target} = \frac{1}{|NB\_Set^\#|} * \sum_{u_a \in NB\_Set^\#} ws.q_a \tag{6}$$

Thus through the above four steps of  $DistSR_{Amplify-LSH}$ , we can recommend the quality-optimal service to the target user, so as to finish the privacy-preserving distributed service recommendation process. More formally, our proposal can be specified by the following pseudo code.

---

**Algorithm:**  $DistSR_{Amplify-LSH}(PF, U, WS, u_{target}, q)$

---

**Inputs:** (1)  $PF = \{pf_1, \dots, pf_z\}$ : platform set;  
 (2)  $U = \{u_1, \dots, u_m\}$ : user set;  
 (3)  $WS = \{ws_1, \dots, ws_n\}$ : web service set;  
 (4)  $u_{target}$ : a target user;  
 (5)  $q$ : a quality dimension of web services.

**Output:**  $ws_{optimal}$ : a candidate service with optimal predicted quality

---

```

/* Step 1: Build user indexes offline based on LSH */
1  for  $j = 1$  to  $r$  do //  $r$  LSH functions in each LSH table
2    for  $k = 1$  to  $n$  do //  $n$ -dimensional vector depicting a user
3       $v_{jk} = \text{random}[-1, 1]$ 
4    end for
5    for  $a = 1$  to  $m$  do //  $m$  users
6      calculate  $h_j(\overline{u_a})$  based on (3)
7    end for
8  end for
9  for  $a = 1$  to  $m$  do
10    $H(\overline{u_a}) = (h_1(\overline{u_a}), \dots, h_r(\overline{u_a}))$ 
11 end for
12 repeat line 1-11  $T$  times to generate  $T$  LSH tables offline

/* Step 2 - Step 3: Define "neighbor" relationship between users based on amplified LSH
and online neighbor finding for  $u_{target}$  */
13 count = 0 // number of LSH tables in which the condition in (4) holds
14 for  $a = 1$  to  $m$  do
15   for  $x = 1$  to  $T$  do //  $T$  LSH tables
16     if  $H_x(\overline{u_a}) = H_x(\overline{u_{target}})$  holds based on (4)
17       then count ++
18     continue
19   else break
20   end if
21 end for
22 if count =  $T$ 
23   then put  $u_a$  into  $NB\_Set$ 
24   end if
25 end for

/* Step 4: Service recommendation */
26 for  $i = 1$  to  $n$  do //  $n$  candidate web services
27   if  $ws_i.q_{target} = 0$  //  $u_{target}$  has never invoked  $ws_i$  before
28   then COUNT = 0 // size of set  $NB\_Set^{\#}$  in (6)
29     for  $K = 1$  to  $|NB\_Set|$  do

```

```

30         if  $ws_r.q_k \neq 0$ 
31         then COUNT ++
32              $ws_r.q_{target} = ws_r.q_{target} + ws_r.q_k$ 
33         end if
34     end for
35      $ws_r.q_{target} = ws_r.q_{target} / \text{COUNT}$ 
36 end if
37 end for
38  $ws_{optimal} = \{ws_i \mid ws_i.q_{target} = \text{OPTIMAL}\{ws_i.q_{target}\}\}$ 
39 return  $ws_{optimal}$  to  $u_{target}$ 

```

---

## 5 Experiments

### 5.1 Experiment Dataset and Deployment

In this section, a set of experiments are conducted to validate the feasibility of our proposed service recommendation approach  $DistSR_{Amplify-LSH}$ . The experiments are based on a real distributed service quality dataset  $WS-DREAM$  [11] which collects the historical quality data of 5825 web services (from different countries) observed by 339 users. In our experiments, each country that hosts a set of web services is regarded as a distributed platform, so as to simulate the distributed service recommendation scenario. Besides, only a quality dimension of services, i.e., *response time* is considered; furthermore, some entries in the user-service quality matrix are removed randomly to simulate the missing quality data.

To demonstrate the feasibility and advantages of our proposed  $DistSR_{Amplify-LSH}$  approach, we compare our proposal with three state-of-the-art approaches:  $UPCC$  [12],  $P-UIPCC$  [13] and  $PPICF$  [14]. Concretely, the following two evaluation measures are examined and compared, respectively (as user privacy can be protected well by the intrinsic nature of LSH technique, we will not evaluate the capability of privacy-preservation of our proposal here).

- (1) *time cost*: consumed time for generating the final recommended results, through which we can test the recommendation efficiency.
- (2) *MAE* (Mean Absolute Error): average difference between predicted quality and real quality of recommended services, through which we can test the recommendation accuracy (the smaller the better).

The experiments are conducted on a Lenovo computer with 2.40 GHz processors and 12.0 GB RAM. The machine runs under Windows 10, JAVA 8 and MySQL 5.7. Each experiment is carried out 10 times and the average experiment results are reported finally.

### 5.2 Experiment Results and Analyses

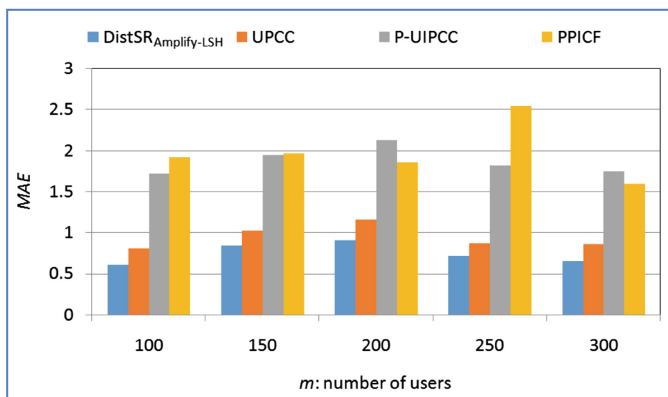
Concretely, the following five profiles are tested and compared in our experiments. Here,  $m$  and  $n$  denote the number of users and number of web services, respectively;

$T$  and  $r$  denote the number of LSH tables and number of hash functions in each LSH table, respectively.

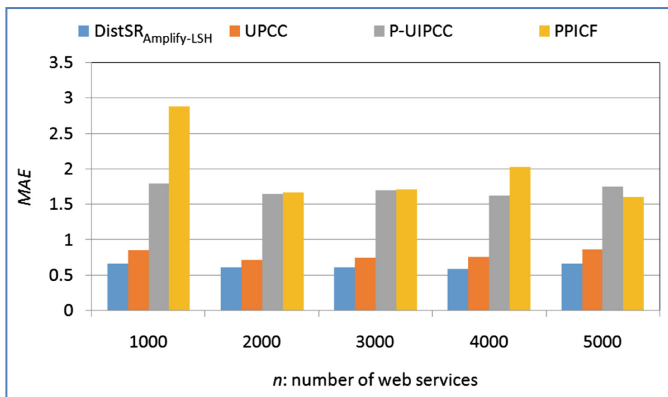
**Profile 1: Recommendation Accuracy Comparison**

In this profile, we test the recommendation accuracy (i.e.,  $MAE$ , the smaller the better) of four approaches. The experiment parameters are set as follows:  $m$  is varied from 100 to 300,  $n$  is varied from 1000 to 5000, and  $T = r = 10$  holds. The experiment results are shown in Fig. 4.

In Fig. 4(a),  $n = 5000$  holds. The experiment results show that the recommendation accuracy values of  $P$ -UIPCC and  $PPICF$  approaches are often low (i.e.,  $MAE$  values are high). This is because several approximation strategies (e.g., data obfuscation, divide-merge) are adopted in these two approaches so as to protect the user privacy, while the approximation strategies reduce the recommendation accuracy significantly. No data approximation strategy is recruited in  $UPCC$  approach; therefore, the recommendation accuracy of  $UPCC$  is higher than  $P$ -UIPCC and  $PPICF$  approaches. While our proposed  $DistSR_{Amplify-LSH}$  approach outperforms the other three ones in



(a)  $n = 5000$



(b)  $m = 300$

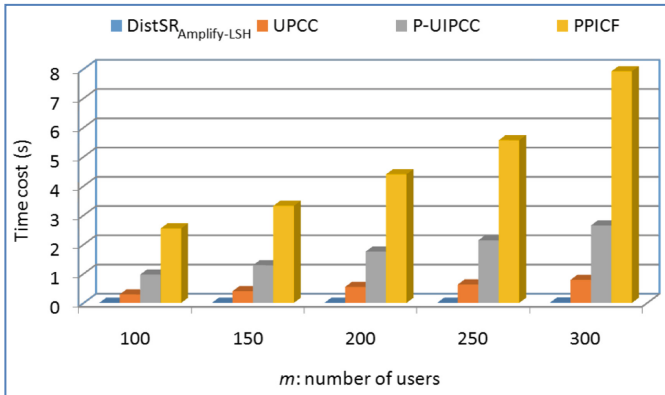
**Fig. 4.** Recommendation accuracy comparison of four approaches

terms of recommendation accuracy, which is due to the following two reasons. First, according to the inherent nature of LSH, only the “most similar” neighbors of a target user can be returned for subsequent service recommendation in our  $DistSR_{Amplify-LSH}$  approach; as a consequence, high recommendation accuracy is often guaranteed. Second, the AND/OR operations are adopted in our approach to amplify LSH, through which the “False-positive” and “False-negative” search results are reduced; accordingly, the recommendation accuracy is improved.

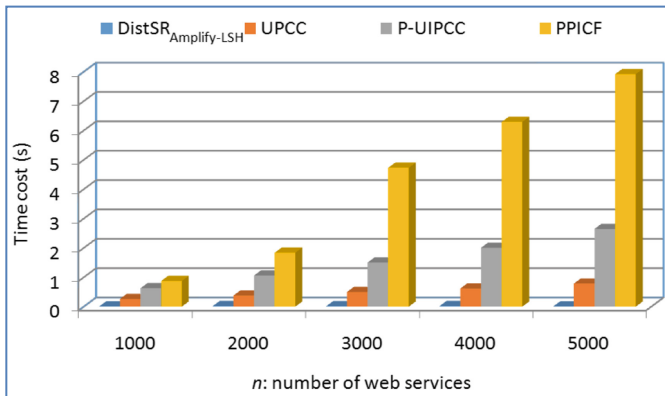
Similar experiment results can be observed from Fig. 4(b) where  $m = 300$  holds and  $n$  is varied from 1000 to 5000. The reasons are the same as those in Fig. 4(a) and are not repeated here.

**Profile 2: Recommendation Efficiency Comparison**

In this profile, we test the recommendation efficiency of four approaches. The recruited experiment parameters are set as below:  $m$  is varied from 100 to 300,  $n$  is varied from 1000 to 5000,  $T = r = 10$  holds. The concrete experiment results are shown in Fig. 5.



(a)  $n = 5000$

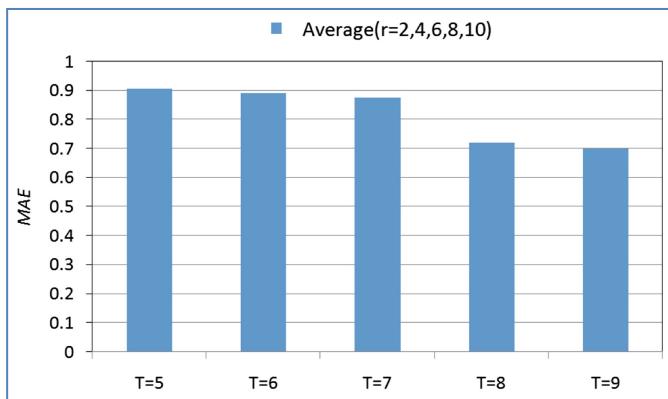


(b)  $m = 300$

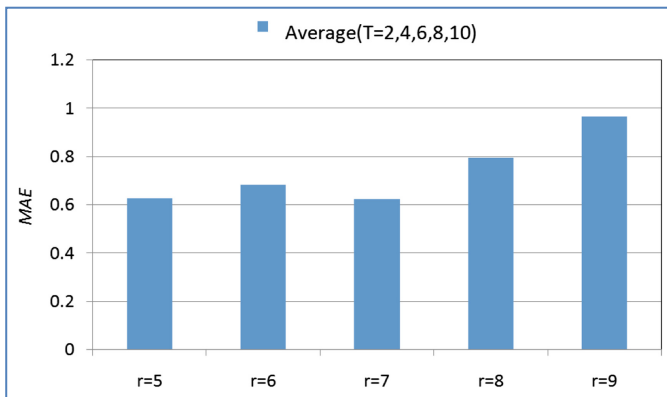
**Fig. 5.** Recommendation efficiency comparison of four approaches

In Fig. 5(a),  $n = 5000$  holds. The experiment results indicate that the time costs of  $P$ -UIPCC and  $PPICF$  approaches both increase with the growth of  $m$  and are often high because of the recruited additional privacy-preservation operations (e.g., data obfuscation, divide-merge). Similar experiment result can be observed from the  $UPCC$  approach in Fig. 5; however,  $UPCC$  performs better than  $P$ -UIPCC and  $PPICF$  in recommendation efficiency as no additional privacy-preservation operations are adopted. While our  $DistSR_{Amplify-LSH}$  approach outperforms the other three approaches in terms of recommendation efficiency, because the recruited LSH tables can be built offline and the rest step of online neighbor finding is very efficient (in the best cases, the time complexity of online neighbor finding is  $O(1)$  [15]). Similar experiment results can be observed from Fig. 5(b), which are not analyzed repeatedly.

**Profile 3: Recommendation Accuracy of  $DistSR_{Amplify-LSH}$  with Respect to  $T$  and  $r$**   
 The number of LSH tables (i.e.,  $T$ ) and the number of LSH functions in each LSH table (i.e.,  $r$ ) are two key parameters in our proposal. Therefore, in this profile, we test the



(a) MAE w.r.t.  $T$



(b) MAE w.r.t.  $r$

**Fig. 6.** Recommendation accuracy of  $DistSR_{Amplify-LSH}$  w.r.t.  $T$  and  $r$

recommendation accuracy of  $DistSR_{Amplify-LSH}$  with respect to  $T$  and  $r$ . The parameters are set as follows:  $m = 200$ ,  $n = 3000$ . The concrete experiment results are shown in Fig. 6.

In Fig. 6(a),  $T = \{5, 6, 7, 8, 9\}$ ,  $r = \{2, 4, 6, 8, 10\}$ . For each  $T$  value, we test the  $MAE$  values corresponding to different  $r$  values, and finally the average  $MAE$  value is adopted. The experiment results indicate that the recommendation accuracy increases (i.e.,  $MAE$  decreases) slightly when  $T$  grows. This is because in *Strategy-2*, the AND operation is taken over the generated  $T$  LSH tables; therefore, more LSH tables often mean a stricter search condition for neighbors as well as higher service recommendation accuracy.

In Fig. 6(b),  $T = \{2, 4, 6, 8, 10\}$ ,  $r = \{5, 6, 7, 8, 9\}$ . For each  $r$  value, we test the  $MAE$  values corresponding to different  $T$  values, and finally the average  $MAE$  value is adopted. The experiment results show that the recommendation accuracy decreases (i.e.,  $MAE$  increases) with the growth of  $r$  approximately. This is because in *Strategy-1*, the OR operation is taken over the chosen  $r$  LSH functions; as a consequence, more LSH functions often mean looser search condition for neighbors as well as lower recommendation accuracy.

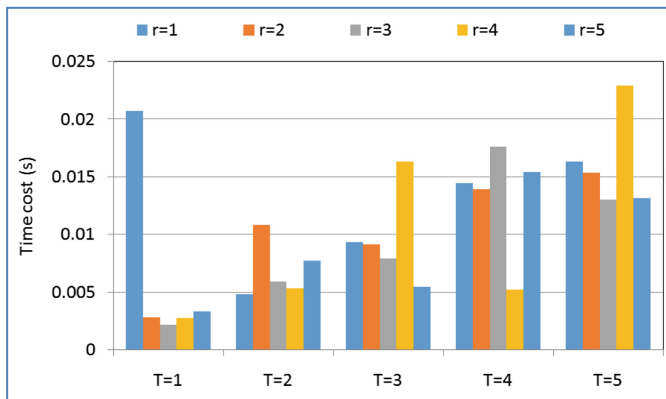
**Profile 4: Recommendation Efficiency of  $DistSR_{Amplify-LSH}$  with Respect to  $T$  and  $r$**

In this profile, we test the efficiency of  $DistSR_{Amplify-LSH}$  with respect to  $T$  and  $r$ . Here,  $m = 200$ ,  $n = 3000$ ,  $T$  is varied from 1 to 5,  $r$  is varied from 1 to 5. The experiment results are shown in Fig. 7. As Fig. 7(a) shows, the time cost of our proposal increases approximately with the growth of  $T$  because of the AND operation (over  $T$  LSH tables) adopted in *Strategy-2*. While as can be seen from Fig. 7(b), the time cost of our proposal does not exhibit a very regular variation tendency when  $r$  grows. Typically, the average time cost (i.e., the blue column) stays approximately the same with the growth of  $r$  because of the OR operation (over  $r$  LSH functions) adopted in *Strategy-1*.

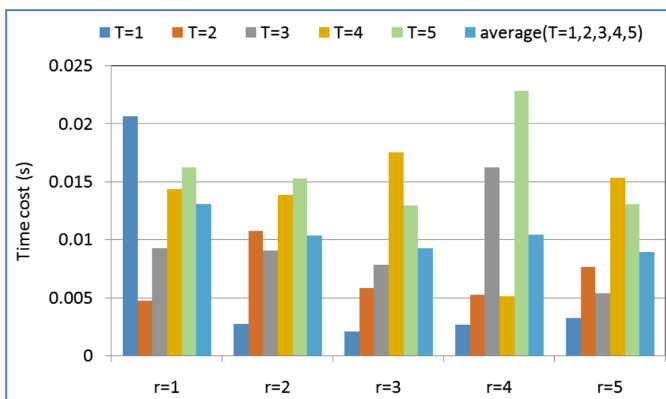
**Profile 5: Recommendation Failure Rate of  $DistSR_{Amplify-LSH}$  with Respect to  $T$**

The number of LSH tables, i.e.,  $T$  plays an important role in the successful service recommendation because of the adopted AND operation in *Strategy-2*. Generally, for a target user, more LSH tables mean a stricter search condition for similar neighbors. Therefore, if  $T$  is large enough, the search condition may become too strict to find a qualified neighbor of the target user; in this situation, service recommendation is failed. In this profile, we test the failure rate of  $DistSR_{Amplify-LSH}$  with respect to  $T$ . The parameters are set as follows:  $m = 200$ ,  $n = 3000$ ,  $r = 1$ ,  $T$  is varied from 5 to 15. Concrete experiment results are presented in Fig. 8.

As Fig. 8 shows, the failure rate of our  $DistSR_{Amplify-LSH}$  approach increases with the growth of  $T$  approximately; this is because more LSH tables (i.e., a larger  $T$  value) often mean a stricter search condition for similar neighbors and hence may lead to a recommendation failure. While when  $T$  is small (e.g., when  $T = 5$  or  $T = 6$ ), the failure rate of  $DistSR_{Amplify-LSH}$  drops to a small value even 0. Therefore, through tuning the parameter value of  $T$ , a low recommendation failure rate can be guaranteed.

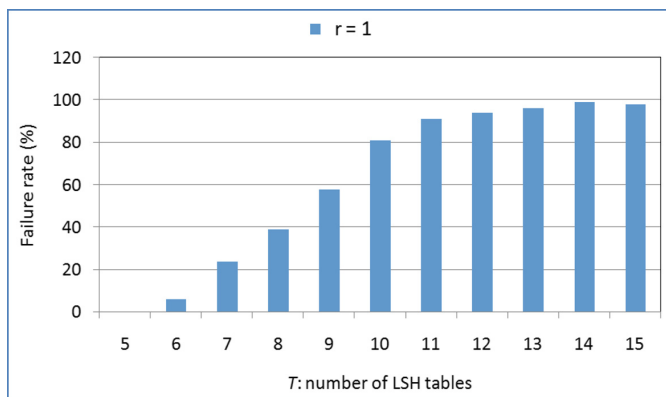


(a) Time cost w.r.t.  $T$



(b) Time cost w.r.t.  $r$

**Fig. 7.** Recommendation efficiency of  $DistSR_{Amplify-LSH}$  w.r.t.  $T$  and  $r$  (Color figure online)



**Fig. 8.** Recommendation failure rate of  $DistSR_{Amplify-LSH}$  w.r.t.  $T$



## 6 Related Work and Comparison Analyses

In this section, we compare our proposal with existing CF-based recommendation approaches. Generally, CF-based service recommendation approaches can be divided into two categories: model-based CF and memory-based CF.

### 6.1 Model-Based CF

Model-based CF approaches recruit the historical service quality data observed by users to build a user-service recommendation model offline, and then make online recommendation based on the obtained model. Several model-based CF approaches are introduced for service recommendation, e.g., *Matrix Factorization* [16], *Latent Dirichlet Allocation* [17] and *clustering* [18]. Generally, the recommendation efficiency of these model-based CF approaches is high as the recommendation model can be trained offline. However, these approaches have two shortcomings. First, they seldom consider the privacy leakage risks in service recommendation process. Second, they often assume that the historical service quality data used for service recommendation are centralized, without considering the distributed service recommendation scenarios where historical service quality data are from multiple independent platforms; therefore, these approaches fall short in handling the distributed service recommendation problems.

### 6.2 Memory-Based CF

Memory-based CF recommendation approaches first employ the historical service quality data to find similar users or similar services, and then utilize them to make service recommendations. User-based CF and item-based CF are employed for service recommendation in [19] and [20], respectively. In order to combine the advantages of user-based CF and item-based CF, a hybrid service recommendation approach named *WSRec* is proposed in [21]. Experiment results show that the recommendation performance is improved. As the running quality of a web service often depends on the service invocation context (e.g., *service invocation time*, *service location*, *user location*), time-aware recommendation and location-aware recommendation are put forward by [22] and [23], respectively. Besides, user preferences also play an important role in the service selection decisions of a target user; in view of this observation, preference-aware CF recommendation approach is introduced in [24], where the users with similar preferences are recruited to make service recommendations.

However, the above approaches do not consider the privacy leakage risks in service recommendation process. In view of this shortcoming, in [25], users are suggested to release only a small portion of their observed service quality data to the public so that the remaining majority of data are still secure. However, the released small portion of data can still reveal part of a user's private information. In order to protect the private information of users better, data obfuscation strategy is adopted in [13] to hide the real service quality data. However, as the service quality data used to make service recommendations have been obfuscated, the recommendation accuracy is reduced accordingly. In view of this drawback, a "divide-merge" strategy is put forward in [14]

where each piece of service quality data is divided into several segments with little private information, and then the segments are recruited for user similarity calculation and service recommendations. However, two shortcomings are present in [14]. First, the recommendation efficiency is decreased severely as the adopted “divide-merge” operations bring additional time cost. Besides, some private information of users cannot be protected well, e.g., *the service intersection commonly invoked by two users*.

In view of the shortcomings of existing research work, a novel amplified LSH-based recommendation approach named *DistSR<sub>Amplify-LSH</sub>* is proposed in this paper, to solve the privacy-preserving distributed service recommendation problems in the distributed environment. Through the extensive experiments conducted on a real distributed service quality dataset *WS-DREAM*, we validate the feasibility and advantages of our proposal in terms of recommendation accuracy and efficiency while guaranteeing privacy-preservation.

## 7 Conclusions

In this paper, we put forward a novel privacy-preserving service recommendation approach based on amplified Locality-Sensitive Hashing, i.e., *DistSR<sub>Amplify-LSH</sub>*, to handle the distributed service recommendation problems. Through Locality-Sensitive Hashing, user indexes can be built offline; as a consequence, the neighbor search efficiency and service recommendation efficiency are improved significantly. Besides, due to the inherent nature of LSH, user privacy can be protected during the distributed service recommendation process. Moreover, we amplify LSH by integrating the AND/OR operations over multiple LSH tables or LSH functions, through which the “False-positive” and “False-negative” recommended results are reduced; as a result, the service recommendation accuracy is improved significantly. Finally, through a set of experiments deployed on the distributed service quality dataset *WS-DREAM*, we validate the feasibility of our proposed *DistSR<sub>Amplify-LSH</sub>* approach. Experiment results demonstrate that our proposal can achieve a good recommendation performance in terms of recommendation accuracy and efficiency while guaranteeing privacy-preservation.

In the future, we will further investigate the distributed service recommendation problems with multiple quality dimensions. Besides, the running qualities of web services are often not static, but dynamic; therefore, we will study the dynamic quality-aware distributed service recommendation problems in the future.

**Acknowledgements.** This paper is partially supported by the National Key Research and Development Program of China (No. 2017YFB1001800), Natural Science Foundation of China (Nos. 61402258, 61672276), UoA Faculty Research Development Fund (No. 3714668), Open Project of State Key Laboratory for Novel Software Technology (No. KFKT2016B22).

## References

1. Blake, M.B., Saleh, I., Wei, Y., Schlesinger, I.D., Yale-Loehr, A., Liu, X.: Shared service recommendations from requirement specifications: a hybrid syntactic and semantic toolkit. *Inf. Softw. Technol.* **57**, 392–404 (2015)

2. Al-Hassan, M., Haiyan, L., Jie, L.: A semantic enhanced hybrid recommendation approach: a case study of e-Government tourism service recommendation system. *Decis. Support Syst.* **72**, 97–109 (2015)
3. Segev, A., Sheng, Q.: Bootstrapping ontologies for web services. *IEEE Trans. Serv. Comput.* **5**(1), 33–44 (2012)
4. Cao, G., Kuang, L.: Identifying core users based on trust relationships and interest similarity in recommender system. In: *IEEE International Conference on Web Services*, pp. 284–291 (2016)
5. Zhong, Y., Fan, Y., Tan, W., Zhang, J.: Web service recommendation with reconstructed profile from mashup descriptions. *IEEE Trans. Autom. Sci. Eng.* (2016)
6. Mashal, I., Chung, T.-Y., Osama, O.: Toward service recommendation in internet of things. In: *IEEE International Conference on Ubiquitous and Future Networks*, pp. 328–331 (2015)
7. Zheng, Z., Ma, H., Lyu, M.R., King, I.: QoS-aware web service recommendation by collaborative filtering. *IEEE Trans. Serv. Comput.* **4**(2), 140–152 (2011)
8. Gionis, A., Indyk, P., Motwani, R.: Similarity search in high dimensions via hashing. *VLDB* **99**(6), 518–529 (1999)
9. Lee Rodgers, J., Nicewander, W.A.: Thirteen ways to look at the correlation coefficient. *Am. Stat.* **42**(1), 59–66 (1988)
10. *Data Mining and Query Log Analysis for Scalable Temporal and Continuous Query Answering* (2015). <http://www.optique-project.eu/>
11. Zheng, Z., Zhang, Y., Lyu, M.R.: Investigating QoS of real world web services. *IEEE Trans. Serv. Comput.* **7**(1), 32–39 (2014)
12. Breese, J.S., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: *International Conference on Uncertainty in Artificial Intelligence*, pp. 43–52 (1998)
13. Zhu, J., He, P., Zheng, Z., Lyu, M.R.: A privacy-preserving QoS prediction framework for web service recommendation. In: *IEEE International Conference on Web Services*, pp. 241–248 (2015)
14. Li, D., Chen, C., Lv, Q., Shang, L., Zhao, Y., Lu, T., Gu, N.: An algorithm for efficient privacy-preserving item-based collaborative filtering. *Future Gener. Comput. Syst.* **55**, 311–320 (2016)
15. Slaney, M., Casey, M.: Locality-sensitive hashing for finding nearest neighbors. *IEEE Sig. Process. Mag.* **25**(2), 128–131 (2008)
16. Yao, L., Sheng, Q.Z., Qin, Y., Wang, X., Shemshadi, A., He, Q.: Context-aware point-of-interest recommendation using tensor factorization with social regularization. In: *International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1007–1010 (2015)
17. Zhong, Y., Fan, Y., Huang, K., Tan, W., Zhang, J.: Time-aware service recommendation for mashup creation in an evolving service ecosystem. In: *IEEE International Conference on Web Services*, pp. 25–32 (2014)
18. Wu, C., Qiu, W., Zheng, A., Wang, X., Yang, X.: QoS prediction of web services based on two-phase k-means clustering. In: *IEEE International Conference on Web Services*, pp. 161–168 (2015)
19. Rong, H., Huo, S., Hu, C., Mo, J.: User similarity-based collaborative filtering recommendation algorithm. *J. Commun.* **35**(2), 16–24 (2014)
20. Chung, K.-Y., Lee, D., Kim, K.J.: Categorization for grouping associative items using data mining in item-based collaborative filtering. *Multimed. Tools Appl.* **71**(2), 889–904 (2014)
21. Jiang, C., Duan, R., Jain, H.K., Liu, S., Liang, K.: Hybrid collaborative filtering for high-involvement products: a solution to opinion sparsity and dynamics. *Decis. Support Syst.* **79**, 195–208 (2015)

22. Wang, X., Zhu, J., Zheng, Z., Song, W., Shen, Y., Lyu, M.R.: A spatial-temporal QoS prediction approach for time-aware web service recommendation. *ACM Trans. Web* **10**(1), 7 (2016)
23. Yu, C., Huang, L.: A web service QoS prediction approach based on time- and location-aware collaborative filtering. *Serv. Oriented Comput. Appl.* **10**(2), 135–149 (2016)
24. Fletcher, K.K., Liu, X.F.: A collaborative filtering method for personalized preference-based service recommendation. In: *IEEE International Conference on Web Services*, pp. 400–407 (2015)
25. Dou, W., Zhang, X., Liu, J., Chen, J.: HireSomeII: towards privacy-aware cross-cloud service composition for big data applications. *IEEE Trans. Parallel Distrib. Syst.* **26**(2), 455–466 (2015)

# Learn to Accelerate Identifying New Test Cases in Fuzzing

Weiwei Gong<sup>(✉)</sup>, Gen Zhang, and Xu Zhou

School of Computer, National University of Defense Technology,  
Changsha 410073, China

IssacGong@outlook.com, zhanggen12@hotmail.com, zhouxu@nudt.edu.cn

**Abstract.** Fuzzing is an efficient testing technique to catch bugs early, before they turn into vulnerabilities. Without complex program analysis, it can generate interesting test cases by slightly changing input and find potential bugs in programs. However, previous fuzzers either are unable to explore deeper bugs, or some of them suffer from dramatic time complexity, thus we cannot depend on them in real world applications. In this paper, we focus on reducing time complexity in fuzzing by combining practical and light-weight deep learning methods, which fundamentally accelerate the process of identifying new test cases and finding bugs. In order to achieve expected fuzzing coverage, we implement our method by extending state-of-the-art fuzzer AFL with deep learning methods and evaluate it on several wide-used and open source executable programs. On all of these programs, efficiency of our method is witnessed and significantly better outcomes are generated.

**Keywords:** Accelerate · New test case · Fuzzing · Deep learning Security

## 1 Introduction

Security of networks and softwares is attracting increasing attention these days. Despite efforts to increase the resilience of software against security flaws, vulnerabilities in software are still commonplace [1]. During the past few decades, security specialists and researchers spared no effort in finding vulnerabilities and fixing them. However, many classes of vulnerabilities, such as functional correctness bugs, are difficult to find without executing a piece of code [2]. With regard to the problem of code executing, there has been much debate about the efficiency of symbolic execution versus more lightweight fuzzers [3]. Symbolic execution tools, such as KLEE, EXE and DART, are capable of automatically generating tests that achieve high coverage on a diverse set of complex and environmentally-intensive programs [2, 4, 5]. While fuzzing is the process of finding security vulnerabilities in input-parsing code by repeatedly testing the parser with modified, or fuzzed, inputs [6].

Symbolic execution is very effective because each test case typically execute the target program along a certain path [7]. However, this effectiveness comes at the cost of spending significant time doing program analysis and constraint solving. It triggers a large number of paths in the target program and will result in path explosion [8]. However, with regard to fuzzing, today most vulnerabilities are exposed by particularly lightweight fuzzers that do not leverage any program analysis [9]. It turns out that even the most effective symbolic execution is less efficient than fuzzing if the time spent generating a test case takes relatively too long [10]. For the above reasons, in this paper, we give up classic symbolic execution and focus on extending a stage-of-the-art fuzzer American Fuzzing Lop (AFL) [11].

There are three main types of fuzzing techniques in use [12]: black-box random fuzzing [13], white-box constraint-based [14] fuzzing and grey-box fuzzing [15]. Black-box fuzzing is a technique of software testing without any knowledge of the internal architecture of the target program. It only examines the fundamental aspects of the system, which treats the software as a Black Box [16]. White-box fuzzing is based on analysis of internal structure of the target program and is very effective and efficient in validating design and assumptions. White-box fuzzing is performed based on the knowledge of how the system is implemented [17]. Grey-box fuzzing tests the program with limited knowledge of the structure of an application. It provides combined benefits of black-box and white-box fuzzing techniques and the tester can design excellent test scenarios [15].

However, since grey-box fuzzing is efficient for real world program and its combined benefits, it is widely applied in software testing. In the range of grey-box fuzzing, there have been several well-known fuzzers, such as automatic generation [18], fuzz [19] and semantic model [20]. Although these methods do give up the time-consuming program analysis, they suffer from low testing accuracy: either take a non-crash point as a crash or ignore the real significant crash. Furthermore, existing grey-box fuzzers have been effective mainly in discovering superficial bugs, close to the surface of software, while struggling with more complex ones [21,22]. On top of all these reasons, we focus on extending a state-of-the-art grey-box fuzzer AFL, which employs evolutionary algorithms to operate valid input generation and a simple feedback loop to assess how good an input is [23]. And in previous works, AFL is proved to have high accuracy and able to reveal deeper bugs in programs [3,23].

AFL is a brute-force fuzzer coupled with an exceedingly simple but rock-solid instrumentation-guided genetic algorithm. However, it would take hours and days to find a unique crash with AFL tool. And the main drawback of AFL is that it actually runs the target program for once with the input to decide whether there is a state transition or new state: in other words, the process of identifying new test cases. When the target program is complex and contains millions of parsing or condition code, running the program for once would take considerable time and the whole process of AFL would take hours and days to find a unique crash (we will discuss this in detail in Sect. 2). To tackle this

problem, in this paper, we integrate efficient deep learning methods such as neural networks into AFL to accelerate the process of identifying new tesecases and finding bugs. On all of our experiments, efficiency of our method is witnessed and significantly better outcomes are generated.

The main contribution of this paper is as follows: (1) we apply grey-box fuzzing, which is efficient and easy to implementation; (2) we extend AFL with deep learning methods, which significantly accelerate the process of identifying new tesecases and finding bugs.

## 2 Background

### 2.1 American Fuzzy Lop

American Fuzzy Lop (AFL) is a brute-force fuzzer coupled with an exceedingly simple but rock-solid instrumentation-guided genetic algorithm. It uses a modified form of edge coverage to effortlessly pick up subtle, local-scale changes to program control flow [24].

The overall algorithm can be summed up as:

- (1) Load user-supplied initial test cases into the queue,
- (2) Take next input file from the queue,
- (3) Attempt to trim the test case to the smallest size that doesn't alter the measured behavior of the program,
- (4) Repeatedly mutate the file using a balanced and well-researched variety of traditional fuzzing strategies,
- (5) If any of the generated mutations resulted in a new state transition recorded by the instrumentation, add mutated output as a new entry in the queue,
- (6) Go to 2.

American Fuzzy Lop does its best not to focus on any singular principle of operation and not be a proof-of-concept for any specific theory. The tool can be thought of as a collection of hacks that have been tested in practice, found to be surprisingly effective, and have been implemented in the simplest, most robust way people could think of at the time [9]. Thus extension on AFL can give us a high level platform to start with.

More specifically, when we look deeply into the running process of AFL, we will figure out there are several procedures that are time consuming, which we actually aim at to make improvement. As described above, AFL takes initial test cases into the queue, trim it down, makes mutations, and actually runs the target program for once with the input to decide weather this test case will cause state transition or new state. If this test case do cause state transition or new state, we can identify it as a new test case. When the target program is complex and contains millions of parsing or condition code, running the program for once would take considerable time and the whole process of AFL would take hours and days to find a unique crash. So we extend AFL with deep learning techniques, without actually running the target program, to accelerate the whole process of identifying new test cases and finding bugs, which is significant improvement of original version of AFL.

## 2.2 Deep Learning

Machine learning and deep learning is under spot light these days, attracting thousands of researchers to work hard on it. Deep and recurrent neural networks (DNNs and RNNs, respectively) are powerful models that achieve high performance on difficult pattern recognition problems in vision, and speech [25]. Deep Learning in Neural Networks (NNs) is relevant for Supervised Learning (SL), Unsupervised Learning (UL), and Reinforcement Learning (RL) [26]. Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction [27].

Deep learning comes with a lot brand new characteristics and the combination of software testing and deep learning would make sense. In this paper, we apply deep learning techniques to improve the performance of AFL and accelerate the process of new test case identification and bug finding, which is a brand-new and interesting research field (to the best of our knowledge so far).

## 3 Model Overview

In this section, we will describe our proposed model in detail and the whole process of our method will be presented.

### 3.1 Train the AFL Results with Deep Learning Methods

As described in Sect. 2, when the target program is complex and contains millions of parsing or condition code, running the program for once would take considerable time and the whole process of AFL would take hours and days to find a unique crash. So we focus on eliminating or cutting down these procedures using deep learning techniques.

In AFL, one test case is mutated with certain fuzzing techniques, such as flipping a certain bit, adding a certain integer, replacing with a certain number and so on. In total, there are 16 kinds of fuzzing techniques in AFL: flip2, flip4, flip8, flip16, flip32, arith8, arith16, arith32, int8, int16, int32, ext-UO, ext-UI, ext-AO, havoc and splicing. After one test case is mutated by one of the above techniques, it will be thrown into the target program to run for once to see if there will be a new state of program or state transition. If there is something new, the mutated test case will be identified as a new test case and saved for later use.

More specifically, the content of one test case can always be translated into a sequence of 0s and 1s. Likewise, the fuzzing techniques can also be seen as a 4-bit binary sequence. For example, “0000” stands for “flip2” and “1111” stands for “splicing”. And the position and value of mutation can be translated in the same way. At last, whether this is a new test case is simply 1 and 0. Table 1 illustrates how we translate the original input of AFL problem into a binary sequence. For example, we have a string test case “3”, and fuzzing technique is “int8”, which will add an 8-bit integer to test case. Moreover, the 8-bit integer



is 4 and the position is just the 0 bit. This mutated test case will cause state transition in AFL so it is a NEW test case. Thus we can translate this situation into our defined binary sequence. “00...0011” for test case, “1000” for fuzzing technique, “00...000” for mutation position, “00...00100” for mutation value and “1” for new.

**Table 1.** Transition to binary sequence

32-bit	4-bit	10-bit	32-bit	1-bit
Test case	Fuzzing techniques	Mutation position	Mutation value	New

After we run AFL for a certain time, the data we collect from the process of AFL will be divided into training data and testing data according to a certain proportion, which will be discussed in detail in Sect. 4. After translating all the training data into binary sequence, the preceding 78-bit can be taken as input and the last 1-bit can be seen as tag. And this binary sequence can perfectly fit into a neural network model, which is a widely-used division of deep learning techniques. A neural network model take a sequence of 1s and 0s as input and simply out put 1 or 0 after calculating in networks. We will use this process to predict weather a mutated test case is a new test case in AFL. More specifically, for training the network,  $78 + 1$  bits are needed as training data and as for testing, 78 bits are taken as input, while the last 1 bit is what we need the network to calculate.

### 3.2 Integrate the Learning Results into AFL

As discussed above, we will eliminate or cut down the time-consuming new test case identifying procedure of AFL by integrating the learning results into AFL. The original version will run the program with the mutated test case, but our new method will simply throw the translated binary sequence into a neural network model and predict weather this test case will result in a new state. To be more specific, when AFL pulls out a test case to make mutation on it and at this time we manually stop further executing on this test case. On the contrary, we put this test case into the neural network in our defined form in Table 1, and make prediction on weather this is a new test case. The result will be sent back to AFL for further running: if this test case is new and will cause state transition, it will be saved for later use; if not, discarded. One thing need to be declared is that: although train the neural network model will take some time, but we only have to train once for one target program. And time complexity of predicting the result in a simple neural network is only  $O(n_1 * n_2 + n_2 * n_3 + \dots)$  [28]. So when the number of nodes is in a small range, such as in our model, the predicting process takes much less time than actually running the target program for once. As a result, our method can accelerate the process of identifying new test case and finding bugs of the original version of AFL.

So the whole process of our method of extending AFL runs as follows:

- (1) Run AFL for a certain amount of time to collect data we need for further training. (One thing has to be made clear, compared to running AFL for hours and days, the time spent in this procedure is much less.)
- (2) Train the data we collect in our neural network.
- (3) Go back to the AFL with the training results to make predictions for each test case, and finally find potential bugs in programs.

Algorithm 1 also illustrates the process of our extended AFL:

---

**Algorithm 1.** Our proposed method of extending AFL

---

**Input:**

$\mathcal{T}_0$ : set of user initial test cases

$P$ : target program

$t$ : fixed time of collecting data

$k$ : k-fold division of data

$NN$ : Neural networks with certain parameters

$T'$ : a mutated test case waiting to be identified

**Output:**

$NEW$ : whether  $T'$  is a new test case and cause state transition

**Procedures:**

- 1: Run AFL for time  $t$  with  $\mathcal{T}_0$  as input in  $P$ , to collect data  $D$  we need for further training
  - 2: Divide  $D$  into training data  $D_{train}$  and testing data  $D_{test}$  with  $k$ -fold
  - 3: Train the data  $D_{train}$  we collect in the neural networks  $NN$  and we get results  $R$
  - 4: Go back to AFL with the training results  $R$  to make predictions for test case  $T'$
  - 5: return  $NEW$
- 

## 4 Experiments

In this section, we will present the experiment results of our method and discuss the improvement over the original version of AFL. We ran our experiments on an Ubuntu 14.04 LTS system equipped with a 64-bit 4-core Intel CPU and 32 GB RAM. For our experiment, we select 8 target programs: bmp2tiff, pal2rgb, tiff2pdf, tiff2ps, gif2png, readelf, nm-new and cxxfix [29–31], which are all widely-used in their areas and also have been tested in previous papers [3, 23].

### 4.1 Training Results of Neural Networks

For each of our target programs, we run AFL with it for a certain time and collect over 500K test cases for later training process. And all of them are stored in binary sequence as shown in Table 1. And the neural network we apply is `sklearn.neural-network.MLPClassifier` [32], which is a widely-used multi-layer perceptron classifier. The hidden layer size is (5, 2) and the solver is “adam”.

Though there are many types of network algorithms, taking time and other factors into consideration, in this paper we only introduce a single neural network to accomplish our experiment. We take training data as input for our neural network and in an acceptably short time it will calculate the results. We obtain training results and they are illustrated in Table 2.

The first column “Target” are the 8 target programs we experiment on. The next column “Total” is the number of test cases we collect for each target program. Then we divide all the collected test case into training and testing data with 5-fold, which means training data is  $\frac{4}{5}$  of the total test case and testing data is  $\frac{1}{5}$  accordingly. After fitting training data into our neural networks, we can predict a new coming test case. And the third column “Error” is the number of wrong predictions and the last column is predicting accuracy of our neural networks. As Table 2 illustrates, our predicting accuracy is between around 85% to 90%. With such high accuracy, our proposed method of identifying new test cases can accurately predict weather a mutated test case will cause state transition and perform almost the same as the original version in accuracy, and perform much better in speed.

**Table 2.** Training results of neural networks

Target	bmp2tiff	pal2rgb	tiff2pdf	tiff2ps	gif2png	readelf	nm-new	cxxfix
Total	718392	1409466	2797254	1282716	1395384	1768752	566442	543078
Error	12320	32415	48820	17220	28880	29175	11814	9780
Accuracy	89.71%	86.20%	89.53%	91.95%	87.58%	90.10%	87.49%	89.19%

## 4.2 Time Performance of Our Method

As described above, after training our neural networks with the data we collect, it is time to predict weather a mutated test case will cause state transition in the process of AFL. And the high accuracy of our prediction is definitely a solid basis for later operation. Again, we run the 8 target programs for a certain time, and the only difference compared to the original AFL is that a test case doesn’t have to go through the target program to decide weather it will be saved or discarded. And the only thing need to be done is to throw this test case to our already trained neural networks to make predictions. Table 3 illustrates comparison with the original version of AFL.

The first column are the 8 target programs as described above. And the second column is the number of crashes and hangs AFL and our method find in a certain time. So the comparison between AFL and our method is with regard to the time of finding a given number of crashes and hangs. The next two columns are the time AFL and our method need to find those crashes and hangs respectively. (All of our execution time is in accordance with the work of Böhme et al. [3] and Rawat et al. [23].) As presented in the table, our method do lead the original AFL in execution time and on average at 5% in advance. As detailed

**Table 3.** Time (in minutes) performance of our method

Target	bmp2tiff	pal2rgb	tiff2pdf	tiff2ps	gif2png	readelf	nm-new	cxxfix
Crash/Hang	27/47	39/12	64/8	48/22	17/11	0/3	0/2	0/1
Before	12.03	14.08	21.05	11.05	9.03	167.03	20.10	25.80
After	11.51	13.40	19.82	10.60	8.68	156.84	18.91	24.25
Improvement	4.41%	4.83%	5.94%	4.07%	3.88%	6.10%	5.90%	5.99%

in our theoretical analysis in the previous section, our method does NOT have to actually run the target program and it only depend on our neural networks to make predictions to see weather a test case will cause state transition. Thus both theoretical analysis and experiments show our method does accelerate AFL in the process of identifying new test cases and finding bugs.

## 5 Conclusion

Facing the severe security issue nowadays, in our proposed method, we first select fuzzing over complicated symbolic execution to focus on. Additionally, in all the implementation of fuzzing, we concentrate on state-of-the-art AFL to make extension for its promised characteristics. In order to accelerate the process of AFL and identifying new test cases, in this paper, we focus on reducing time complexity in fuzzing by combining practical and light-weight deep learning methods. In order to achieve expected fuzzing coverage, we implement our method by extending AFL with deep learning methods and evaluate it on 8 wide-used and open source target programs. On all of these programs, our method to extend AFL shows prominent improvement over the original version. However, our method does not focus on improving code coverage to find more bugs. Thus in later work, we will try to make research on these fields. Moreover, though there are many types of network algorithms, taking time and other factors into consideration, in this paper we only introduce a single neural network to accomplish our experiment. In the future we would use different neural networks on different situations to evaluate the performance of different neural networks.

**Acknowledgments.** The work is supported by The National Key Research and Development Program of China (No. 2016YFB0200401).

## References

1. Stephens, N., Grosen, J., Salls, C., Dutcher, A., Wang, R., Corbetta, J., Shoshitaishvili, Y., Kruegel, C., Vigna, G.: Driller: augmenting fuzzing through selective symbolic execution. In: Proceedings of the Network and Distributed System Security Symposium (2016)
2. Cadar, C., Dunbar, D., Engler, D.R., et al.: KLEE: unassisted and automatic generation of high-coverage tests for complex systems programs. In: OSDI, vol. 8, pp. 209–224 (2008)

3. Böhme, M., Pham, V.-T., Roychoudhury, A.: Coverage-based greybox fuzzing as markov chain. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 1032–1043. ACM (2016)
4. Cadar, C., Ganesh, V., Pawlowski, P.M., Dill, D.L., Engler, D.R.: EXE: automatically generating inputs of death. *ACM Trans. Inf. Syst. Secur. (TISSEC)* **12**(2), 10 (2008)
5. Godefroid, P.: Compositional dynamic test generation. In: *ACM SIGPLAN Notices*, vol. 42, pp. 47–54. ACM (2007)
6. Godefroid, P., Peleg, H., Singh, R.: Learn&fuzz: machine learning for input fuzzing. arXiv preprint [arXiv:1701.07232](https://arxiv.org/abs/1701.07232) (2017)
7. Zalewski, M.: Symbolic execution in vulnerability research (2016)
8. Boonstoppel, P., Cadar, C., Engler, D.: RWset: attacking path explosion in constraint-based test generation. In: Ramakrishnan, C.R., Rehof, J. (eds.) *TACAS 2008*. LNCS, vol. 4963, pp. 351–366. Springer, Heidelberg (2008). [https://doi.org/10.1007/978-3-540-78800-3\\_27](https://doi.org/10.1007/978-3-540-78800-3_27)
9. Zalewski, M.: American fuzzy lop (AFL) fuzzer-technical details (2016)
10. Böhme, M., Paul, S.: A probabilistic analysis of the efficiency of automated software testing. *IEEE Trans. Softw. Eng.* **42**(4), 345–360 (2016)
11. Zalewski, M.: American fuzzy lop (AFL) fuzzer (2016)
12. Khan, M.E.: Different forms of software testing techniques for finding errors. *Int. J. Comput. Sci. Issues* **7**(3), 11–16 (2010)
13. Sutton, M., Greene, A., Amini, P.: *Fuzzing: Brute Force Vulnerability Discovery*. Pearson Education, London (2007)
14. Godefroid, P., Levin, M.Y., Molnar, D.A., et al.: Automated whitebox fuzz testing. In: *NDSS*, vol. 8, pp. 151–166 (2008)
15. Khan, M.E., Khan, F., et al.: A comparative study of white box, black box and grey box testing techniques. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, **3**(6) (2012)
16. Khan, M.E.: Different approaches to black box testing technique for finding errors. *Int. J. Softw. Eng. Appl.* **2**(4), 31 (2011)
17. Khan, M.E.: Different approaches to white box testing technique for finding errors (2011)
18. Bird, D.L., Munoz, C.U.: Automatic generation of random self-checking test cases. *IBM Syst. J.* **22**(3), 229–245 (1983)
19. Forrester, J.E., Miller, B.P.: An empirical study of the robustness of windows NT applications using random testing. In: *Proceedings of the 4th USENIX Windows System Symposium*, Seattle, pp. 59–68 (2000)
20. Offutt, A.J., Hayes, J.H.: A semantic model of program faults. In: *ACM SIGSOFT Software Engineering Notes*, vol. 21, pp. 195–200. ACM (1996)
21. Clark, S., Frei, S., Blaze, M., Smith, J.: Familiarity breeds contempt: the honeymoon effect and the role of legacy code in zero-day vulnerabilities. In: *Proceedings of the 26th Annual Computer Security Applications Conference*, pp. 251–260. ACM (2010)
22. Dolan-Gavitt, B., Hulin, P., Kirda, E., Leek, T., Mambretti, A., Robertson, W., Ulrich, F., Whelan, R.: LAVA: large-scale automated vulnerability addition. In: *2016 IEEE Symposium on Security and Privacy (SP)*, pp. 110–121. IEEE (2016)
23. Rawat, S., Jain, V., Kumar, A., Cojocar, L., Giuffrida, C., Bos, H.: Vuzzer: application-aware evolutionary fuzzing (2017)
24. Zalewski, M.: American fuzzy lop (AFL) fuzzer-readme (2016)
25. Sutskever, I., Martens, J., Dahl, G.E., Hinton, G.E.: On the importance of initialization and momentum in deep learning. *ICML* **3**(28), 1139–1147 (2013)

26. Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015)
27. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436–444 (2015)
28. Orponen, P.: Neural networks and complexity theory. *Math. Found. Comput. Sci.* **1992**, 50–61 (1992)
29. Welles, M., Warmerdam, F., Kiselev, A., Kelly, D.: Tag image file format. <http://www.libtiff.org/>
30. GNU. GNU binutils. <http://www.gnu.org/software/binutils/>
31. Raymond, E.S.: GIFs to PNGs. <http://www.catb.org/esr/gif2png/>
32. sklearn. sklearn.neural\_network.mlpclassifier. [http://scikit-learn.org/stable/modules/generated/sklearn.neural\\_network.MLPClassifier.html](http://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html)

# Service Selection Based on User Privacy Risk Evaluation

Mingdong Tang<sup>1,2(✉)</sup>, Sumeng Zeng<sup>2</sup>, Jianxun Liu<sup>2</sup>,  
and Buqing Cao<sup>2</sup>

<sup>1</sup> School of Information Science and Technology,  
Guangdong University of Foreign Studies, Guangzhou 510006, China  
mdtang@126.com

<sup>2</sup> School of Computer Science and Engineering,  
Hunan University of Science and Technology, Xiangtan 411201, China

**Abstract.** Requesting a service on the Internet may require the user's privacy data, and thus raising the risk of the user's privacy leakage and violation. Hence, it is necessary for users to select services that protect their privacy information. However, previous studies on service selection usually focused only on the quality of service, seldom had they considered the user's privacy concern. As such, their results may be unable to meet the user's privacy protection requirement. Aiming at reducing the privacy risk of users in service selection, this paper proposes a fuzzy logic service selection approach. The approach uses a fuzzy model to allow a service user specifying personalized privacy preference and a service provider specifying flexible privacy requirements; then it leverages the service's reputation, privacy policy and the user's privacy preference to compute the privacy risk for each service candidate; finally, it ranks all service candidates based on their privacy risk degrees. Examples and evaluations show that the proposed approach is effective and efficient for reducing privacy risk in service selection.

**Keywords:** Service selection · Service ranking · Privacy protection  
Reputation · Fuzzy logic

## 1 Introduction

The Internet has become a platform for numerous services provision and consumption, varied from e-commerce services, e-government services, social networking services, Web services to cloud services, which have brought great convenience to Internet users. With the large amount of services, how to choose one that most satisfies the user's personal requirements is a challenging issue. To attack this issue, dozens of research studies have been done on service selection. Given a set of service candidates satisfying the user's functional requirements, previous service selection approaches focused on evaluating the quality of service candidates and used the evaluation results to rank services [1–4]. The quality of services (QoS) on the Internet is typically referred to as a group of criteria like availability, reliability, latency, cost, reputation, etc. Though QoS-based service selection approaches has received the most attention,

seldom did they take the user's privacy concern into consideration, which, however, has become one of the most important issues of the information age [5].

As a matter of fact, many services on the Internet require users to supply personal or private information, such as name, telephone number, credit card number, email and location, to benefit their performance. For example, an online meal ordering and delivery service may ask a user to supply exact location and phone number, so that it can recommend the nearest restaurants to the user and deliver the food successfully as soon as possible. Once the private information of a user was obtained by the service provider, the service provider may use it improperly or leak it out intentionally or unintentionally, thus causing loss of the user. Therefore, protecting the user's privacy when requesting services on the Internet has become a requisite and many studies have been conducted on this topic. Most privacy protecting techniques are based on anonymization [6, 7] or encryption [8]. Their basic ideas are to anonymize or encrypt some key values of the user data, so that for each user contained in the data it cannot be identified. The anonymization or encryption techniques can prevent the user's privacy being exposed in data release to some extent. However, they assume that the service provider is trustworthy and will always use the users' data properly, which is unrealistic. Moreover, data leakage or violation may occur inevitably because decryption and anti-anonymization techniques can be applied by malicious users or organizations. Therefore, to reduce the privacy risk or loss of users in service request, policy-based privacy protecting approaches have been proposed for service selection [9]. They allowed a service provider defining privacy requirements using policies and did matchmaking between the user's privacy preference and the service's privacy requirements [10–12] or between two services' privacy requirements for service composition purpose [13–16].

However, previous work did not allow a flexible specification of users' privacy preferences and services' privacy policies, and usually uses a simple way (e.g., weighted summation) to compute the privacy risk or satisfactory degree. In reality, different users probably have different privacy preferences. For instance, some users may regard his name more important than his phone number when releasing his/her personal information, while some other users may not regard so. Moreover, the privacy policy of a service may involve the combination or tradeoff of multiple privacy attributes. For example, a service may ask a user to provide either "name + phone number" or "name + email" when being requested.

To be line with real situations, this paper proposes a privacy-aware service selection approach, which supports personalized specification of privacy preferences and policies and uses a fuzzy logic method to compute the privacy risk of the user on each service candidate. To validate the proposed approach, we use both examples and experiments to evaluate its effectiveness and performance.

The rest of this paper is organized as follows. Section 2 presents the definitions of the concepts appeared in this paper and the service selection problem. Section 3 firstly overviews the proposed approach and then describes in detail the procedures of the approach. Section 4 uses some examples to illustrate the process of service selection. Section 5 conducts analytical and experimental evaluations to show the performance of the proposed approach. Finally, Sect. 6 concludes this paper with an outlook of future work.



## 2 Definitions

### 2.1 User Privacy Preference

The privacy information of a user usually has many attributes, such as name, phone number, birthdate, income, email, address and so on. Different persons could have different privacy preference, i.e., they are likely to have different views on a privacy attribute in terms of its importance and sensitiveness. The following definitions are used to formally describe privacy attributes and user privacy preference:

**Definition 1** (Privacy attributes): The privacy attributes concerned by a user are denoted by the set  $C = \{c_1, c_2, \dots, c_k\}$ , where  $c_i (1 \leq i \leq k)$  represents a privacy attribute.

**Definition 2** (Privacy Sensitiveness): Privacy sensitiveness can be described using a partially ordered set  $P = (PC, \leq)$ , where  $PC$  denotes the set of privacy sensitiveness levels, which can be expressed using linguistic terms, i.e.,  $\{very\ low, low, medium, high, very\ high\}$ . The operator  $\leq$  is a partial order relation based on  $PC$ , so that we have  $very\ low \leq low \leq medium \leq high \leq very\ high$ .

**Definition 3** (Privacy Preference): A user's privacy preference can be denoted by  $F = \{(c, p) \mid c \in C, p \in PC\}$ , where  $c$  is a privacy attribute concerned by the user, and  $p$  represents a privacy sensitiveness level.

### 2.2 Service Reputation and Privacy Policy

In reality, the confidence of a service user on privacy protection usually depends a lot on the reputation and privacy policy of the service. Generally, the higher is the service's reputation, the more confidence will the user have in that the service will properly use his/her privacy. Therefore, this work takes the reputation of services into consideration. The privacy policy of a service usually contains the information about what privacy the service will gather from a user and its promise to use it properly. A privacy policy may be quite complicated in practice, while this work simplifies it by only considering its privacy data requirement. In the following, we formally define the reputation and privacy policy of a service.

**Definition 4** (Service Reputation): Service reputation can be described using a partially ordered set  $R = (RC, \leq)$ , where  $RC$  denotes the set of reputation levels, i.e.,  $\{very\ low, low, medium, high, very\ high\}$ . The operator  $\leq$  is a partial order relation based on  $RC$ , so that we have  $very\ low \leq low \leq medium \leq high \leq very\ high$ .

**Definition 5** (Privacy Policy): The privacy policy (requirement) of a service ( $s$ ) for users can be described using a subset of privacy attributes and an aggregation operator (will be discussed in detail later)  $\amalg$ , as follows:

$$PR(s) = \amalg_{j=1}^k c_j \quad (1)$$

where  $\coprod$  is the fuzzy connective operators  $\wedge$  or  $\vee$ . For instance, a service requires a user's name (denoted by  $c_{name}$ ) and phone number (denoted by  $c_{phone}$ ) or email address ( $c_{email}$ ), its privacy requirement (policy) can be specified as  $c_{name} \wedge (c_{phone} \vee c_{email})$ .

### 2.3 Privacy Risk

**Definition 6** (Individual Privacy Risk): The individual privacy risk is defined as the risk of an individual privacy attribute ( $c_j$ ) to a user caused by a service, denoted by  $\omega_j$ . For instance, if a service requires a user's phone number which is a very sensitive privacy attribute to him/her, there would be a significant privacy risk for the user requesting the service. The individual privacy risk of a user's phone number can be represented by  $\omega_{phone}$ .

The computation of individual privacy risk depends on the user's preference on the privacy attribute and the service's reputation. Basically, the more the user cares about the privacy attribute and the lower is the service's reputation, the higher is the individual privacy risk (more details will be discussed later).

**Definition 7** (Overall Privacy Risk): The overall privacy risk is an aggregation of the individual privacy risks of all privacy attributes to the user caused by a service. For instance, suppose that a service requires the user to provide privacy information such as name, phone number and email address. The overall privacy risk of the service to the user is an aggregation of the individual privacy risks  $\omega_{name}$ ,  $\omega_{phone}$ , and  $\omega_{email}$ .

### 2.4 Privacy-Aware Service Selection

Given a set of services that satisfy the user's functional requirements, the privacy-aware service selection process can be modeled as a ranking in terms of the privacy risk so that for any two services  $S_i$  and  $S_j$ , the following is true:

$$S_i > S_j \iff PR(S_i) \leq PR(S_j)$$

where  $PR(S_i)$  and  $PR(S_j)$  represents the overall privacy risk to the user in terms of  $S_i$  and  $S_j$  respectively. The lower is the privacy risk of a service to the user, the better is the service to the user.

## 3 The Service Selection Approach

Figure 1 is an overview of the proposed privacy-aware service selection approach. Given a set of service candidates with reputation and privacy policies (privacy requirements), and a set of privacy attributes concerned by the user, based on the user's privacy preference, the proposed approach uses the following three steps to rank service candidates based on their assessed privacy risks. First, for each service and each privacy attribute required by the service, according to the user's privacy preference and the service's reputation, the violation risk of the individual privacy attribute is computed. Second, for each service, fuzzy operators are used to aggregate the individual privacy risk on every privacy attribute to compute the overall privacy risk. Finally, all

the service candidates are ranked based on their privacy risk degrees to the user, and the service candidates with the least privacy risks would be recommended to the user.

Please note that the proposed approach is based on the assumption that all the service candidates meet the user's functional requirement, thus we can focus on the privacy risk of service selection. Actually, searching for such a set of service candidates is not difficult to implement in reality, since many functional matching algorithms have been proposed (please refer to [17] for more details).

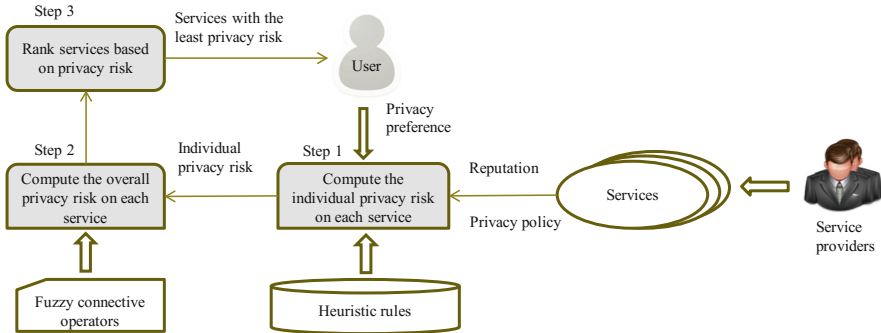


Fig. 1. Overall process of the privacy-aware service selection approach

### 3.1 Computation of Individual Privacy Risk

Different privacy attributes may weight quite differently in a user's view, and thus leads to different risk levels to the user. The sensitiveness of a privacy attribute can be measured using one of the following five levels: *very low*, *low*, *medium*, *high*, *very high*, as defined in Definition 2. The reputation of a service is also key to the privacy risk computation for a user, because it is generally believed that a service with higher reputation would be more trustworthy in using the user's privacy data. Also, the reputation of a service can be expressed using five linguistic terms: *very low*, *low*, *medium*, *high*, *very high*, as defined in Definition 4.

Taking the user's privacy preference and the service's reputation into consideration, a set of heuristic rules are developed to infer the privacy risk of a privacy attribute caused by a service to the user, and they are summarized in Table 1. Actually, the twenty-five heuristic rules are extended from the general heuristic rules as follows:

- (1) If the sensitiveness of the privacy attribute is high and the reputation of the service is low to the user, the privacy risk is high to the user;
- (2) If the sensitiveness of the privacy attribute is low and the reputation of the service is high to the user, the privacy risk is low to the user;
- (3) If the sensitiveness of the privacy attribute is low and the reputation of the service is low to the user, or the sensitiveness of the privacy attribute is high and the reputation of the service is high to the user, the privacy risk is on a level between low and high (e.g., medium).

Since both privacy sensitiveness and service reputation have five linguistic terms, the privacy risk also has such five values: *very low, low, medium, high, very high*.

**Table 1.** Heuristic rules for privacy risk inference based on user preference and service reputation

Reputation	Sensitiveness				
	Very low	Low	Medium	High	Very high
Very low	Medium	High	Very high	Very high	Very high
Low	Low	Medium	High	Very high	Very high
Medium	Very low	Low	Medium	High	Very high
High	Very low	Very low	Low	Medium	High
Very high	Very low	Very low	Very low	Low	Medium

### 3.2 Computation of Overall Privacy Risk

A service may require user data that involve multiple privacy attributes. Different services are likely to have different privacy policies, which require different privacy data from users. For example, an online shopping service may require a user’s name, phone number, credit card number and delivery address, while another online shopping service, in addition to the four kinds of user data, may also require the user’s identity card number and email address. Moreover, as stated in Definition 5, a privacy policy can use fuzzy connective operators,  $\wedge$  or  $\vee$ , to define complex and flexible privacy requirements. In order to properly compute the overall privacy risk of a service to the user and rank services based on their risk degrees, we specify the following requirements.

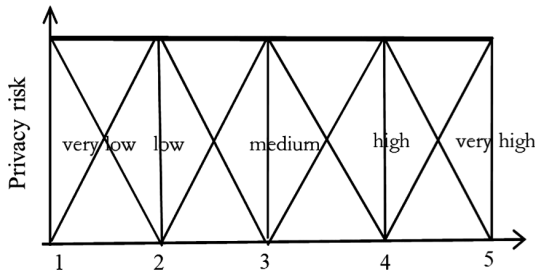
- (1) It would be better for the overall privacy risk uses crisp values instead of linguistic terms, so that services with the same linguistic overall privacy risk can be further distinguished in ranking.
- (2) More privacy data required by a service typically result in a higher privacy risk to the service user. Meanwhile, more sensitive privacy data required by a service typically result in a higher privacy risk to the service user.
- (3) The fuzzy connective operator  $\wedge$  would generate larger privacy risk values than the fuzzy connective operator  $\vee$  when aggregating two privacy attributes.

The computation of the overall privacy risk of a service to the user is based on the above requirements. To obtain crisp privacy risk values, we first need to transform the individual privacy risk degrees computed in the above procedure from linguistic terms to real number values. In this work, we use the centroid method (CM) [18] for defuzzifying linguistic term-based values. The CM, also known as either the center of gravity (CoG) or center of area (CoA) method, is the most commonly used defuzzification technique. This technique provides a crisp value based on the CoG of the fuzzy set. It also determines the best point for dividing the fuzzy set into exactly two masses. Because linguistic terms can be decomposed in a triangular shape, the CM becomes a

suitable approach for defuzzifying the linguistic privacy risk terms. For a triangular fuzzy number  $F = (\mu, \lambda, \rho)$ , its real number value ( $\omega$ ) can be calculated as

$$\omega = \lambda + \frac{(\mu - \lambda) + (\rho - \lambda)}{3} \tag{2}$$

In this work, five linguistic terms are decomposed into triangular fuzzy numbers using the triangular fuzzy set shown in Fig. 2. These linguistic terms, their fuzzy numbers and corresponding real number values are presented in Table 2.



**Fig. 2.** Triangular membership functions for the five linguistic terms of privacy risk

**Table 2.** The five linguistic terms, their fuzzy numbers, and corresponding real number values

Linguistic terms	Fuzzy numbers	Real number value
Very high	(4, 5, 5)	4.67
High	(3, 4, 5)	4.00
Medium	(2, 3, 4)	3.00
Low	(1, 2, 3)	2.00
Very low	(1, 1, 2)	1.33

Multiple privacy attributes must be aggregated based on fuzzy connective operators to obtain an overall privacy risk value. The fuzzy connective operators  $\wedge$ ,  $\vee$  are discussed in the following.

The fuzzy connective operators  $\wedge$  is used to connect two privacy attributes when they both are required by a service. If  $c_i$  and  $c_j$  are two privacy attributes required by a service, and their personalized privacy risks are  $\omega_i$  and  $\omega_j$  respectively, their privacy risks can be aggregated using the summation operator as follows:

$$\omega_i \wedge \omega_j = \omega_i + \omega_j \tag{3}$$

The fuzzy connective operators  $\vee$  is used to connect two privacy attributes when either of them is required by a service. If  $c_i$  and  $c_j$  are two privacy attributes, and the service user is asked to supply either of them to the service,  $\vee$  can be used to connect

them in privacy policy specification. In this case, the privacy risks of  $c_i$  and  $c_j$  can be aggregated using the minimum operator as follows:

$$\omega_i \vee \omega_j = \min\{\omega_i, \omega_j\} \tag{4}$$

### 3.3 Service Ranking

After the privacy risk of each service candidate is evaluated for the user, the service candidates are ranked in order of increasing privacy risk, and the top-ranked services are recommended to the user. If several service candidates have the same privacy risk value, the services with higher reputation will be placed prior to the others.

## 4 Illustrative Examples

This section uses some examples to illustrate our proposed service selection approach. We consider the online purchase services, which may require users to supply personal and sensitive information like name, address, email, phone number, ID, credit card, mobile payment account, etc. These privacy attributes (data items) and their notations are tabulated in Table 3. Given five online purchase services that meet the user’s functional requirement, their reputation values and privacy requirements are supposed to be different and are presented in Table 4.

**Table 3.** Privacy attributes and their notations

Privacy attributes	Name	Address	Identity number	Email	Phone number	Credit card	Mobile payment account
Notation	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$

**Table 4.** Five services with privacy requirements and reputation degrees

Service ID	Privacy requirements	Reputation
1	$c_1 \wedge c_2 \wedge c_4 \wedge (c_6 \vee c_7)$	High
2	$c_1 \wedge c_5 \wedge c_6$	Medium
3	$c_3 \wedge (c_4 \vee c_5) \wedge c_7$	Very high
4	$c_2 \wedge c_3 \wedge c_5 \wedge c_6$	Low
5	$c_1 \wedge c_3 \wedge c_6 \vee c_5 \vee c_6$	High

Suppose there are three users interested in the online purchase services, and they have different preferences on the privacy attributes, as shown in Table 5. We will illustrate in the following, how to calculate the privacy risk of each service candidate to each user.

**Table 5.** Three users and their preferences on different privacy attributes

User ID	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$
1	Low	Low	Medium	High	Very high	High	Medium
2	Very low	Very low	Low	Very high	Medium	Low	High
3	Medium	Low	Very low	Low	High	Medium	Low

Tables 6, 7 and 8 respectively shows the individual privacy risk degrees of the privacy data items supplied by the three users to the services. This is done based on our privacy risk heuristic rules tabulated in Table 1. We can see that the privacy risks of a privacy data item caused by a service to different users are probably different because their privacy preferences are different.

Based on the defuzzification technique and fuzzy connective operators (discussed in Sect. 3.2), the individual privacy risks presented in Tables 6, 7 and 8 can be aggregated to compute the overall privacy risk value of each service to a user, as shown in Table 9. From Table 9, we can identify the best service with least privacy risk for each user, as shown in Table 10.

**Table 6.** The individual privacy risks of different services to user 1

Service	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$
1	Very low	Very low	Low	Medium	High	Medium	Low
2	Low	Low	Medium	High	Very high	High	Medium
3	Medium	Low	Very low	Low	High	Medium	Low
4	Medium	Medium	Very high	Very high	Very high	Very high	Very high
5	Very low	Very low	Low	Medium	High	Medium	Low

**Table 7.** The individual privacy risks of different services to user 2

Service	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$
1	Very low	Very low	Very low	High	Low	Very low	Medium
2	Very low	Very low	Low	Very high	Medium	Low	High
3	Very low	Very low	Very low	Medium	Low	Medium	Very low
4	Low	Low	Medium	Very high	High	Medium	Very high
5	Very low	Very low	Very low	High	Low	Very low	Medium

**Table 8.** The individual privacy risks of different services to user 3

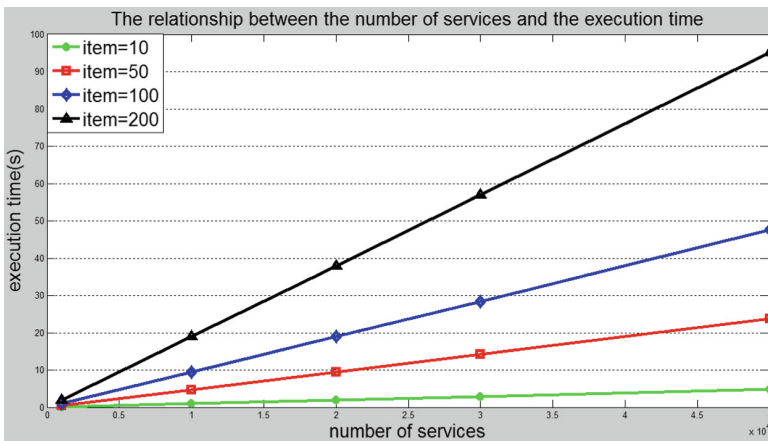
Service	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$
1	Low	Very low	Very low	Very low	Medium	Low	Very low
2	Medium	Low	Very low	Low	High	Medium	Low
3	Very low	Very low	Very low	Very low	Low	Very low	Very low
4	High	Medium	Low	Medium	Very high	High	Medium
5	Low	Very low	Very low	Very low	Medium	Low	Very low

**Table 9.** The overall privacy risk of different services to the users

	Service 1	Service 2	Service 3	Service 4	Service 5
User 1	7.66	10.67	7.00	16.01	6.33
User 2	7.99	6.33	4.66	12	3.33
User 3	5.99	10	3.99	13.67	5

**Table 10.** The service with least privacy risk to the users

User 1	Service 5
User 2	Service 5
User 3	Service 3

**Fig. 3.** The execution time with respect to number of services

## 5 Evaluation

### 5.1 Analytical Evaluation

The performance of the entire approach can be analyzed as follows. First, there is a search for services satisfying functional requirements, which is upper bounded by the number of services (denoted  $n$ ), i.e.,  $O(n)$ . Then, the privacy risk degree of each privacy data item for each service is computed. This is also upper bounded by the number of user-concerned privacy data items being considered (denoted  $k$ ), i.e.,  $O(k)$ . Thirdly, the individual privacy risk degrees of each service are aggregated, which is upper bounded by the number of privacy attributes and services, i.e.,  $O(kn)$ . Finally, the services are ranked based on their privacy risk degrees to the user, which has a time complexity of



$O(n \log n)$ . Therefore, the entire service selection system is of  $O(kn + n \log n)$ , with the bottle necks being the number of available services and the number of privacy attributes under consideration.

### 5.2 Experimental Evaluation

The experimental evaluation was performed to see the scalability of the proposed approach with respect to the number of privacy attributes and the number of available services. We run the approach with a varied number of privacy attributes (10, 50, 100 and 200) against different number of services (1K, 10K, 20K, 30K, and 50K). Figure 3 shows the execution time versus the number of services with respect to the different number of privacy attributes. We can see that the execution time increases slowly with respect to the number of services, which meets our expectations. This indicates that our proposed approach has a good scalability.

### 5.3 Impact of User Preference on Privacy Risk Evaluation

This analysis is performed to determine the impact on the actual results if the user’s privacy preference is tweaked. Our expectations were that the change of the user’s privacy preference would have significant influences on the privacy risk computation. We use the examples discussed in Sect. 4 to do analysis. Let the privacy sensitiveness of user 1 on the privacy attribute  $c_1$  be changed from *low* to *medium*, as shown in Table 11, the privacy risks of the user on different services are recomputed and the results are shown in Table 12. Compared with Table 9, we can see that for those services that require the privacy  $c_1$  their privacy risk values are significantly changed with respect to the change of the user’s sensitiveness on  $c_1$ . The privacy risk of service 3 on the user is significantly reduced and becomes as low as service 5, thus it can also be recommended to the user. From the above results, it can be concluded that the user’s preference has influence on the privacy risk evaluation on services.

**Table 11.** Tweak the user 1’s privacy preference

	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$
Before tweak	Low	Low	Medium	High	Very high	High	Medium
After tweak	Medium	Low	Medium	High	Very high	High	Medium

**Table 12.** The privacy risk of user 1 after privacy preference tweaked

	Service 1	Service 2	Service 3	Service 4	Service 5
User 1	8.33	11.67	7.00	16.01	7

**Table 13.** Tweak the service 5’s reputation

	Reputation degree
Service 5	High $\rightarrow$ medium

**Table 14.** The privacy risk of users on service 5 after its reputation tweaked

	Service 5
User 1	8.67
User 2	5
User 3	7

#### 5.4 Impact of Service Reputation on Privacy Risk Evaluation

This analysis is performed to determine the impact on the actual results if the service's reputation is tweaked. Our expectations were that the change of the service's reputation would have significant influences on the privacy risk computation. Again, we use this the examples discussed in Sect. 4 to do analysis.

Let the reputation of servicer 5 be changed from *high* to *medium*, as shown in Table 13, its privacy risks on different users are recomputed and the results are shown in Table 14. Compared with Table 9, we can see that the privacy risk values of service 5 are significantly changed to all users. Especially for user 1 and user 2, service 5 has the least privacy risk to them before its reputation change, while after change, it is not the best candidate any more. From this observation, it can be concluded that the service's reputation has influence on the privacy risk evaluation on services.

## 6 Conclusion

This paper presents a privacy-aware service selection approach for reducing the privacy risk of users when requesting services on the Internet. The approach evaluates a service's risk degree to a user's privacy based on its reputation, privacy policy and the user's privacy preference. It thus can help a user identify the service with least privacy risk among a set of service candidates. Examples and experiments have been provided to evaluation the proposed approach, and it is demonstrated that the proposed approach is effective and can perform efficiently. The future work will take quality of service into consideration, and will investigate the tradeoff relation between service quality and privacy requirement.

**Acknowledgments.** The work described in this paper is supported by the National Natural Science Foundation of China under Grant Nos. 61572186 and 61572187.

## References

1. Jeong, B., Cho, H., Lee, C.: On the functional quality of service (FQoS) to discover and compose interoperable web services. *Expert Syst. Appl.* **36**(3), 5411–5418 (2009)
2. Wang, P.: QoS-aware web services selection with intuitionistic fuzzy set under consumer's vague perception. *Expert Syst. Appl. Part 1* **369**(3), 4460–4466 (2009)
3. Fletcher, K.K., Liu, X.F., Tang, M.: Elastic personalized non-functional attribute preference and trade-off based service selection. *ACM Trans. Web (TWEB)* **9**(1), 1–27 (2015)

4. Fan, W., Yang, S., Perros, H., et al.: A multi-dimensional trust-aware cloud service selection mechanism based on evidential reasoning approach. *Int. J. Autom. Comput.* **12**(2), 208–219 (2015)
5. Featherman, M.S., Miyazaki, A.D., Sprott, D.E.: Reducing online privacy risk to facilitate e-service adoption: the influence of perceived ease of use and corporate credibility. *J. Serv. Mark.* **24**(3), 219–229 (2010)
6. Sweeney, L.: k-anonymity: a model for protecting privacy. *Int. J. Uncertainty Fuzziness Knowl.-Based Syst.* **10**(5), 557–570 (2002)
7. Ammar, N., Malik, Z., Medjahed, B., et al.: K-anonymity based approach for privacy-preserving web service selection. In: 2015 IEEE International Conference on Web Services (ICWS), pp. 281–288. IEEE (2015)
8. Li, X., Jung, T.: Search me if you can: privacy-preserving location query service. In: IEEE INFOCOM 2013, Turin, Italy, 14–19 April 2013, pp. 2760–2768 (2013)
9. Lin, L., Liu, T., Hu, J., Ni, J.: PQsel: combining privacy with quality of service in cloud service selection. *Int. J. Big Data Intell.* **3**(3), 202–214 (2016)
10. Kapitsaki, G.M.: Reflecting user privacy preferences in context-aware web services. In: IEEE 20th International Conference on Web Services (ICWS), pp. 123–130. IEEE (2013)
11. Squicciarini, A., Carminati, B., Karumanchi, S.: A privacy-preserving approach for web service selection and provisioning. In: 2011 IEEE International Conference on Web Services (ICWS), pp. 33–40. IEEE (2011)
12. Lin, L., Liu, T., Hu, J., Zhang, J.: A privacy-aware cloud service selection method toward data life-cycle. In: 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS), Hsinchu, Taiwan, pp. 752–759, 16–19 December 2014
13. Costante, E., Paci, F., Zannone, N.: Privacy-aware web service composition and ranking. In: 2013 IEEE 20th International Conference on Web Services (ICWS), pp. 131–138. IEEE (2013)
14. Carminati, B., Ferrari, E., Tran, N.H.: A privacy-preserving framework for constrained choreographed service composition. In: 2015 IEEE International Conference on Web Services (ICWS), 297–304. IEEE (2015)
15. Squicciarini, A.C., Carminati, B., Karumanc, S.: Privacy aware service selection of composite web services. In: 9th International Conference on Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), pp. 260–268. IEEE (2013)
16. Tbahrity, S.-E., Ghedira, C., Medjahed, B., Mrissa, M.: Privacy-enhanced web service composition. *IEEE Trans. Serv. Comput.* **7**(2), 210–222 (2014)
17. Sajjanhar, A., Hou, J., Zhang, Y.: Algorithm for web services matching. In: Yu, J.X., Lin, X., Lu, H., Zhang, Y. (eds.) APWeb 2004. LNCS, vol. 3007, pp. 665–670. Springer, Heidelberg (2004). [https://doi.org/10.1007/978-3-540-24655-8\\_72](https://doi.org/10.1007/978-3-540-24655-8_72)
18. Opricovic, S., Tzeng, G.H.: Defuzzification within a multicriteria decision model. *Int. J. Uncertainty Fuzziness Knowl.-Based Syst.* **11**(5), 635–652 (2003)

# An Efficient Lattice-Based Proxy Signature with Message Recovery

Faguo Wu, Wang Yao, Xiao Zhang<sup>(✉)</sup>, and Zhiming Zheng

Key Laboratory of Mathematics, Informatics and Behavioral Semantics,  
Ministry of Education, School of Mathematics and Systems Science,  
Beihang University, Beijing 100191, China  
09621@buaa.edu.cn

**Abstract.** Proxy signature scheme is an important cryptographic primitive in which an entity can delegate its signing rights to another entity, the purpose of proxy signature with message recovery is to shorten the length of proxy signatures which can effectively reduce the communication overhead. Although message recovery proxy signature scheme based on conventional number-theoretic problems has been proposed for a long time, the message recovery technique draws no attention to proxy signature scheme from lattice. In this paper, we firstly propose a proxy signature scheme with message recovery from lattice which is more efficient than previous proxy signature schemes in signature size, time and energy cost, and we prove that in the random oracle, our scheme is secure model under the hardness assumption of SIS. Our proxy signature scheme with message recovery would work well in the quantum age based on the underlying lattice problems.

**Keywords:** Lattice-based · Proxy signature · Message recovery  
Communication overhead · Quantum age

## 1 Introduction

Proxy signatures, proposed by Mambo et al. [1], are significant cryptographic systems that are widely used in different situations, such as E-commerce, cloud computing and Electronic election. In a proxy signature scheme, Original signer A delegates his signing power and responsibility to another one which is called proxy signer. According to the hardness of traditional Number Theory problems, the researchers have proposed many effective proxy signature schemes such as integer factoring-based schemes, discrete logarithm schemes and elliptic curve schemes [2,3]. However, in a long run, all these proxy signature schemes are not secure, because both discrete logarithms and large prime factorization algorithms can be solved in polynomial-time [4] when we take quantum computing into account. For purpose of reducing the threat from the quantum age, many researchers pay their attention to Post-Quantum Cryptography, some proxy signature schemes, like hash-based schemes, MPKC schemes, lattice-based schemes, which are considered

as Post-Quantum Cryptography, and many corresponding efficient Post-Quantum proxy signature schemes have been proposed, such as [5–9].

Lattice-based signature occupies a position of particular interest, as it relies on well-studied problems and comes with uniquely strong security guarantees [10]. Organizations and research groups are looking forward for efficient lattice-based cryptography schemes to replace RSA and ECC based schemes, and many creative and constructive results show that properly optimized lattice schemes may be competitive with, or even outperform, classical factoring and discrete logarithm-based cryptography.

The first signature scheme with message recovery was proposed by Nyberg and Rueppel [11]. In a signature scheme with message recovery, we only need to transmit the signature without the signed message, because verifier can easily recover the signed message from the signature. This construction quietly adapts to situations where small signed message to be transmitted or strict bandwidth requirements [12, 13]. The existent proxy signature schemes with such property can be categorized into two different types: discrete logarithm based and RSA based [14–16] which efficiently improve the performance of previous signature without message recovery. Although there are many very effective latticed-based proxy signature schemes, based on the hardness assumption of lattice, have been proposed [17–19], as far as we know, there are no efficient lattice-based proxy signature schemes with message recovery have been proposed.

In this paper, benefit from the techniques [20, 21], we propose an efficient lattice-based proxy signature scheme with message recovery, and prove that it has existentially security in the random oracle model, in addition, contribution to message recovery technology and mainly multiplication of matrix and vector given in our scheme, our scheme enjoy higher performance when compared with others' proxy signature scheme, finally, when we take piratical situation into consideration [22], we are surprised to find that this kind of scheme consume less energy even when we add some operations for message recovery, which means our proxy signature schemes are extremely suitable for system with low energy and low bandwidth. As the hardness assumption of lattice problem SIS, our lattice-based message recovery proxy signature scheme would work well in the quantum age.

The remainder of our paper is organized as follow. In Sect. 2, we provide necessary preliminaries of our scheme. In Sect. 3, we describe two models of our lattice-based proxy signature scheme with message recovery: syntax model and security model. In Sect. 4, we propose our efficient message recovery proxy scheme from lattice. In Sect. 5, we present the formal security analysis of our scheme. In Sect. 6, we introduce some necessary criterions, and give detailed comparisons between our scheme and some existing proxy schemes from lattice.

## 2 Preliminaries

### 2.1 Notations

In this paper, following notations would be used:

- $x \parallel y$  denotes the connection of two string  $x$  and  $y$ , and they are effectively recoverable.

- $M^{n \times (k_1+k_2)} = M_1^{n \times k_1} \parallel M_2^{n \times k_1}$  denotes the concatenation of Matrices  $M_1, M_2$ .
- $\|v\|_p$  denotes the  $l_p$  norm of  $v$ .
- $|x|$  denotes the quantity of bits of  $v$ .
- $|x|^{l_1}$  denotes the first left  $l_1$  bits of  $x$ .
- $|x|_{l_2}$  denotes the first right  $l_2$  bits of  $x$ .

### 2.2 Lattice

**Definition 1.** A lattice  $L$  is a discrete subgroup of some space  $\mathbb{R}^n$ , it is generated by independent vector  $v_1, v_2, \dots, v_k \in \mathbb{R}^n$  through the following way:

$$A = L(v_1, v_2, \dots, v_k) = \left\{ \sum_{i=1}^k a_i v_i \mid a_i \in \mathbf{Z} \right\}$$

The basis of  $L$  are vectors  $v_1, v_2, \dots, v_n$ , lattice's rank is the integer  $n$  where  $k < n$  and  $a_i$  is coefficient.

**Definition 2.** Given integers  $q, m, n$  and a matrix  $A \in Z_q^{n \times m}$ , for some  $S \in Z^m$ .

$$A_q(A) = \{x \in Z^m : x = A^T S = u(\text{mod}q)\}$$

$$A_q^\perp(A) = \{x \in Z^m : x = A^T S = 0(\text{mod}q)\}$$

From the above definition, these two types of lattices are dual to each other.

### 2.3 Gaussian on Lattice

In lattice-based signature scheme, Gaussian series are very effective techniques which are widely used, and we have a briefly review of it here.

**Definition 3** (Discrete Gaussian distribution).  $\sigma \in \mathbb{R}^m$  is standard deviation, vector  $c \in Z^m$  is center, Continuous Gaussian distribution  $\rho_{c,\sigma}^m(x)$  and Discrete Gaussian distribution  $D_{c,\sigma}^m(x)$  are defined as follow:

$$\rho_{c,\sigma}^m(x) = \left( \frac{1}{\sqrt{2\pi\delta^2}} \right)^m e^{-\frac{\|x-c\|^2}{2\sigma^2}}$$

$$D_{c,\sigma}^m(x) = \frac{\rho_{c,\sigma}^m(x)}{\sum_{z \in Z^m} \rho_{c,\sigma}^m(z)}$$
(1)

When  $c = 0$ , we can simply write  $\rho_{c,\sigma}^m(x), D_{c,\sigma}^m(x)$  as  $\rho_\sigma^m, D_\sigma^m$ , and from [21], an important theorem of Discrete Gaussian distribution is described as follow

**Theorem 1.**  $\forall \sigma > 0$  and  $m \in Z^+$

- (1)  $P[x \in D_\sigma^1 : |x| > 12\sigma] < 2^{-100}$ ;
- (2)  $P[x \in D_\sigma^m : \|x\| > 2\sigma\sqrt{m}] < 2^{-m}$ .

---

**Algorithm 1.** Rejection sampling technique

---

**Input:**  $H : \{\{0, 1\}^* \rightarrow v : v \in -1, 0, 1^k, \|v\| < c\}$  (Where  $k \in Z$  and  $\ll m$ ), message  $u$ , a matrix  $A$  randomly sampled from  $Z_q^{m \times n}$ ,  $S$  (signature key) sampled from  $\{-d, \dots, 0, \dots, d\}^{m \times k_1}$

**Output:** Vector  $v$  and  $c$

- 1: Obtain  $y$  randomly from  $D_\sigma^m$
  - 2:  $C = H(Ay, u)$
  - 3:  $Z = SC + y$
  - 4: return  $(Z, C)$  with probability  $\min(\frac{D_\sigma^m(Z)}{MD_{SC, \sigma}(Z)}, 1)$
- 

**2.4 RST: Rejection Sampling Technique**

For a lattice-based signature scheme, the most important conception of the RST is to eliminate the relationship between signing key and output signature’s distribution [10], the algorithm as follow (Algorithm 1).

**2.5 Small Integer Solution (SIS) and its Hardness Assumption**

**Definition 4.** For an integer modular homogeneous scheme  $As = 0 \text{ mod } q$ , get a proper solution  $s \in Z^m$  where  $q$ , matrix coefficient, small solution  $s$  satisfy  $q \in Z^m$ ,  $A \in Z_q^{n \times m}$  and  $\|s\| \leq \beta$  where  $\beta$  is a real value.

In reference [10,23], they proved that for any polynomial-bound  $m, \beta$  and any prime  $p$ , with small factors and the Gaussian measure, there is no difference between the hardness of some worst-case approximation and average-case harness of SIS. Even the hardness of SIS problem has been proved, there still exist overwhelming probability that anyone can solve some case if the trapdoor of  $f = Ax(\text{mod } q)$  is got.

**3 Lattice-Based Proxy Signature Scheme with Message Recovery**

Syntax model and security model of our lattice-based proxy signature scheme with message recovery are proposed in this section.

**3.1 Syntax**

**Definition 5.** In such lattice-based proxy signature scheme with message recovery scheme, there are three participants: An original signer with  $ID_o$ , a proxy signer:  $ID_p$ , and a verifier, and this scheme is consists of six PPT algorithm (Setup, KeyGen, DelGen, DelVer, Psign, Pver), where:

1. Setup: Given a security parameters  $n$ , and select appropriate parameters and functions.  $(par, n) \leftarrow Setup$ .

2. KeyGen: Given a security parameters  $n$ , this algorithm output the secret and public key of original signer and proxy signer: Original signer's  $= (PK_o, SK_o)$ , Proxy signer'  $= (PK_p, SK_p)$ .  $(PK_o, SK_o, PK_p, SK_p) \leftarrow KeyGen(n, M)$ .
3. DelGen: Given the original signer's public key  $PK_o$  and secret key  $SK_o$  hash function  $H_i$ , proxy signer's  $ID_p$ , this algorithm output the Delegation Key  $(PK_D, SK_D)$  for original signer, original signer sends this pair to proxy signer in security channel.  $(PK_D, SK_D) \leftarrow DelGen(H_i, ID_p, PK_D, SK_D)$ .
4. DelVer: Given the original signer's public key  $PK_o$  and Delegation Key  $(PK_D, SK_D)$ , and check  $DelVer(PK_o, PK_D, SK_D) = 1$  or not. If it outputs 1, it's a valid delegation of original signer.  $\{0, 1\} \leftarrow DelVer(PK_o, PK_D, SK_D)$ .
5. Psign: Given the secret delegation key  $SK_D$ , proxy signer's secret key  $SK_p$  and the message  $u = u_1 \parallel u_2$ , output the proxy signature  $\theta$ , that is  $\theta \leftarrow Psign(u, SK_D, SK_p)$ .
6. Pver: Given the public key  $PK_o$  of original signer, public key  $PK_o$  of proxy signer, public delegation key  $PK_D$ , proxy signer's  $ID_p$ , hash function  $H_i$ , partial message  $u_2$ , and proxy signature  $\theta$ , if the proxy signature is valid, output 1, otherwise, output 0, that is  $(m, \{0, 1\}) \leftarrow Pver(PK_o, PK_o, PK_D, ID_p, H_i, \theta, u_2)$ .

**Remark 1.** For consistency requirements, partial message  $u_2$ , the proxy signature  $\theta$  of secret Delegation Key  $SK_D$  and proxy signer's secret key  $SK_p$  must hold with overwhelming probability with following equation  $Verify(Sign(SK_D, SK_p, u), u_2, PK_D, SK_D) = 1$ .

### 3.2 Security Model

In a lattice-based proxy signature scheme with message recovery, the properties of Unforgeability, Verifiability, Strong identifiability, Strong undeniability and Key dependence are satisfied naturally. Therefore, we consider in this lattice-based proxy signature scheme under adaptive chosen message and identity attack. To have a formal security definition for this scheme, the security model is an security game played between a adversary A and a challenger C:

1. Setup: In this game, the challenger C firstly run the algorithm Setup( $n$ ), get the necessary parameters and send them to the adversary A.
2. Queries: In such query-game, following types of queries can be adaptively issued by adversary A within polynomial bound number of questions.
  - KeyGen-query: The adversary A can issue a query on the ID which he want to get the secret key, and the challenger run the algorithm KeyGen, and return A with  $SK_{ID}$  in response.
  - DelGen-query: To get the delegation key  $SK_D$ , the adversary A input two secret key corresponding to the identity  $ID_o$  and  $ID_p$ , in response, the challenger C run the algorithm DelGen, and return A with  $SK_D$ .
  - Psign-query: When adversary issues such on  $ID_p$  with message  $u$ , the challenger C run the algorithm Psign, and return A with signature.



3. By the above queries, the adversary A generate a valid proxy signature  $\theta'$  on message  $u^*$  under the identity  $ID'$ , if the following holds, Adversary A wins the game: (i)  $Vfy(Sign(SK_D, SK_P, m), PK_D, SK_D, par) = 1$ ; (ii)  $u^*$  has never been send to the Psign-query; (iii) all identities which is related to  $ID'$  have never been sent to KeyGen-query.

An lattice-based proxy signature scheme with message recovery is considered as existential unforgeable if the advantage of Adversary A wins the above query game in polynomial time is negligible.

## 4 Our Lattice-Based Proxy Signature Scheme with Message Recovery

We gave a detailed account of our efficient message recovery proxy signature scheme from lattice in this section. Like the traditional signature systems, our scheme have three participants: an original signer with  $ID_o$ , a proxy signer with identity  $ID_p$ , and a verifier, and this scheme is consists of six probabilistic polynomial time (PPT) algorithm (Setup, KeyGen, DelGen, DelVer, Psign, Pver), where:

1. Setup: Given the security parameter  $n$  of this system, we select  $l_1, l_2, k_1, k_2, m, q \in N$ , where  $q$  is a prime, select five hash function:  $H_1 = Z_q^n \rightarrow \{0, 1\}^{l_1+l_2}$ ,  $H_2 = \{0, 1\}^* \rightarrow \{0, 1\}^{k_1+k_2}$ ,  $H_3 = ID \rightarrow \{-1, 0, 1\}^{k_1 \times k_2}$ ,  $F_1 = \{0, 1\}^{l_2} \rightarrow \{0, 1\}^{l_1}$ ,  $F_2 = \{0, 1\}^{l_1} \rightarrow \{0, 1\}^{l_2}$ .  $H_1, H_2$  are seen as a random oracle.
2. KeyGen: On input security parameter  $n$ , randomly choose  $A \in Z_q^{n \times m}$  together with two secret matrix basis  $S_1, S_2 \in \{-d, \dots, 0, \dots, d\}^{m \times k_1}$ . Original signer computes  $T_1 = AS_1 \text{ mod } q$ , and keep  $PK_o = (A, T_1)$  as his own public key,  $SK_o = S_1$  as secret key. Proxy signer computes  $T_2 = AS_2 \text{ mod } q$ , and keep  $PK_p = (A, T_2)$  as his own public key,  $SK_p = S_2$  as secret key.
3. DelGen: Original signer computes  $t = H_3(ID_p)$ ,  $S_3 = S_1 t \in Z^{m \times k_2}$ ,  $T_3 = T_1 t \in Z^{n \times k_2}$ , and send  $(S_3, T_3)$  to Proxy signer via an safe authenticated channel, Proxy signer keep the delegation key  $PK_D = S_3$ , proxy signer can use it to generate valid proxy signatures stand for original signer, and the corresponding public key is  $PK_D = T_3$ .
4. DelVer: The Proxy signer receives  $(S_3, T_3)$  from original signer, and checks if  $AS_3 = T_1 t$  holds, where  $t = H_3(ID_p)$ .
5. Psign: The Proxy signer with identity  $ID_p$  do the following:
  - Divide the message  $u$  into two parts  $u = u_1 \parallel u_2$  where  $|u_1| = l_2$ , if  $|u| < l_2$ , let  $u_2 = \perp$ .
  - Select a random  $y \in D_\sigma^m$ , and compute  $\alpha = H_1(Ay)$ .
  - Let  $u'_1 = F_1(u_1) \parallel (F_2(F_1(u_1)) \oplus u_1)$ ,  $r = \alpha \oplus u'_1$ .
  - Compute  $C = H_2(r, u_2)$ .
  - Compute  $Z = (S_2 \parallel S_3)C + y \in Z_q^m$ .
  - The proxy signature on the message  $u$  is  $(r, Z, u_2)$  with probability  $\min(\frac{D_\sigma^m(Z)}{MD_{(S_2 \parallel S_3)C, \sigma}(Z)}, 1)$ .

6. Pver: Given  $(r, Z, u_2)$ , a verifier does as explained in the succeeding text:
- Compute  $\alpha = H_1(AZ - (T_2 \parallel T_3)H_2(r, u_2))$ .
  - Compute  $u'_1 = r \oplus \sigma$ ,  $u_1 = |u'_1|_{l_2} \oplus F_2(|u'_1|^{l_1})$ , and then recover message  $u = u_1 \parallel u_2$ .
  - Check if  $\|Z\| < 2\sigma\sqrt{m}$  and  $F_1(u_1) = |u'_1|^{l_1}$  hold at same time, if so, accept signature from the proxy signer, otherwise, reject it.

**Theorem 2.** *Our lattice-based proxy signature scheme with message recovery is correctness.*

*Proof.* From the above detailed construction, we can easily have following equations where  $u$  message

$$\begin{aligned} &AZ - (T_2 \parallel T_3)H_2(r, u_2) \\ &= A((S_2 \parallel S_3)C + y) - (T_2 \parallel T_3)H_2(r, u_2) \\ &= Ay \end{aligned}$$

the distribution of  $(S_2 \parallel S_3)C + y$  is statically closed to the distribution  $D_\sigma^m$ , and from the Theorem 1,  $\|(S_2 \parallel S_3)C + y\| \leq 2\sigma\sqrt{m}$  with probability at least  $1 - 2^{-m}$ . On the other hand,  $u_1 = F_1(u_1) \parallel (F_2(F_1(u_1)) \oplus u_1)$ , we can recover  $u_1 = |u'_1|_{l_2} \oplus F_2(|u'_1|^{l_1})$  with  $F_1(u_1) = |u'_1|^{l_1}$  hold.

## 5 Security Analysis

As a signature scheme, security is the most important factor, we prove that our lattice-based proxy signature scheme with message recovery is secure enough (unforceability) under the hardness assumption of SIS in this section.

When we proving the unforceability of proxy signature scheme with message recovery, we should take two types adversary into consideration:

**Type(i)** Adversary  $A$  can not only have the public key  $PK_o$  of original signer and public key  $PK_p$  of proxy signer, but also have the original signer’s secret key  $SK_o$ .

**Type(ii)** Adversary  $A$  has neither the original signer’s secret key  $SK_o$  nor the proxy signer’s secret key  $SK_p$ .

It’s obvious that adversary in **Type(i)** knows more information than the adversary in **Type(ii)**, so we only need to take **Type(i)** adversary into consideration. We suppose there is a polynomial time, adversary  $A$  forge a valid proxy signature by at most  $q_{H_1}$  times  $H_1$  query,  $q_{H_2}$  times  $H_2$  query,  $q_{F_1}$  times  $F_1$  query,  $q_{F_2}$  times  $F_2$  query, and  $q_s$  times signature query with non-negligible probability, it means there exist an algorithm  $C$  (Challenger  $C$ ) which can solve a  $SIS_{q,n,m,C}$  problem “with the help of” Adversary  $A$  in polynomial time. Algorithm  $C$  (Challenger  $C$ ) has a game with  $A$ , the following is the simulation.

**Queries.** The adversary  $A$  issues the following types of queries adaptively,  $A$  has random oracle  $H_1 - query$  before any other queries.

- $H_3$  – query. Challenger C maintains a list  $L_0 = (ID_{p_i}, PK_{p_i}, SK_{p_i})$ , and the initial value is null, when adversary A issues a query on  $ID_{p_i}$ , Challenger C search it in list first, if there exist corresponding tuple  $(ID_{p_i}, PK_{p_i}, SK_{p_i})$ , return  $PK_{p_i}$ ; otherwise, Challenger C randomly chooses matrix  $S \in Z_q^{m \times k_2}$ , then Challenger C computes  $PK = AS$ , Update the list  $L_0$  as  $L_0 = (L_0, (ID, PK, SK))$ , return  $PK$ .
- $KeyGen$  – query. When adversary A issues a query on  $ID_p$ , Challenger C look it up in  $L_0$ , find a match tuple  $(ID, PK, SK)$ , and output  $SK$  as response.
- $DelGen$  – query. On receiving the secret key  $SK_P$ , Challenger C outputs  $SK_D$  as response.
- $H_1$  – query. Challenger C maintains a list  $L_1 = (Ay, \alpha_i)$ , and the initial value is null. When the adversary A issue a query for  $Aymodq$ , Challenger C search it in list first, if there exist corresponding tuple  $(Ay, \alpha)$ , return  $\alpha$ ; otherwise, randomly chooses  $\alpha \in \{0, 1\}^{k_1+k_2}$ . Update the list  $L_1$  as  $L_1 = (L_1, (Ay, \alpha))$ , then return  $\alpha$ .
- $F_1$  – query. Challenger C maintains a list  $L_2 = (u_1, F_1(u_1))$ , and the initial value is null. When the adversary A issue a query for  $u_1$ , Challenger C search it in list first, if there exist corresponding tuple  $(u_1, F_1(u_1))$ , return  $F_1(u_1)$ , otherwise, randomly chooses  $F_1(u_1) \in \{0, 1\}^{l_1}$ , Update the list  $L_2 = (L_2, (u_1, F_1(u_1)))$ .
- $F_2$  – query. Challenger C maintains a list  $L_3 = (F_1(u_1), F_2(F_1(u_1)))$ , and the initial value is null. When the adversary A issue a query for  $F_1(u_1)$ , Challenger C search it in list first, if there exist corresponding tuple  $(F_1(u_1), F_2(F_1(u_1)))$ , return  $F_2(F_1(u_1))$ , otherwise, randomly chooses  $F_2(F_1(u_1)) \in \{0, 1\}^{l_2}$ , Update the list  $L_3 = (L_3, (F_1(u_1), F_2(F_1(u_1))))$ .
- $H_2$  – query. Challenger C maintains a list  $L_4 = (r_i, u_i, z_i, c_i)$ . When the adversary A issues a query for  $(r, u = u_1 \parallel u_2)$ , the Challenger C search it in list first, if there exist corresponding tuple  $(r, u, c, z)$ , return C; otherwise, randomly chooses vector  $Z \in D_\sigma^m, C \in \{v : v \in \{-1, 0, 1\}^m\}$ ,  $H_1$  – query  $AZ - (T_2 \parallel T_3)modq$  for  $\alpha$ , let  $u_1 = \alpha \oplus r$ , and according to  $u_1 = F_1(u_1) \parallel (F_2(F_1(u_1)) \oplus u_1)$ , and Update  $L_3 = (L_3, (F_1(u_1), F_2(F_1(u_1)) \oplus u_1))$ ,  $L_2 = (L_2, (u_1, F_1(u_1)))$ . Update  $L_1 = (r, u, c, z)$  where  $r, u$  satisfied  $H_2(r, u_2) = C$ , then return C.
- $Psign$  – query. To obtain a proxy signature on message  $u$ , adversary search it in  $L_4$ , if Challenger C find a match tuple  $(r, u, c, z)$ , then output  $(r, z)$  as a response; otherwise, randomly choose vector  $C \in \{-1, 0, 1\}^{k_1+k_2}, Z \in D_\sigma^m$ , adversary A  $H_1$  – query  $AZ - T_2 \parallel T_3C$  for  $\alpha$ ,  $F_1$ –query and  $F_2$ –query for  $(u_1, F_1(u_1))$  and  $(F_1(u_1), (F_2(F_1(u_1)) \oplus u_1))$ , let  $r = \alpha \oplus u_1'$ . Update  $L_1 = (L_1, (r, u, c, z))$  where  $H(r, u_2) = C$ . then return  $(r, u_2)$  and  $u_2$ .

**Forgery.** The adversary finally outputs a valid forgery  $(r, z, u_2^*)$  on message  $u^* = u_1^* \parallel u_2^*$ .

The specific example SIS problem: In order to solve SIS problem, adversary A should find a small vector  $x \in \Lambda_q^\perp(A)$ , Challenger responses to adversary A with different results when adversary A repeats his queries. According to General Forking Lemma [24], adversary A finally gets a valid forgery  $(r^\#, Z^\#, u_2^*)$  on

same message  $u^* = u_1^* \parallel u_2^*$  with non-negligible probability, and the following equation satisfied, where  $C^\# = H_2(r^\#, u_2^*), H_2(r^*, u_2^*) = C^*$ .

According to the above construction, we can see that for any message  $u$ :

- $H_2(r^\#, u_2^*) \neq H_2(r^*, u_2^*)$
- $AZ^\# - (T_2 \parallel T_3)H_2(r^\#, u_2^*) - (AZ^* - (T_2 \parallel T_3)H_2(r^*, u_2^*)) = A(Z^\# - Z^* - (S_2 \parallel S_3)C^\# + (S_2 \parallel S_3)C^*)$
- $\| Z^\# \| \leq 2\sigma\sqrt{m}, \| Z^* \| \leq 2\sigma\sqrt{m}, \| (S_2 \parallel S_3)C^\# \| \leq d_{k_1+k_2}\sqrt{m}$ , and  $\| (S_2 \parallel S_{+3})C^* \| \leq d_{k_1+k_2}\sqrt{m}$
- $\| Z^\# - Z^* - (S_2 \parallel S_{+3})C^\# + (S_2 \parallel S_{+3})C^* \| \leq (4\delta + 2d_{k_1+k_2})\sqrt{m}$

According to [21],  $Z^\# - Z^* + (S_2 \parallel S_3)C^\# - (S_2 \parallel S_3)C^* \neq 0$  with probability at least 1/2, so  $Z^\# - Z^* + ((S_2 \parallel S_3)C^\# - (S_2 \parallel S_3)C^*) \neq 0$  with non-negligible ability. Our lattice-based proxy signature scheme with message recovery is Unforgeability.

## 6 Efficiency Analysis

When we have a efficiency analysis of our scheme, We take signature size, computation time and energy cost into consideration. In this section, we analyse our scheme’s efficiency by comparing it with some existing proxy signature schemes from the length of secret delegation key, proxy signature message and total time cost. In order to simplify the presentation, We define R = Rejection sampling algorithm computation cost, T = TrapGen algorithm computation cost, S = SamplePre algorithm computation cost, B = BasisDel algorithm computation cost, E = ExtBasis algorithm computation cost, M = Multiplication of matrix, M = Multiplication of vector,  $M = m = O(n \log n), q = O(n^2), M = \omega(\sqrt{\log m}), S_1, S_2 \in \{-1, \dots, 0, \dots, 1\}^{m \times k_1}, \sigma = 12k_2\sqrt{m}$ . Table 1 are given the detailed sizes and time of the comparison.

**Table 1.** Comparison of related proxy signature schemes

Proxy signature scheme	Delegation key length	Signature proxy	Time
[17]	$4m^2 \log(LM\sqrt{2m})$	$ u  + 2m \log(LM^2 2m)$	$2T + 2S + E$
[18]	$m^2 \log(LM\sqrt{2m})$	$ u  + m \log(LM^5 m^2)$	$T + 3S + 3B$
[19]	$ml' \log q$	$ u  + k + l + 2m \log(12\delta)$	$R + M + V$
Our’s	$ml'' \log q$	$ u_2  + l_1 + l_2 + m \log(12\delta)$	$\approx R + (M + V)/2$

From Table 1, it is obvious that the total length (signed message and signature) of our message recovery scheme is less than other proposed schemes which is the foremost advantage of a lattice-based proxy signature scheme with message recovery, besides, we can find that in proxy signature [17, 18], they mainly use TrapGen algorithm, SamplePre algorithm, BasisDel algorithm and ExtBasis algorithm which are very time consuming, while in our scheme and [14], based on

Lyubashevsky's rejection sampling algorithm, we mainly use the multiplication of matrix and vector, and due to different technique used in verification, we are almost twice as fast as [14].

When we let  $k + l = l_1 + l_2$  and take security parameters mentioned in [21] into consideration ( $n = 512, q = 2^{57}, d = 1$ ), we can have a direct comparison with [19] as following

$$\Delta_{\text{Length of Proxy Signature and Message}} \approx l_2 + m \log(12\delta) = (l_2 + 163000) \text{ bits} \quad (2)$$

Refer to the comparison proposed in [22], 1 bit transmission cost more energy than 32 bits simple operation, in that case, even we increase simple computation (like hash and XOR) in message recovery technology, our scheme still cost much less energy than [19]'s in pracRef1tical situation. According to the above analysis, our message recovery proxy signature scheme is more efficient than these existing schemes in signature size, time and energy cost.

## 7 Conclusion

With the development of quantum computers, constructing an efficient quantum-secure proxy signature scheme enjoys priority. Lattice-based signature occupies a position of particular interest, as it relies on well-studied problems and comes with uniquely strong security guarantees, such as worst-case to average case reductions. In this work, we proposed an efficient lattice-based proxy signature with message recovery which is possible to be the first proxy signature scheme with message recovery that can resist known quantum attack, and we give a formal proof of it's security in the random oracle model. In addition, compared with some existing proxy signature schemes, our scheme is more efficient than others in signature length, signature time and energy cost. Contribution to rich theoretical foundation of Lattice Cryptography, we will design more efficient lattice-based signature schemes with message recovery in the future.

**Acknowledgements.** This work was supported by NSFC (61402030), the Major Program of National Natural Science Foundation of China (11290141), and Fundamental Research of Civil Aircraft no. MJ-F-2012-04.

## References

1. Mambo, M., Usuda, K., Okamoto, E.: Proxy signatures: delegation of the power to sign messages. *IEICE Trans. Fundam. A* **79**(9), 1338–1354 (1996)
2. Kumar, R., Verma, H.K., Dhir, R.: Analysis and design of protocol for enhanced threshold proxy signature scheme based on rsa for known signers. *Wirel. Pers. Commun.* **80**(3), 1281–1345 (2015)
3. Xiao, Y.M.: Improvement of an Elliptic curve based threshold proxy signature scheme (2016)
4. Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. In: *Quantum Entanglement and Quantum Information-Proceedings of Ccast*, pp. 303–332 (1999)

5. Tang, S., Xu, L.: Proxy signature scheme based on isomorphisms of polynomials. In: Xu, L., Bertino, E., Mu, Y. (eds.) NSS 2012. LNCS, vol. 7645, pp. 113–125. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-34601-9\\_9](https://doi.org/10.1007/978-3-642-34601-9_9)
6. Yang, C., Qiu, P., Zheng, S., Wang, L.: An efficient lattice-based proxy signature scheme without trapdoor. In: International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 189–194 (2016)
7. Chen, Y.Z., Liu, Y., Wen, X.J.: A quantum proxy weak blind signature scheme. *Chin. J. Quantum Electron.* **54**(4), 1325–1333 (2011)
8. Zhang, L., Ma, Y.: A lattice-based identity-based proxy blind signature scheme in the standard model. *Math. Probl. Eng.* **2014**(1) (2014)
9. Wang, T.Y., Wei, Z.L.: Analysis of forgery attack on one-time proxy signature and the improvement. *Int. J. Theor. Phys.* **55**(2), 1–3 (2015)
10. Micciancio, D., Regev, O.: Worst-case to average-case reductions based on Gaussian measures. In: IEEE Symposium on Foundations of Computer Science, pp. 372–381 (2004)
11. Nyberg, K., Rueppel, R.A.: A new signature scheme based on the DSA giving message recovery. In: Proceedings of the ACM Conference on Computer and Communications Security, CCS 1993, Fairfax, Virginia, USA, pp. 58–61. November 1993
12. Simoens, P., Vankeirsbilck, B., Deboosere, L., Ali, F.A., Turck, F.D., Dhoedt, B., Demeester, P.: Upstream bandwidth optimization of thin client protocols through latency-aware adaptive user event buffering. *Int. J. Commun. Syst.* **24**(5), 666–690 (2011)
13. Liu, C.X., Liu, Y., Zhang, Z.J., Cheng, Z.Y.: High energy-efficient and privacy-preserving secure data aggregation for wireless sensor networks. *Int. J. Commun. Syst.* **26**(3), 380–394 (2013)
14. Padhye, S., Tiwari, N.: ECDLP-Based Certificateless Proxy Signature Scheme with Message Recovery. Wiley, Hoboken (2015)
15. Zhou, C.: An improved ID-based proxy signature scheme with message recovery. *Int. J. Secur. Appl.* **9**(9), 151–164 (2015)
16. Asaar, M.R., Salmasizadeh, M., Susilo, W.: A Short ID-Based Proxy Signature Scheme. Wiley, Hoboken (2016)
17. Xia, F., Yang, B., Ma, S., Sun, W.W., Zhang, M.W.: Lattice-based proxy signature scheme. *J. Hunan Univ.* **38**(6), 84–88 (2011)
18. Kim, K.S., Hong, D., Jeong, I.R.: Identity-based proxy signature from lattices. *J. Commun. Netw.* **15**(1), 1–7 (2013)
19. Jiang, M.M., Yupu, H.U., Baocang, Y.U., Lai, J.J.: Efficient lattice-based proxy signature. *J. Beijing Univ. Posts Telecommun.* (2014)
20. Tian, M., Huang, L.: Lattice-based message recovery signature schemes. *Int. J. Electron. Secur. Digit. Forensics* **5**(3/4), 257–269 (2013)
21. Lyubashevsky, V.: Lattice signatures without trapdoors. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 738–755. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-29011-4\\_43](https://doi.org/10.1007/978-3-642-29011-4_43)
22. Barr, K.C.: Energy-aware lossless data compression. *ACM Trans. Comput. Syst.* **24**(3), 250–291 (2006)
23. Ajtai, M.: Generating hard instance of lattice problems. In: Twenty-Eighth ACM Symposium on Theory of Computing ACM, pp. 99–108 (1996)
24. Bellare, M., Neven, G.: Multi-signatures in the plain public-key model and a general forking lemma. In: ACM Conference on Computer and Communications Security, CCS 2006, Alexandria, VA, USA, pp. 390–399, October 30–November 03 2006

# FABAC: A Flexible Fuzzy Attribute-Based Access Control Mechanism

Yang Xu<sup>1</sup>(✉), Wuqiang Gao<sup>1</sup>, Quanrun Zeng<sup>1</sup>, Guojun Wang<sup>2</sup>, Ju Ren<sup>1</sup>,  
and Yaoxue Zhang<sup>1</sup>

<sup>1</sup> School of Information Science and Engineering, Central South University,  
Changsha 410083, China

{xuyangcsu, forrestgao, zengquanrun, renju, zyx}@csu.edu.cn

<sup>2</sup> School of Computer Science and Educational Software, Guangzhou University,  
Guangzhou 510006, China  
csgjwang@gmail.com

**Abstract.** Attribute-Based Access Control (ABAC) is a promising approach for addressing intricate management requirements in dynamic and distributed environments. Nevertheless, because of lacking flexible access exception handling mechanism, rigid rules in ABAC influence the resource availability and ultimately the working efficiency. In this paper, we propose a novel fuzzy ABAC model (FABAC) that extends the ABAC with better usability. We introduce the fuzzy mechanism into decision-making process. Based on the membership grades of requests to rules and the spare credits of respective subjects, our framework permits additional requests failing in rule matching, thus enhancing the information flows in business processes. Furthermore, we develop the credit system with history-based recovery mechanism, wherein the subject's credits and corresponding recovery rate are impacted by the past authorizations on sub-standard requests, for maintaining the risk of abuse under control. The analysis reveals that our model contributes to attaining better tradeoff between security and usability.

**Keywords:** ABAC · Fuzzy mathematics · Credit · Flexibility  
Security

## 1 Introduction

Recent computing technologies, for instance mobile computing and cloud computing [1] have introduced improvements in the resource availability of organizations to their workforce. Nevertheless, this trend unavoidably results into more intricate, dynamic and distributed access scenarios, for instance, data access in mobile office or Bring-Your-Own-Devices (BYOD) environments.

Classical access control models like Discretionary Access Control (DAC) [2], Mandatory Access Control (MAC) [3] and Role-Based Access Control (RBAC) [4] are so inflexible and oversimplified that they fade in such unpredictable scenarios where user's permissions require continuous readjustments. In the meantime, even more models throwing focus on multiple security-relevant factors

(subject states, object states, environmental states, etc.), have been studied in order to attain effective access control. Characteristically, the Attribute-Based Access Control (ABAC) is a powerful fine-grained policy-based technique that has been gradually developed from the research phase [5–9] to practical stage [10]. Through the establishment of access policies composed of multi-dimensional attributes for the control of access requests, ABAC implements more accurate and adaptive control on recourses and better management extensibility that is capable of addressing more intricate management requirements in nomadic, dynamic and distributed environments.

Nevertheless, this kind of access control technique has far more rigid (sometimes even poorly designed) policy rules that in turn pose influence on the resource availability and ultimately the efficiency of business, particularly for the entities like brokerages, emergency centers and intelligence departments. For instance, one of the Fortune 500 companies constrains a portion of sensitive data to be accessed only by the employees situated within the office site during a specific time period. Unfortunately, the real-time location data can be both incorrect and unstable while it is quite easy for someone to miss the time, such as, just by two minutes. Accordingly, the employees easily meet failures in matching the policies and then unable to attain corresponding data in time as inefficient manual interventions by administrators are always needed for the purpose of permitting exceptional requests in this kind of scenarios.

It goes without saying that there is a need for a bit more liberty with the authorization of exceptional access requests so that information flows could be improved in business processes. Hence, more advanced mechanism is needed in order to make the ABAC a more flexible and effective access control technique, which is practical for complex, dynamic and distributed environments.

Since access control is essentially a tradeoff between risk and benefit, the risk mechanism, together with its other side trust, has already been incorporated into access control models for the purpose of effective and flexible decision-making. The risk assessment mechanism based on fuzzy techniques has met success in extending conventional frameworks such as MAC and RBAC [11,12] with the ability to permit additional accesses by assessments, which cut down the rigidity of models to some extent. In quite recent works about access control [13–16], risk and trust assessment mechanisms are frequently applied in conjunction with some other factors such as context, delivering more adaptive and flexible features.

Inspired by these facts, we have the belief that the assessment mechanism is perfectly appropriate for the ABAC in order to enhance its flexibility and efficiency in handling exceptional access requests. And as an evaluation method succeed in numerous areas, the fuzzy mathematical tool [17] that infers results from ambiguous evidences is supposed to be a good candidate for the evaluation match quality of access requests with rules in the policy-based ABAC model.

Therefore, in this paper, we propose a fuzzy ABAC model (FABAC), wherein the original ABAC model is enhanced to attain better flexibility in handling exceptional access requests. We equip the original decision-making mechanism with a fuzzy mechanism and treat the pre-defined ABAC rules as measurement criteria.



In respect of any request, we perform the evaluation of its maximum match degree with rules by membership grade calculation. And when it does not exactly match any rule, a preset rejection threshold is applied to determine whether it requires further assessment. The final decision for fuzzy request is made in accordance with whether the initiating subject spare enough credits. Accordingly, our model is capable of granting additional requests that fail to match any rule precisely without manual intervention that leads to the improvement of information flows in business processes. Additionally, we develop the credit system with the adjustable recovery mechanism in order to mitigate the risk posed by our model. By reducing the credit of a subject as well as the recovery rate every time when its exceptional request is granted, our model is reasonably constrained from granting too many fuzzy requests for any subject. The case study describes how our model functions in detail, and the analysis of FABAC model shows that the usability of resources has been improved, while the risk it imposed is under control.

Summarily, our major contributions are twofold:

- (1) We equip the ABAC model with fuzzy mechanism in order to make it flexible enough to handle exceptional access requests. Through the evaluation of the match quality of access request with rules, our model brings forth opportunities to the almost matched requests to be granted without manual intervention, for improving data availability and attaining better business efficiency.
- (2) We not only set a threshold to filter out unreasonable requests with low match degrees, but also deliver an effective credit system with adjustable recovery mechanism based on extra approval histories for requests, so that the fuzzy mechanism could be prevented from being abused and thus the risk is kept under control.

The rest of this paper is organized as follows. Section 2 performs the reviews of the related work. Section 3 introduces the preliminary knowledge about fuzzy mathematics. In Sect. 4, we describe the FABAC model followed by a case study. Section 5 brings forth a discussion about our model. Finally, we conclude this paper and outline the future improvements in Sect. 6.

## 2 Related Work

Access control is a fundamental technique of security and plays a predominant role in protecting private and confidential information from unauthorized access. Focusing on confidentiality, integrity, usability and administration flexibility, dozens of access control models have been developed over the decades. Among them, DAC [2], MAC [3] and more recently RBAC [4] are the three most commonly used ones. However, these classical models above are inflexible and take no account of additional factors (e.g. time, location or user IP). So they become gradually unable to meet emerging demands in geographical, temporal and context-aware information systems.

Differs from passive subject-object models, more expressive access control schemes are proposed from other perspectives.

One interesting attempt is introducing risk factor to achieve balance between system security and usability. Fuzzy mechanism has been introduced to extend RBAC [12] for attaining more flexibility in response to exceptions. This strategy seeks to obtain a higher degree of flexibility at the expense of security. Cheng et al. [11] introduced Fuzzy MLS, a risk-adaptive access control model, which performed quantitative risk analysis and achieved a better equilibrium between risk and benefit. Meanwhile, trust mechanism, closely connected to the concept of risk, has also been taken into consideration. Dimmock et al. [13] extended existing access control schemes to incorporate trust-based assessment and ratiocination in order to attain a more resilient authorization mechanism. Mahalle et al. [14] presented a fuzzy approach to the Trust Based Access Control with the notion of trust levels for identity management. Context awareness is a significant prerequisite for accurately perceiving and properly handling risks. Feng et al. [15] integrated user behaviors and context environments to propose a trust and context based access control model which is scalable, and suitable for dynamic distributed systems. Through consideration of both factors of trust and environmental perception, Bhatti et al. [16] outlined a mechanism to develop a trust-based, context-aware model for web services based on the X-GTRBAC (XML based Generalized Temporal Role-based Access Control) framework.

As intra/inter-organizational collaboration and information sharing become more common, there is growing concern about a multi-dimensional access control paradigm where access requests are granted or denied based on a set of policies composed of all kinds of attributes, including subject attributes (e.g. age, department, position), operation attributes (e.g. read, delete, write), object attributes (e.g. owner, size, sensitivity) and environment attributes (e.g. time, location), specifically ABAC [5]. It is scalable and manageable, and also ensures the access control are applied consistently. As a result, ABAC is considered as “next generation” authorization model for its fine-granularity, context-awareness, rich semantics and other beneficial features. Therefore, it draws wide attention and has an attractive application prospect.

Standards organizations and industrial community [5,10] have made great efforts to promoting further improvement and practical deployment of ABAC. Moreover, large quantities of academic researches have also flourished [18]. Li et al. [6] explored in depth about the relationships among access request, attribute authority, policies and decision procedure of ABAC. Jin [7] has presented a relatively formal model about ABAC. Sookhak et al. [8] provided a comprehensive survey on ABAC schemes suitable for cloud and distributed computing.

In spite of the advantages of ABAC, its rigid policies involving too many attributes may produce challenges in management and deployment which lead to the reduction of usability and efficiency of information flows. Ngo et al. [9] presented a multi-tenant ABAC model for cloud services which can support multiple levels of delegations with improved flexibility for inter-tenant collaborations.

This scheme increases the flexibility of ABAC to a certain extent, however, its applicability is limited to specific scenario. From a more essential perspective, it reflects the fact that ABAC lacks a general approach to handling exceptional accesses flexibly without time-consuming human approval.

Motivated by such challenges in nomadic, dynamic and distributed scenarios, our work tends to present an innovative approach FABAC based on fuzzy theory and credit mechanism in purpose of improving the flexibility of native ABAC.

### 3 Preliminary

In this section, we introduce some background knowledge about fuzzy theory [17].

**Fuzzy Set.** A fuzzy set is a pair  $(U, \mu)$  where  $U$  is a set while  $\mu$  is a mapping relation called membership function, as follows:

$$\begin{aligned} \mu : U &\rightarrow [0, 1] \\ x \in U &\rightarrow \mu(x) \in [0, 1] \end{aligned} \tag{1}$$

For each  $x \in U$ , the value  $\mu(x)$  is called the membership grade of  $x$  in  $(U, \mu)$ . If  $\mu(x) \in (0, 1)$ , then  $x$  is called a fuzzy member in the fuzzy set. Besides,  $x$  is called fully included in the fuzzy set if  $\mu(x) = 1$ , or called not included in the fuzzy set if  $\mu(x) = 0$ .

**Mamdani-Type Fuzzy Inference.** Fuzzy logic is a kind of many-valued logic which mimics Boolean logic. The most popular definition of fuzzy synthetic is proposed by Mamdani [19], in which AND operator takes the minimum value of all antecedents, while OR operator takes the maximum one. There is also a NOT operator which transits the value of membership function from positive to negative and then adds an integer “1”.

## 4 FABAC

In this section, we firstly introduce some notations and then describe the overall framework and working procedures, followed by two of its major parts: the matching module and credit mechanism. At last, we give a further explanation of FABAC model through a case study.

For easy and concise description, we simplify the problem by assuming that only positive rules are involved with “deny” preferable strategy<sup>1</sup>, i.e., an access request should be granted if and only if it matches at least one rule.

### 4.1 FABAC Model

We extend the original ABAC model and present a new fuzzy attribute-based access control model (FABAC), which takes fuzzy logic and credit mechanism into consideration and achieves better flexibility and higher usability.

<sup>1</sup> Any negative rule can be transformed to positive rule.

**Formalization.** To further describe the FABAC model, we use the following notations and definitions:

$Attr = \{attr_i\}$  is the universe of attributes.

$Sub, Obj, Env, Op$  are sets of attributes of subject, object, environment and operation respectively, and  $Sub \cup Obj \cup Env \cup Op \subseteq Attr$ .

$req_i$  is a certain request consists of values of corresponding attributes represented by a vector  $\langle v_{req_i}(attr_1), v_{req_i}(attr_2), \dots, v_{req_i}(attr_n) \rangle$ , where  $v_{req_i}(attr_k)$  represents the concrete value of  $attr_k$  in  $req_i$ .

$RS_j = (R, \mu_{rule_j})$  is a fuzzy set corresponds to a rule  $rule_j$  in the FABAC model, in which  $R = \{req_i\}$  is the value domain of requests, and  $\mu_{rule_j}$  is the corresponding membership function.

$AS_{j,k} = (D(attr_k), \mu_{j,k})$  is a fuzzy set to represent the matching degree of a value of  $attr_k$  with the restriction for  $attr_k$  involved in  $rule_j$ , where  $D(attr_k)$  is the value domain of  $attr_k$ , and  $\mu_{j,k}$  is the corresponding membership function.

$\mu_{all}(req_i)$  is the overall membership function reflecting the matching degree of  $req_i$  with its best matching rule.

$T \in (0, 1)$  is the rejection threshold. Any request whose membership grade below  $T$  will be rejected directly.

$crd_x$  is a credit value of an subject  $x$ . Whenever  $x$  successfully access without meeting all attributes requirements of the rules, the credit value will be consumed according to certain policy. The value of credit is gradually approaching its upper limit as time goes by.

**Framework and Procedures.** We build the FABAC framework upon the original ABAC model by equipping it with fuzzy matching module as well as a credit mechanism, shown in Fig. 1, wherein the former offers additional authorization opportunities to exceptional requests, while the latter helps to maintain the risk of abuse under control. Both of these modules are loosely coupled with the original framework and thus are quite easy to deploy.

When a subject initiates an access request, the FABAC decision module firstly gathers concrete values of all kinds of attributes of this request, including Subject Attributes, Object Attributes, Environment Attributes and Operation Attributes (Steps 1–2). And then it calculates the degree of compliance of this request to each rule by calculating the membership degree. According to the maximum membership grade, a exactly matched request will be directly permitted, otherwise, it is considered as a **fuzzy request** and will be denied or pended for further decision by the threshold filter (Step 3). Then the decision module will turn to credit mechanism for support (Step 4). If corresponding subject spare enough credit for this exceptional request, a “clearance” suggest will be returned with the consumption of corresponding credit along with the reduction of the recovery rate. Otherwise, this request will eventually be rejected (Step 5). Finally, subject will be granted/denied access to object accordingly (Step 6). Note that a granted fuzzy request is called a **fuzzy access**.

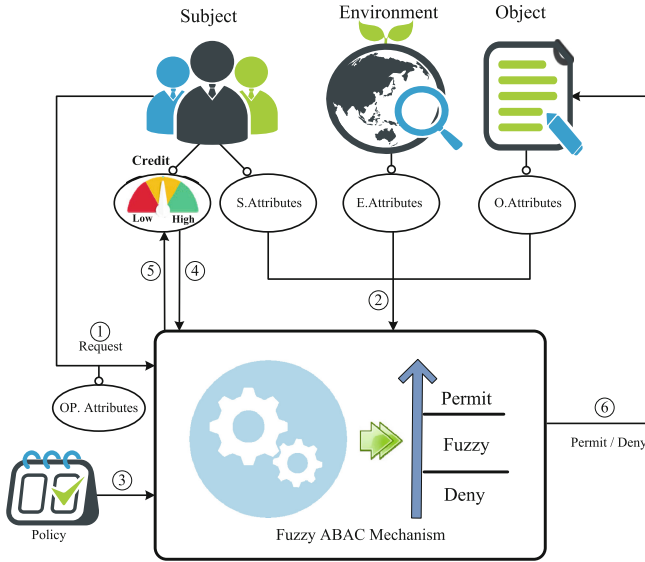


Fig. 1. Framework and working procedures of FABAC

**Matching Module.** In our model, when an access request  $req_i$  is initiated, we will calculate the degree of compliance of  $req_i$  to each rule by the membership degree calculation, the complete process is as follows:

First, for each  $attr_k$  involved in the  $rule_j$ , we build a corresponding fuzzy set  $AS_{j,k} = (D(attr_k), \mu_{j,k})$  to represent the set of attribute values acceptable to the rule. To reduce the burden on administrators, the membership function could be automatically generated from some function templates instead of leaving the design task to administrators.

Then we list the formula for calculating the degree of membership of  $req_i$  to  $rule_j$  as follows:

$$\mu_{rule_j}(req_i) = \sum_{k=1}^n \frac{\mu_{j,k}(v_{req_i}(attr_k))}{n} \tag{2}$$

When there exist multiple rules, we choose the Mamdani-type fuzzy system for synthesis. Then the overall membership degree  $\mu_{all}(req_i)$  is listed as follows:

$$\mu_{all}(req_i) = \max_{j=1}^n \{\mu_{rule_j}(req_i)\} \tag{3}$$

The formula above represents request’s conformity to the most consistent rule in the system. If the membership grade  $\mu_{all}(req_i) = 1$ , the request will be granted directly; if it is less than  $T$ , the request will be refused; if  $\mu_{all}(req_i) \in (T, 1)$ , we will make further decision according to the subject’s credit  $crd$ , which will be described in the following part.

**Credit Mechanism.** The matching module provides more relaxed access control, so that some requests which fail to completely match the rules can be granted. To prevent the matching module from being abused, we employ the credit mechanism to restrict users who immoderate perform fuzzy request. Any time a subject is granted a fuzzy request, its credit will be consumed accordingly.

To achieve this, we hold a real number  $crd_x \in (0, crd_{max})$  to represent the credit value of a certain subject  $x$ . If a request  $req_i$  initiated by  $x$  satisfies  $\mu_{all}(req_i) \in (T, 1)$ , we will attempt to consume credit values in exchange for successful access. If  $crd_x > 1 - \mu_{all}(req_i)$ , request will be allowed with the withdrawing of  $1 - \mu_{all}(req_i)$  from  $crd_x$ .

Obviously, when  $crd_x$  is used up, the FABAC model will be almost equivalent to the usual ABAC model. In order for the system to be functional persistently, we introduce recovery mechanism into FABAC. The  $crd_x$  should gradually increase to its upper bound as time goes by, which ensures that the whole system does not degenerate into a normal ABAC model quickly. Moreover, to further restrict the abuse of credit mechanism, we introduce a recovery rate adjustment mechanism. When a user consumes credits frequently, his credit recovery rate will be slower than usual. We refer to the logistic function [20], which is often used to describe population recovery and easily adjusted, to construct our credit mechanism. Then we get the implementation of the mechanism listed as follows:

$$\frac{d(crd_x)}{dt} = r \cdot crd_x (1 - crd_x) \tag{4}$$

Here,  $r$  could be a mutable parameter affecting the recovery rate of  $crd_x$ . For any subject, if its request has been successfully granted without exactly matching a rule, its  $r$  should be decreased. A possible implementation is simply half the  $r$  whenever a fuzzy request  $req$  of subject  $x$  is granted.

## 4.2 Case Study

In this subsection, we give a further explanation of the FABAC model by a case.

First, we define a FABAC model with the threshold  $T = 0.85$ , the initial max credit  $crd_{max} = 0.2$  and the initial recovery rate  $r = 1$ . We assume that there exist two rules in the system:

$rule_1 : Env_{location} \text{ in office } (205.395583333332, 57.9323888888888)$   
 $rule_2 : (Sub_{job} \text{ is manager}) \text{ and } (Env_{time} \text{ in } (9 : 00, 17 : 00))$

In this case, the request is defined as a 4-tuple  $req_i = \langle user\_id, t, l, p \rangle$ , where  $user\_id$  is the identification of user,  $t$  is the timestamp of the request,  $l$  represents the location where the request is initialized ( $l$  is given in latitude and longitude), and  $p$  is the subject's job position. Then the above rules are converted in the following way:

$$rule_1 = \langle AS_{1,1} \rangle$$

$$rule_2 = \langle AS_{2,1}, AS_{2,2} \rangle$$

Here, the  $AS_{j,k} = (D(attr_k), \mu_{j,k})$  is the fuzzy set representing the matching degree between the restriction for  $k$ -th attribute involved in  $rule_j$  and corresponding value of the request.

The expressions of membership functions  $\mu_{rule_1}$  of  $RS_1$  and  $\mu_{rule_2}$  of  $RS_2$  are as follows:

$$\begin{aligned} \mu_{rule_1}(req_i) &= \mu_{1,1}(l \text{ of } req_i) \\ \mu_{rule_2}(req_i) &= \frac{\mu_{2,1}(t \text{ of } req_i) + \mu_{2,2}(p \text{ of } req_i)}{2} \end{aligned}$$

Then we detail the subfunctions  $\mu_{1,1}$ ,  $\mu_{2,1}$  and  $\mu_{2,2}$  as follows:

$$\begin{aligned} \mu_{1,1}(l) &= \begin{cases} 1 & , dis(l, office) \in [0, 30) \\ \frac{50-dis(l, office)}{20} & , dis(l, office) \in [30, 50) \\ 0 & , otherwise \end{cases} \\ \mu_{2,1}(t) &= \begin{cases} 2(t - 8.5) & , t \in (8.5, 9] \\ 1 & , t \in (9, 17] \\ 2(17.5 - t) & , t \in (17, 17.5] \\ 0 & , otherwise \end{cases} \\ \mu_{2,2}(p) &= \begin{cases} 1, & p \in \{manager, staff\} \\ 0, & otherwise \end{cases} \end{aligned}$$

Here, the function  $dis(x, y)$  represents the distant of two location in meters. We assume that a subject  $x$  initiates  $req_1$  and  $req_2$  in separate days:

$$\begin{aligned} req_1 &= \langle 003, (205.395298933332, 57.9323888888888), 2017.08.13 \ 21 : 35, manager \rangle \\ req_2 &= \langle 003, (206.395583333332, 57.9323888888888), 2017.08.14 \ 17 : 03, manager \rangle \end{aligned}$$

We assume that when the first visit occurs, the credit  $cred_x = cred_{max} = 0.2$ . Then for the first request  $req_1$ , we have:

$$\begin{aligned} \mu_{all}(req_1) &= \max(\mu_{rule_1}, \mu_{rule_2}) \\ &\approx \max(0.9, 0.5) \\ &= 0.9 \end{aligned} \tag{5}$$

According to the result, the  $req_1$  will consume 0.1 of  $cred_x$ . Since  $cred_x = 0.2 > 0.1$  now,  $req_1$  is granted and the new  $cred_x$  is 0.1.

When it comes to  $req_2$ , similarly, we get the membership degree  $\mu_{all}(req_2) = 0.85$ . Besides, as the recovery rate  $r$  is reduced to 0.5 due to  $req_1$ , according to formula (5), we get the following expression of current  $cred_x$  about the  $\Delta t$ , where  $\Delta t = (time_2 - time_1) \approx 0.81(day)$  is the time period since  $req_1$  was granted.

$$\begin{aligned} cred_x(\Delta t) &= \frac{cred_{max}cred_x(t_0)e^{r\Delta t}}{cred_{max} + cred_x(t_0)(e^{r\Delta t} - 1)} \\ &= \frac{0.2 \cdot 0.1e^{0.5\Delta t}}{0.2 + 0.1(e^{0.5\Delta t} - 1)} \\ &\approx 0.12 \end{aligned} \tag{6}$$

The  $crd_x$  is restored to 0.12 as the time goes by. However, to grant  $req_2$  needs a consumption of 0.15 of  $crd_x$ , so this request should be rejected.

## 5 Discussion

In this section, we analyze the FABAC model from the perspectives of usability and security.

First, we discuss the usability of the system. Let  $U$  be the usability of a certain FABAC model, we believe that  $U$  is positively associated to the granted rate of requests in the system, that is,  $U \propto P$  (Here the  $\propto$  represents the **positive relationship** and  $P$  represents the granted rate). Further, from the definition of the request granted rate, we can get the following formula:

$$U \propto P = P_{fuzzy} + P_{normal} \tag{7}$$

In the above formula,  $P_{fuzzy}$  and  $P_{normal}$  represent the probabilities that a fuzzy request is granted, and a request exactly matches at least one rule respectively. As  $P_{normal}$  is an unchanged part in FABAC, the increment of system usability in FABAC is no doubt positively correlated with  $P_{fuzzy}$ .

Given a request  $req$  initiated by a subject  $sub$ , assuming its overall membership degree  $\mu_{all}(req) = x$ , then the probability that  $crd_{sub} > \mu_{all}(req)$  (which means  $req$  will be granted) must subject to a distribution function  $a(x)$ . After that, we assume that the number of requests satisfying  $\mu_{all} = x$  follows a distribution  $f(x)$ . Now we can get an expression for  $P_{fuzzy}$  from  $f(x)$  and  $a(x)$ , as shown below:

$$P_{fuzzy} \propto \int_T^1 a(x)f(x)dx \tag{8}$$

Obviously, as  $a(x) > 0$ , a lower  $T$  will make more requests granted. i.e., the usability  $U$  is negatively associated with  $T$  (and when  $T \rightarrow 1$ , the FABAC model will degenerate into the normal ABAC).

Besides, the fuzzy mechanism gets higher usability at the limited expense of security. The main threshold  $T$  filter out a part of substandard requests. Furthermore, as a remediation, we also introduce the credit mechanism to ensure the security of FABAC. Through the above mechanism, the subjects who carry out too many fuzzy accesses will be restricted more strictly on its future requests. As a result, no one can abuse his credits for continual illegal accesses. Hence, the FABAC still keep reasonable security.

In conclusion, by extending the granting domain with fuzzy mechanism, the FABAC model gets higher usability compared with the ABAC model with limited expense of security.

## 6 Conclusion

In this paper, we have demonstrated how to upgrade existing ABAC model into the novel FABAC model by incorporating fuzzy process and credit mechanism, for attaining better flexibility while keeping the risk within acceptable



bounds. Fuzzy mechanism is proposed to flexibly handle exceptions without manual intervention by assessing the matching degree of access request to preset rules. Credit mechanism is presented to prevent fuzzy mechanism from causing excessive security risks by systematically evaluating the fuzzy decision histories. Then we elaborate why this model is more suitable for nomadic, dynamic and distributed scenarios through a concrete case. Furthermore, the usability and security of our FABAC model is analyzed, which finally shows that our scheme does achieve a better balance.

For our future work, we would like to implement this model with XACML (eXtensible Access Control Markup Language) on our medical big data platform. Furthermore, we also plan to experiment with our solution to investigate the full impacts of model parameters.

**Acknowledgments.** This work is supported in part by the scholarship from China Scholarship Council under the Grant 201506370106, Hunan Provincial Innovation Foundation for Postgraduate under the Grant CX2015B047, the National Natural Science Foundation of China under Grants 61632009 and 61472451, and the Joint Project of Central South University and Shenzhen Tencent Computer Systems CO., LTD.

## References

1. Dinh, H.T., Lee, C., Niyato, D., Wang, P.: A survey of mobile cloud computing: architecture, applications, and approaches. *Wirel. Commun. Mob. Comput.* **13**(18), 587–1611 (2013)
2. Li, N.: Discretionary access control. In: van Tilborg, H.C.A., Jajodia, S. (eds.) *Encyclopedia of Cryptography and Security*, pp. 353–356. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-1-4419-5906-5\\_798](https://doi.org/10.1007/978-1-4419-5906-5_798)
3. Lindqvist, H.: Mandatory access control. Master's thesis, Umea University, Sweden (2006)
4. Sandhu, R.S., Coyne, E.J., Feinstein, H.L., Youman, C.E.: Role-based access control models. *IEEE Comput.* **29**(2), 38–47 (1996)
5. Hu, C., Ferraiolo, D., Kuhn, R., Schnitzer, A., Sandlin, K., Scarfone, K.: Guide to attribute based access control (ABAC) definition and considerations (draft). NIST Special Publication 800-162 (2014)
6. Li, X., Feng, D., Chen, Z., Fang, Z.: Model for attribute based access control. *J. Commun.* **29**(4), 90–98 (2008). (in Chinese)
7. Jin, X.: Attribute-based access control models and implementation in cloud infrastructure as a service. Ph.D. dissertation, The University of Texas at San Antonio, America (2014)
8. Sookhak, M., Yu, F.R., Khan, M.K., Xiang, Y., Buyya, R.: Attribute-based data access control in mobile cloud computing: taxonomy and open issues. *Future Gener. Comput. Syst.* **72**, 273–287 (2017). Elsevier
9. Ngo, C., Demchenko, Y., de Laat, C.: Multi-tenant attribute-based access control for cloud infrastructure services. *J. Inf. Secur. Appl.* **27**, 65–84 (2016). Elsevier
10. Axiomatics. <https://www.axiomatics.com/>
11. Cheng, P.C., Rohatgi, P., Keser, C., Karger, P.A., Wagner, G.M., Reninger, A.S.: Fuzzy multi-level security: an experiment on quantified risk-adaptive access control. In: *Proceedings of IEEE Symposium on Security and Privacy*, pp. 222–230. IEEE (2007)

12. Martínez-García, C., Navarro-Arribas, G., Borrell, J.: Fuzzy role-based access control. *Inf. Process. Lett.* **111**(10), 483–487 (2011). Elsevier
13. Dimmock, N., Belokosztolszki, A., Eyers, D., Bacon, J., Moody, K.: Using trust and risk in role-based access control policies. In: Proceedings of 9th ACM Symposium on Access Control Models and Technologies, pp. 156–162. ACM (2004)
14. Mahalle, P.N., Thakre, P.A., Prasad, N.R., Prasad, R.: A fuzzy approach to trust based access control in internet of things. In: Proceedings of 3rd International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE), pp. 1–5. IEEE (2013)
15. Feng, F., Lin, C., Peng, D., Li, J.: A trust and context based access control model for distributed systems. In: Proceedings of 10th IEEE International Conference on High Performance Computing and Communications, pp. 629–634. IEEE (2008)
16. Bhatti, R., Bertino, E., Ghafoor, A.: A trust-based context-aware access control model for web-services. *Distrib. Parallel Databases* **18**(1), 83–105 (2005). Springer
17. Zadeh, L.A.: Fuzzy sets. *Inf. Control* **8**(3), 338–353 (1965). Elsevier
18. Servos, D., Osborn, S.L.: Current research and open problems in attribute-based access control. *ACM Comput. Surv.* **49**(4), 65–107 (2017)
19. Mamdani, E.H., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man-Mach. Stud.* **7**(1), 1–13 (1975)
20. McKendrick, A., Pai, M.K.: XLV.—the rate of multiplication of micro-organisms: a mathematical study. *Roy. Soc. Edinb.* **31**, 649–653 (1912). Cambridge

# Security Analysis and Improvement of An Anonymous Attribute-Based Proxy Re-encryption

Hongjian Yin and Leyou Zhang(✉)

School of Mathematics and Statistics, Xidian University, Xi'an 710126, China  
lyzhang@mail.xidian.edu.cn

**Abstract.** The ciphertext-policy attribute-based proxy re-encryption (CP-AB-PRE) is a flexible proxy re-encryption (PRE), which makes the encryptor control its encrypted data at a fine-grained level and update the access policy. However, most of constructions focuses only on the data security, rather than on user privacy protection. In order to protect users' attribute privacy, recently, a novel secure CP-AB-PRE named anonymous CP-AB-PRE was first proposed by Zhang et al. However, we found that their scheme fails to achieve anonymity, which means that their scheme cannot realize users' attribute privacy protection. In order to remedy this security gap, a novel anonymous CP-AB-PRE scheme is proposed, which can protect user attribute privacy by hiding the access policy. Theoretical analysis and simulation results demonstrate that our proposed scheme is secure and efficient.

**Keywords:** Attribute-based proxy re-encryption · Anonymity  
Security · Privacy protection · Ciphertext-policy

## 1 Introduction

With the development of cloud computing, it has become a trend that more and more companies and individual users store their sensitive data by third parties such as Amazon, Google and Alibaba. In order to guarantee data confidentiality, these data should be encrypted before uploading. However, in many cases, it requires complex access control for protected data, that is, only can the user who satisfies some attributes get the protected data. Sahai and Waters [1] solved this issue by introducing the attribute-based encryption (ABE). Generally, there are two kinds of ABE involving ciphertext-policy ABE (CP-ABE) and key-policy ABE (KP-ABE). In a CP-ABE scheme, the user secret keys are associated with a set of attributes and ciphertext is embedded with an access structure. The user will be able to decrypt the ciphertext only if the attributes associated with his secret keys satisfy the access structure of the ciphertext. In a KP-ABE scheme, users' secret keys are embedded with an access structure and ciphertext is associated with a set of attributes. Successful decryption of the ciphertext requires if

and only if the ciphertext attribute set satisfies the access structure associated with the user secret keys.

In 1998, Blaze et al. [5] proposed the first PRE scheme in public key cryptosystem. In their scheme, it allows a proxy to translate a ciphertext encrypted under Alice's public key into one that can be decrypted by Bob's secret key. In 2008, Guo et al. [9] proposed the first attribute-based proxy re-encryption (AB-PRE) scheme in the type of key policy. Subsequently, the first ciphertext policy attribute-based proxy re-encryption scheme was proposed by Liang et al. [10]. In CP-AB-PRE schemes, a proxy is specified by users in advance. The proxy can receive a proxy re-encryption key to transform a ciphertext associated with an access policy into another one with a new policy. After the first CP-AB-PRE scheme, many PRE schemes based on CP-ABE have been presented [11, 18, 19]. In these schemes, access structures are embedded in the ciphertext and whoever obtains the ciphertext can see the content of the access structure. However, in some cases, the access structure may be sensitive. For example, in military applications, the access structures may contain military secrets, such as sensitive information like organizational structure.

In order to protect users' privacy, Zhang et al. [15] proposed an anonymous CP-AB-PRE recently. In this scheme, they claimed that their scheme was the first one that made access structure hidden and access policy update simultaneously. Unfortunately, we found that their scheme cannot achieve anonymity, in other words, any malicious user can only utilize public parameters to determine whether the ciphertexts are encrypted under the given access structure or not.

In this paper, we first give the security analysis of Zhang et al.'s scheme and show that their scheme does not realize users' attribute privacy protection. Then we propose an improvement scheme to remedy their security gap. In our proposed scheme, the attribute privacy is protected by hiding the access policy. In the standard model, the proposed scheme is proved to be secure.

## 2 Preliminaries

### 2.1 Hardness Assumptions

**Decisional Bilinear Diffie-Hellman (DBDH) assumption.** Let  $a, b, c, z \in_R \mathbb{Z}_p$ , and  $g \in_R \mathbb{G}$  be a generator. The DBDH assumption holds in group  $\mathbb{G}$  if no probabilistic polynomial-time (PPT) algorithm can distinguish the tuple  $[g, g^a, g^b, g^c, e(g, g)^{abc}]$  from  $[g, g^a, g^b, g^c, g^z]$  with non-negligible advantage.

**The Decision Linear (D-Linear) Assumption.** Suppose  $g \in_R \mathbb{G}$  is a generator and  $a_1, a_2, a_3, a_4, z \in_R \mathbb{Z}_p$ . The D-Linear assumption is said to hold in  $\mathbb{G}$  if no PPT algorithm can distinguish the tuple  $[g, g^{a_1}, g^{a_2}, g^{a_1 a_3}, g^{a_2 a_4}, g^{a_3 + a_4}]$  from  $[g, g^{a_1}, g^{a_2}, g^{a_1 a_3}, g^{a_2 a_4}, g^z]$  with an advantage non-negligible.

**The Computational Bilinear Diffie-Hellman (CBDH) Assumption.** Suppose  $g \in_R \mathbb{G}$  is a generator and  $a, b, c \in_R \mathbb{Z}_p$ . We say that the CBDH assumption holds in  $\mathbb{G}$  if given  $[g, g^a, g^b, g^c]$ , no PPT algorithm can compute  $e(g, g)^{abc}$  with an advantage non-negligible.

## 2.2 Definitions Anonymous of CP-AB-PRE Scheme

An anonymous CP-AB-PRE scheme consists of the following six algorithms:

**Setup:** This algorithm takes the security parameter  $\lambda$  as input and generates a public key  $PK$ , a master secret key  $MK$ .

**KeyGen:** This algorithm takes  $MK$  and a set of attributes  $L$  as input and generates a secret key  $SK_L$  associated with  $L$ .

**Encrypt:** This algorithm takes  $PK$ , a message  $\mathcal{M}$ , and an access policy  $W$  as input, and generates a ciphertext  $CT_W$ .

**RKGen:** This algorithm takes a secret key  $SK_L$  and an access policy  $W'$  as input and generates a re-encryption key  $RK_{L \rightarrow W'}$ .

**Reencrypt:** This algorithm takes a re-encryption key  $RK_{L \rightarrow W'}$  and a ciphertext  $CT_W$  as input, first checks whether the attribute list in  $RK_{L \rightarrow W'}$  satisfies the access policy of  $CT_W$  or not. If check passes, it generates a re-encrypted ciphertext  $CT'_{W'}$ ; otherwise, it returns  $\perp$ .

**Decrypt:** This algorithm takes  $CT_W$  and  $SK_L$  associated with  $L$  as input and returns the message  $\mathcal{M}$  if the attribute list  $L$  satisfies the access policy  $W$  specified for  $CT_W$ . If  $L \not\models W$ , it returns  $\perp$ .

## 2.3 Security Model

The indistinguishability against selective ciphertext-policy and chosen-plaintext attacks (IND-sCP-CPA) model [2,3,5] and the selective master key security (sMKS) model [6] are described as follows.

### IND-sCP-CPA Game

**Init:**  $\mathcal{A}$  submits two challenge ciphertext policies  $W_0^*$  and  $W_1^*$  to the challenger.

**Setup:** The challenger runs the Setup algorithm and outputs the public key  $PK$  to the adversary  $\mathcal{A}$ .

**Phase 1:** The adversary  $\mathcal{A}$  queries the following oracles in polynomial time:

(i) KeyGen oracle:  $\mathcal{A}$  submits an attribute list  $L$ , the challenger returns  $SK_L$  if  $(L \not\models W_0^* \wedge L \not\models W_1^*)$  or  $(L \models W_0^* \wedge L \models W_1^*)$ . Otherwise, it outputs  $\perp$ .

(ii) RKGen oracle: The adversary  $\mathcal{A}$  submits  $L$  and  $W$ , if  $(L \not\models W_0^* \wedge L \not\models W_1^*)$  or  $(L \models W_0^* \wedge L \models W_1^*)$ , the challenger returns  $\mathcal{A}$  the re-encryption key  $RK_{L \rightarrow W}$ . Otherwise, it outputs  $\perp$ .

(iii) Reencrypt oracle:  $\mathcal{A}$  submits  $L$ ,  $W'$  and an anonymous ciphertext  $CT_W$  under an access policy  $W$ , if  $((L \not\models W_0^* \wedge L \not\models W_1^*)$  or  $(L \models W_0^* \wedge L \models W_1^*)$ ) and  $L \models W$ , the challenger gives  $\mathcal{A}$   $CT_{W'}$ . Otherwise, it outputs  $\perp$ .

**Challenge:** Once Phase 1 is over,  $\mathcal{A}$  outputs two equal length messages  $\mathcal{M}_0$  and  $\mathcal{M}_1$ . It is required that  $\mathcal{M}_0 = \mathcal{M}_1$  if any secret key on  $L$  satisfying  $L \models W_0^* \wedge L \models W_1^*$  has been queried. The challenger randomly chooses a bit  $\nu \in \{0, 1\}$ , computes  $CT_{W_\nu^*} = \text{Encrypt}(PK, \mathcal{M}_\nu, W_\nu^*)$  and sends  $CT_{W_\nu^*}$  to  $\mathcal{A}$ , where  $W_\nu^*$  is hidden.

**Phase 2:** It is similar to Phase 1.

**Guess:**  $\mathcal{A}$  outputs a bit  $\nu' \in \{0, 1\}$  as a guess of  $\nu$  and it wins the aforementioned game if  $\nu' = \nu$ .

In the above game, we define the advantage of  $\mathcal{A}$  as  $Adv_0^{\mathcal{A}} = |Pr[\nu' = \nu] - 1/2|$ . In this model, it is required in the challenge phase that  $\mathcal{M}_0 = \mathcal{M}_1$  if the adversary obtains a secret key  $SK_L$  matching both  $W_0^*$  and  $W_1^*$ . Otherwise, the adversary can directly decrypt the challenge ciphertext to get the bit value  $\nu$ . Hence, it easily follows that the security models of anonymous CP-AB-PRE would make no sense without such a restriction.

### sMKS Game

**Init:**  $\mathcal{A}$  submits an attribute list  $L^*$  to the challenger.

**Setup:** The same as that of IND-sCP-CPA game.

**Queries:**  $\mathcal{A}$  queries the following oracles in polynomial time:

- (i) KeyGen oracle:  $\mathcal{A}$  submits an attribute list  $L$ , if  $L \neq L^*$ ,  $SK_L$  is returned by the challenger to  $\mathcal{A}$ . Otherwise, it outputs  $\perp$ .
- (ii) RKGen oracle:  $\mathcal{A}$  submits  $L$  and  $W$ , the challenger generates a proxy re-encryption key  $RK_{L \rightarrow W}$  for  $\mathcal{A}$ .
- (iii) Reencrypt oracle:  $\mathcal{A}$  submits  $L$ ,  $W'$ , and  $CT_W$  under an access policy  $W$ , the challenger returns  $CT'_W$  as a re-encryption ciphertext to  $\mathcal{A}$ . Note that, the access policy  $W$  is not revealed in  $CT'_W$  to achieve anonymity.

**Output:**  $\mathcal{A}$  outputs an attribute secret key  $SK_{L^*}$ , and it succeeds if  $SK_{L^*}$  is valid for  $L^*$ .

In this game, the advantage of  $\mathcal{A}$  is defined  $Adv_1^{\mathcal{A}} = Pr[\mathcal{A} \text{ succeeds}]$ .

**Definition 1.** *An anonymous CP-AB-PRE scheme is selective master key security if no probabilistic polynomial time adversary  $\mathcal{A}$  has a non-negligible advantage in winning the master key security game.*

## 3 Security Analysis of Zhang et al.'s Scheme

In this section, we will show that Zhang et al.'s scheme has the weakness of anonymity, in other words, it cannot realize ciphertext policy hidden. In the following, we explain why the above scheme is not anonymous.

Some parts of ciphertexts and public keys  $g$ ,  $T_{i,t}$ ,  $C_{i,t,\Delta}$  and  $C'_0$  is a Decision DiffieHellman (DDH) tuple, from which the attribute information  $T_{i,t}$  can be revealed. More precisely, for an attribute  $T_{i,t}$ , an attacker can run the DDH test  $e(\prod_{i \in W} C_{i,t,\Delta}, g) \stackrel{?}{=} e(\prod_{i \in W'} T_{i,t}, C'_0)$  because  $g$  and  $T_{i,t}$  are public keys, then the attacker can determine whether  $W'$  is used in ciphertext or not, that is, for attribute  $T_{i,t}$  or not. In conclusion, the DDH test attack work successfully due to  $C_{i,t,\Delta}$  and  $C'_0$  which are used in matching phase only.

## 4 An Improved Scheme and Proof of Security

### 4.1 An Improved Scheme

From the security analysis, we can know that, the main issue is in original match-then-re-encrypt technique. To fill this security gap, we propose a novel

match-then-re-encrypt technique replace the one used in Zhang et al.'s scheme. From this simple modification, our scheme achieves anonymity without losing any feature of anonymous CP-AB-PRE. The following is our construction.

**Setup** ( $1^\lambda$ ). Attribute authority chooses  $g_2, g_3 \in_R \mathbb{G}$  and  $y \in_R \mathbb{Z}_p$ . For each attribute  $w_i$  with  $1 \leq i \leq n$ , it chooses  $a_{i,t}, b_{i,t}, \tau_{i,t} \in_R \mathbb{Z}_p$  and computes  $T_{i,t} = g^{\tau_{i,t}}$ ,  $g_1 = g^y$ ,  $Y = e(g_1, g_2)$ ,  $A_{i,t} = T_{i,t}^{a_{i,t}}$ ,  $B_{i,t} = T_{i,t}^{b_{i,t}}$ , where  $1 \leq i \leq n, 1 \leq t \leq n_i$ . Then it chooses  $w \in_R \mathbb{G}$ ,  $\alpha, t_0, t_1, t_2, t_3 \in_R \mathbb{Z}_p$ , and  $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in_R \mathbb{Z}_p$  and computes  $v = g^{t_1}$ ,  $u = g^{t_2}$ ,  $V = g^{t_3}$ ,  $w = g^{t_0}$ ,  $g_0 = v^\alpha$ ,  $g_{0i} = g^{\beta_i}$ ,  $g_{1i} = g^{\alpha_i}$ ,  $A = e(g_0, g_2)$ . The system public key  $PK = \langle g, g_0, g_1, g_2, g_3, Y, A, u, v, V, w, \{g_{0i}, g_{1i}, \{T_{i,t}, A_{i,t}, B_{i,t}\}_{1 \leq t \leq n_i}\}_{1 \leq i \leq n} \rangle$  and master key  $MK = \langle g_2^y, y, \{\alpha_i, \beta_i, \{a_{i,t}, b_{i,t}, \tau_{i,t}\}_{1 \leq t \leq n_i}\}_{1 \leq i \leq n} \rangle$ .

**KeyGen** ( $PK, MK, L$ ). After receiving an attribute  $A^* = [A_1^*, A_2^*, \dots, A_n^*]$ , where  $A_i^* \in \{0, 1\}$  with  $i \in [1, n]$ , the attribute authority computes  $h_i = (h_{i-1})^{\alpha_i^* \beta_i^{1-A_i^*}}$ ,  $d_\Delta = g_3^{\hat{d}}$ ,  $d_0 = g_2^s h_n^\gamma$ ,  $d_1 = v^\gamma$ ,  $d_2 = w^\gamma$ ,  $d_3 = u^{\gamma t_1}$  and  $d_4 = V^{\gamma t_1}$  where  $\gamma \in_R \mathbb{Z}_p$ ,  $h_0 = g$ . Then it picks random elements  $\{r_i, \lambda_i \in_R \mathbb{Z}_p\}_{1 \leq i \leq n}$ , sets  $r = \sum_{i=1}^n r_i$ . For  $1 \leq i \leq n$ , it computes  $[D_0, D_{i,0}, D_{i,1}, D_{i,2}] = [g_2^{y-r}, g_2^{r_i} B_{i,k_i}^{\lambda_i \alpha_i, k_i}, g^{\lambda_i \alpha_i, k_i}, g^{\lambda_i b_i, k_i}]$ . Then,  $SK_L = \langle d_\Delta, d_0, d_1, d_2, d_3, d_4, D_0, \{D_{i,0}, D_{i,1}, D_{i,2}\}_{1 \leq i \leq n} \rangle$ .

**Encrypt** ( $PK, M, W$ ). To encrypt  $M \in \mathbb{G}_T$  under an access policy  $W = [W_1, W_2, \dots, W_n]$ , data owner chooses  $s, s', s_{1i}, s_{2i} \in_R \mathbb{Z}_p$ ,  $s_1 = \sum_{i=1}^n s_{1i}$ ,  $s_2 = \sum_{i=1}^n s_{2i}$  and computes  $\tilde{C} = MY^s$ ,  $C_0 = g^s$ ,  $C_{RE} = g_3^s$ ,  $C'_i = h_n^{s_{1i}} u^{s_{2i}}$ ,  $C'_0 = A_1^s$ ,  $C'_1 = w^{s_1} V^{s_2}$ ,  $C'_2 = g^{s_2}$ ,  $C'_3 = v^{s_1}$ . And data owner computes  $[C_{i,t,1}, C_{i,t,2}] = [B_{i,t}^{s-s'}, A_{i,t}^{s'}]$ , if  $v_{i,t} \in W_i$ . Otherwise,  $[C_{i,t,\Delta}, C_{i,t,1}, C_{i,t,2}]$  are random elements in  $\mathbb{G}$ . Then the ciphertext of  $M$  with respect  $W$  is  $CT_W = \langle C_0, C_{RE}, \tilde{C}, C'_0, C'_1, C'_2, C'_3, \{\{C_{i,t,1}, C_{i,t,2}\}_{1 \leq t \leq n_i}, C'_i\}_{1 \leq i \leq n} \rangle$ .

**RKGen** ( $SK_L, W'$ ). This algorithm takes  $SK_L$  and  $W'$  as input. Then it chooses  $d \in_R \mathbb{Z}_p$  and computes  $g^d$ ,  $\hat{d}_\Delta = d_\Delta g_3^d$ ,  $\hat{d}_0 = d_0$ ,  $\hat{d}_1 = d_1$ ,  $\hat{d}_2 = d_2$ ,  $\hat{d}_3 = d_3 d_\Delta$ ,  $\hat{d}_4 = d_4 g_3^d$ ,  $\hat{D}_0 = D_0$ ,  $\hat{D}_{i,0} = D_{i,0} g_3^d$ ,  $\hat{D}_{i,1} = D_{i,1}$ ,  $\hat{D}_{i,2} = D_{i,2}$ ,  $\mathbb{C} = \text{Encrypt}(PK, E(g^d), W')$ . Then, the proxy re-encryption key corresponding  $W'$  is  $RK_{L \rightarrow W'} = \langle \hat{D}_\Delta, \{\hat{D}_{i'}\}_{i'=1,2,3,4}, \{\hat{D}_0, \hat{D}_{i,0}, \hat{D}_{i,1}, \hat{D}_{i,2}\}_{1 \leq i \leq n}, \mathbb{C} \rangle$ .

**Reencrypt** ( $RK_{L \rightarrow W'}, CT_W$ ). Upon receiving  $RK_{L \rightarrow W'}$  for  $W'$ , and  $CT_W$  under  $W$ , proxy server does the following without knowing  $W$ :

**Matching Phase:** The proxy server computes  $C'_\Delta = e(C'_2, \hat{d}_\Delta)$  and checks whether  $L \models W$  in terms of the following Equation (1). Specifically,  $L \models W$  if and only if Equation (1) holds

$$C'_0 C'_\Delta = \frac{e(\hat{d}_0, C'_3) e(\hat{d}_3 \hat{d}_4, C'_2) e(\hat{d}_2, C'_3)}{e(C'_1, \hat{d}_1) e(\prod_{i=1}^n C'_i, \hat{d}_1)}, \quad (1)$$

where suppose the indexes satisfy  $L_i = v_{i, k_i}$ . It returns  $\perp$  if  $L \not\models W$ . Otherwise, it initiates the Re-encryption Phase.

**Re-encryption Phase:** The proxy server computes

$$E_i = \frac{e(C_0, \hat{D}_{i,0})}{e(C_{i,t,1}, \hat{D}_{i,1}) e(C_{i,t,2}, \hat{D}_{i,2})} = e(g, g_2)^{sr_i} e(g, g_3)^{sd}. \quad (2)$$

Subsequently, it computes  $\tilde{C} = e(C_0, D'_0) \prod_{i=1}^n E_i = e(g, g_2)^{ys} e(g, g_3)^{nsd}$ , and outputs a proxy re-encryption ciphertext  $CT_{W'} = \langle \tilde{C}, C_{RE}, \tilde{C}, \mathbb{C} \rangle$ . It follows from the subsequent Decrypt algorithm that the decryptor of re-encryption ciphertext only needs  $g^d$  to decrypt the proxy re-encryption ciphertext. Thus, we can obtain a two-time proxy re-encryption ciphertext  $CT_{W''} = \langle \tilde{C}, C_{RE}, \tilde{C}, \mathbb{C}' \rangle$ , where  $\mathbb{C}'$  is generated based on the algorithm Reencrypt with the ciphertext  $\mathbb{C}$  and another proxy re-encryption key  $RK_{L' \rightarrow W''}$  as inputs. Specifically,  $\mathbb{C}' = \text{Reencrypt}(PK, RK_{L' \rightarrow W''}, \mathbb{C})$ . Similarly, the multiple time proxy re-encryption ciphertexts can be generated.

**Decrypt** ( $CT_W, SK_L$ ). The ciphertext  $CT_W$  is tested and decrypted by a user with secret key  $SK_L$  as follows:

1. If  $CT_W$  is an original ciphertext, the user does the following:

**Matching Phase:** The user checks whether  $L \models W$  in terms of the following Eq. (3). Suppose the indexes satisfy  $L_i = v_{i,k_i}$ ,  $L \models W$  if and only if Eq. (3) holds

$$C'_0 = \frac{e(d_0, C'_3)e(d_3d_4, C'_2)e(d_2, C'_3)}{e(C'_1, d_1)e(\prod_{i=1}^n C'_i, d_1)}. \quad (3)$$

If  $L \not\models W$ , it returns  $\perp$ . Otherwise, it initiates the Decryption Phase.

**Decryption Phase:** Suppose  $L_i = v_{i,k_i}$ , the user computes

$$\mathcal{M} = \frac{\tilde{C} \prod_{i=1}^n e(C_{i,k_i,1}, D_{i,k_i,1})e(C_{i,k_i,2}, D_{i,k_i,2})}{e(C_0, D_0) \prod_{i=1}^n e(C_0, D_{i,0})} \quad (4)$$

2. Else if  $CT_W$  is a one-time proxy re-encryption ciphertext consists of  $\langle \tilde{C}, C_{RE}, \tilde{C}, \mathbb{C} \rangle$ , the user does the following:

**Matching Phase:** The user checks whether  $L \models W$  in accordance with  $\mathbb{C}$  by using the method in Eq. (3). If  $L \not\models W$ , the algorithm returns  $\perp$ . Otherwise, it initiates the Decryption Phase.

**Decryption Phase:** The user does

- Performs a decryption of the original ciphertext  $\mathbb{C}$  of  $E(g^d)$  using the secret key  $SK_L$  and decodes it to  $g^d$ .
  - Computes  $\frac{\tilde{C}e(C_{RE}, g^d)^n}{\tilde{C}} = \mathcal{M}$ .
3. Else, if  $CT_W = \langle \tilde{C}, C_{RE}, \tilde{C}, \mathbb{C}' \rangle$  is an  $N + 1$  time proxy re-encryption ciphertext, where  $\mathbb{C}'$  is an  $N$  time proxy re-encryption ciphertext of  $E(g^d)$ , then the user does the following:
    - Performs a decryption of the  $N$  time proxy re-encryption ciphertext  $\mathbb{C}'$ . If the algorithm does not return  $\perp$ , the user recovers  $E(g^d)$ , decodes it to  $g^d$  and proceeds.
    - Computes  $\frac{\tilde{C}e(C_{RE}, g^d)^n}{\tilde{C}} = \mathcal{M}$ .

## 4.2 Proof of Security

**Theorem 1.** *The anonymous CP-AB-PRE scheme is secure in the IND-sCP-CPA model, under the DBDH assumption and the D-Linear assumption.*



**Theorem 2.** *The anonymous CP-AB-PRE scheme is sMKS secure Under the CBDH assumption.*

The Theorems 1 and 2 can be proved secure according to [15]. Due to space limitations, detailed proofs will be given in the full version.

## 5 Performance Analysis

In this section, we compare our work with previous work with regard to security and efficiency. For convenience,  $|PK|$ ,  $|SK|$ ,  $|OCT|$  and  $|CT|$  denote the size of the public key, the secret key, the original ciphertext and the re-encryption ciphertext.  $|G|$  and  $|G_T|$  denote the bit-length of the elements in  $\mathbb{G}$  and  $\mathbb{G}_T$ . Let  $n$  be the number of all attributes in the system,  $N$  express the total number of possible values of all attributes.  $m$  is the number of attribute held by user and  $k$  denotes the total number of attributes required by the ciphertext (Tables 1 and 2).

**Table 1.** Security comparison among different AB-PRE schemes

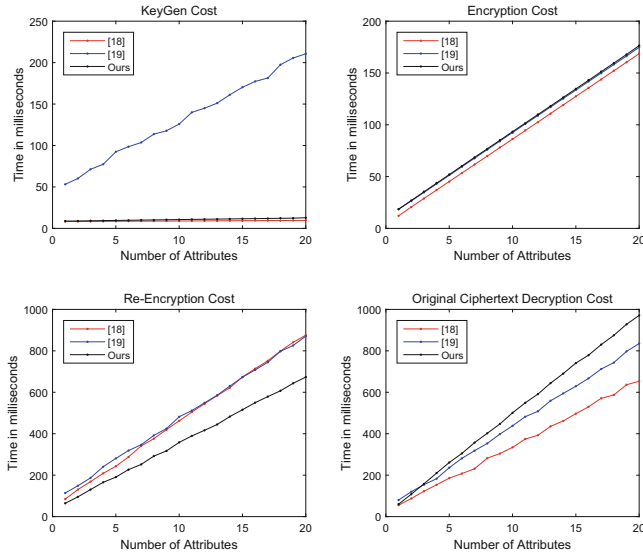
Scheme	Access structure	Anonymity	Security	Hardness
[14]	LSSS	No	Adaptive	DBDH, SUF.Sig*
[15]	AND	No	Selective	DBDH, D-Linear, CBDH
[19]	LSSS	No	Adaptive	$q$ -parallel BDHE
Ours	AND	Yes	Selective	DBDH, D-Linear, CBDH

\*Strong unforgeability of one-time signature.

**Table 2.** Communication cost comparison among different AB-PRE schemes

Scheme	$ PK $	$ SK $	$ OCT $	$ RCT $
[15]	$(3N + 4) G  +  G_T $	$(4m + 4) G $	$(3k + 3) G  + 2 G_T $	$ G  + 3 G_T $
[19]	$(n + 6) G  +  G_T $	$(m + 3) G $	$(2k + 5) G  +  G_T $	$(2k + 5) G  +  G_T $
Ours	$(3N + n + 9) G  +  G_T $	$(3m + 7) G $	$(3k + 5) G  + 2 G_T $	$ G  + 3 G_T $

Our proposed construction is efficient in the size of the re-encryption ciphertext. Although, other schemes are little more efficient in the size of the public key, the secret key and the original ciphertext, these schemes cannot achieve user privacy protection. In order to simulate, we use the pairing-based cryptography library. The simulation experiment is done on a Windows machine with 2.67 GHz Intel(R) Core(TM) 2 Quad CPU and 4 GB ROM. The Fig. 1 shows that the re-encryption cost of our scheme is better than other schemes. In addition, our Key-Gen cost and Encryption cost is good. However, in order to achieve anonymity, it costs more time to decrypt the original ciphertext.



**Fig. 1.** Computational costs comparison among different CP-AB-PRE schemes

## 6 Conclusions

In this paper, we give the security analysis of Zhang et al.'s scheme and show why their scheme does not realize users' attribute privacy protection. In order to remedy this security gap, a novel scheme is proposed based on the match-then-re-encrypt and match-then-re-decrypt technique. In our scheme, the attribute privacy is protected in access policy. And it is efficient in re-encryption ciphertext size. From the analysis of safety and efficiency, the improved scheme achieves anonymity without losing any feature of Zhang et al.'s scheme.

## References

1. Sahai, A., Waters, B.: Fuzzy identity-based encryption. In: Cramer, R. (ed.) EUROCRYPT 2005. LNCS, vol. 3494, pp. 457–473. Springer, Heidelberg (2005). [https://doi.org/10.1007/11426639\\_27](https://doi.org/10.1007/11426639_27)
2. Bethencourt, J., Sahai, A., Waters, B.: Ciphertext-policy attribute-based encryption. In: IEEE Symposium on Security and Privacy, pp. 321–334 (2007)
3. Goyal, V., Pandey, O., Sahai, A., et al.: Attribute-based encryption for fine-grained access control of encrypted data. In: ACM Conference on Computer and Communications Security, pp. 89–98 (2006)
4. Boneh, D., Waters, B.: Conjunctive, subset, and range queries on encrypted data. In: Vadhan, S.P. (ed.) TCC 2007. LNCS, vol. 4392, pp. 535–554. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-70936-7\\_29](https://doi.org/10.1007/978-3-540-70936-7_29)
5. Blaze, M., Bleumer, G., Strauss, M.: Divertible protocols and atomic proxy cryptography. In: Nyberg, K. (ed.) EUROCRYPT 1998. LNCS, vol. 1403, pp. 127–144. Springer, Heidelberg (1998). <https://doi.org/10.1007/BFb0054122>

6. Ateniese, G., Fu, K., Green, M., et al.: Improved proxy re-encryption schemes with applications to secure distributed storage. *ACM Trans. Inf. Syst. Secur.* **9**(1), 1–30 (2006)
7. Canetti, R., Hohenberger, S.: Chosen-ciphertext secure proxy re-encryption. In: *ACM Conference on Computer and Communications Security*, pp. 185–194 (2007)
8. Green, M., Ateniese, G.: Identity-based proxy re-encryption. In: Katz, J., Yung, M. (eds.) *ACNS 2007*. LNCS, vol. 4521, pp. 288–306. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-72738-5\\_19](https://doi.org/10.1007/978-3-540-72738-5_19)
9. Guo, S., Zeng, Y., Wei, J., et al.: Attribute-based re-encryption scheme in the standard model. *Wuhan Univ. J. Nat. Sci.* **13**(5), 621–625 (2008)
10. Liang, X., Cao, Z., Lin, H., et al.: Attribute based proxy re-encryption with delegating capabilities. In: *International Symposium on Information, Computer, and Communications Security*, pp. 276–286 (2009)
11. Luo, S., Hu, J., Chen, Z.: Ciphertext policy attribute-based proxy re-encryption. In: Soriano, M., Qing, S., López, J. (eds.) *ICICS 2010*. LNCS, vol. 6476, pp. 401–415. Springer, Heidelberg (2010). [https://doi.org/10.1007/978-3-642-17650-0\\_28](https://doi.org/10.1007/978-3-642-17650-0_28)
12. Liu, Q., Wang, G., Wu, J.: Time-based proxy re-encryption scheme for secure data sharing in a cloud environment. *Inf. Sci.* **258**(3), 355–370 (2014)
13. Zhang, L., Wu, Q., Mu, Y., et al.: Privacy-preserving and secure sharing of PHR in the cloud. *J. Med. Syst.* **40**(12), 1–13 (2016)
14. Kawai, Y., Takashima, K.: Fully-anonymous functional proxy-re-encryption. *IACR Cryptology EPrint Archive 2013*, p. 318 (2013)
15. Zhang, Y., Li, J., Chen, X., et al.: Anonymous attribute-based proxy re-encryption for access control in cloud computing. *Secur. Commun. Netw.* **9**(14), 2397–2411 (2016)
16. Shao, J.: Anonymous ID-based proxy re-encryption. In: Susilo, W., Mu, Y., Seberry, J. (eds.) *ACISP 2012*. LNCS, vol. 7372, pp. 364–375. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-31448-3\\_27](https://doi.org/10.1007/978-3-642-31448-3_27)
17. Abdalla, M., Catalano, D., Fiore, D.: Verifiable random functions from identity-based key encapsulation. In: Joux, A. (ed.) *EUROCRYPT 2009*. LNCS, vol. 5479, pp. 554–571. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-01001-9\\_32](https://doi.org/10.1007/978-3-642-01001-9_32)
18. Liang, K., Fang, L., Susilo, W., et al.: A ciphertext-policy attribute-based proxy re-encryption with chosen-ciphertext security. In: *International Conference on Intelligent Networking and Collaborative Systems*, pp. 552–559 (2013)
19. Liang, K., Man, H.A., Liu, J.K., et al.: A secure and efficient ciphertext-policy attribute-based proxy re-encryption for cloud data sharing. *Future Gener. Comput. Syst.* **52**(C), 95–108 (2015)
20. Zhang, Y., Chen, X., Li, J., et al.: Anonymous attribute-based encryption supporting efficient decryption test. In: *ACM SIGSAC Symposium on Information, Computer and Communications Security*, pp. 511–516 (2013)

# Relacha: Using Associative Meaning for Image Captcha Understandability

Songjie Wei<sup>1</sup>(✉), Qianqian Wu<sup>1</sup>, and Milin Ren<sup>2</sup>

<sup>1</sup> School of Computer Science and Engineering,  
Nanjing University of Science and Technology, Nanjing 210094, China  
{swei, qqwu}@njust.edu.cn

<sup>2</sup> Beijing Engineering Research Center of NGI and Its Major Application  
Technologies Co. Ltd., Beijing 100084, China  
rmilin@163.com

**Abstract.** Text-based CAPTCHA has been used over decades with increasing difficulty to remain effective with OCR technique advance. Image-based CAPTCHA is supposed to step in as a better alternative. However, recognition-based image CAPTCHA is not robust enough to resist against either computer pattern recognition algorithms or brute-force attacks with exhaustivity approach. We present a new CAPTCHA design called Relacha to distinguish humans from bots by an image content correlation test. The new construction scheme adopts random walk among images with correlated contents, and utilizes human reasoning ability on inferring the relevance of images. Relacha challenges are generated dynamically by using images from real-time online search engine. The usability and robustness of the proposed scheme has been evaluated by both numerical analysis and empirical evidence. The results show that humans can solve Relacha conveniently and effectively with a high pass rate, while bot programs may succeed with slim chance.

**Keywords:** CAPTCHA · Web authentication · Image correlation  
Associative meaning · Random walk

## 1 Introduction

CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Humans Apart) is a type of interactive computer challenge based on Artificial Intelligence (AI) problems that cannot be easily solved by current computer programs or bots, but is effortlessly solvable by humans [1]. CAPTCHA has been widely deployed for preventing malicious users from conducting automated network attacks and resource abuse, such as Denial-of-Service (DoS), which may lead to fatal exhaustion of server resources. Widely-used existing CAPTCHA techniques are either text-based or image-based ones, all request a match between the challenge and the inputted or selected answers to pass the test.

Text-based CAPTCHA as the most popular ones in current Internet, has been adopted for decades [2]. However, in order to survive along the continuously improved Optical Character Recognition (OCR) capability, text-based CAPTCHA challenges

have become extremely distorted and complex, which blocks not only computer programs but also human users from recognizing text-based CAPTCHAs. Instead, image-based CAPTCHA has been introduced as an alternative for better user experience, which relies on human ability to understand visual representations, such as distinguishing images of animals or objects. The gap between human and computational ability in recognizing visual content has been termed by Smeulders et al. as the semantic gap [3]. However, recent advance in pattern recognition foresees a promising migration of this gap between human and machine performance [4], which threatens almost all of the existing match-based CAPTCHA techniques. Furthermore, most of the existing image-based CAPTCHAs typically rest on a fixed collection of images, so an exhaustive attack by automated bots is possible [5].

By further surveying the machine capability in natural language and image processing technology, we find that inferring relevance between images is still an AI-hard problem in today's technology. Recent advance in AI image processing and indexing has shown that AI algorithms can effectively index and compare images for similarities without understanding the image contents. Relacha avoids using *is-a* or *similar-as* metric when presenting CAPTCHA questions and choices, but mimics human divergent thinking to utilize the underlying associative meanings in images to pair questions with correct answers in CAPTCHA test construction and result evaluation. Understandability of such associative meanings among images more relies on context, culture and history owned by humans than in memory, calculation specialized by computers. For example, an image of "suitcase" reminds humans more about "hotel" than "food store". When putting all together as CAPTCHA images, humans tend to pass the test by choosing the more closely associated pair.

We propose to exploit human's reasoning ability on image semantic correlation to create an image-based relevance-oriented CAPTCHA named Relacha (Relevance-oriented CAPTCHA) as a reformation of the existing match-based CAPTCHA. Relacha is constructed based on the correlation and dependency of word-annotated images from real-time online search engine. The correlation degree of word tags is measured based on their frequency of occurrence in Internet contents. Challenge words are visualized as images in Relacha construction. Answer images are retrieved from the Internet on-the-fly with tags as keywords to generate a dynamic resource library. Relacha challenge is presented with a visualized text question and a layout of multiple image choices, which are selected by randomly walking through the semantic relevance graph and retrieved online to avoid regular patterns.

Human involved experiments show that humans can solve Relacha tasks with high accuracy and efficiency, by quickly recognizing the relevance across images, where computer programs constantly fail.

Following we first summarize the up-to-date related image CAPTCHA solutions briefly in Sect. 2, with analysis on the security and robustness limitations of each. We present and explain the design details of Relacha in Sect. 3. The proposed new scheme is evaluated and validated in Sect. 4 with experiments and results, which are followed by a conclusion of the paper contributions in Sect. 5.

## 2 Related Works

As an alternative to text, latest CAPTCHA applications utilize image classification or recognition tasks as part of the challenge, as the examples shown in Fig. 1. ESP-PIX is an image-based solution in which a collection of images are displayed, and users are requested to select a correct description from a predefined list of categories. Another image-based CAPTCHA, KittenAuth, uses a fixed database to present images of cats to users [6]. These image-based CAPTCHAs are vulnerable to brute-force attack due to their static and fixed collection of images and descriptions.



(a) KittenAuth

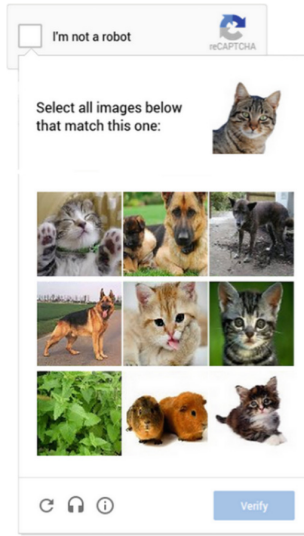


(b) ESP-PIX

**Fig. 1.** Examples of image-based CAPTCHAs.

The “no captcha reCaptcha” developed by Google analyzes various aspects of a user’s interaction with a displayed captcha and calculates a confidence score, then returns CAPTCHA challenges at different difficulty levels. For lower scores, a user may be presented with an image-based challenge as in Fig. 2, in which users are required to select proper images with similar content from the question image. ReCaptcha system

has been widely used. However, Sivakorn et al. [7] design a novel attack for image-based CAPTCHAs that extract semantic information from images. By using image annotation services and libraries, the attack approach is capable of identifying the content of images and selecting those depicting similar objects.



**Fig. 2.** A cat test of no captcha reCaptcha.

So far, all these CAPTCHA challenges are typically composed of questioning keywords and image choices. CAPTCHA answers are usually another representation of the questioning keywords or challenges, i.e. the correct answers match the question with identical semantic meaning in content. This is vulnerable to pattern recognition based image understanding, indexing, and matching. Therefore, using harder AI problems for web security is necessary.

Zhu et al. introduce a new security primitive based on hard AI problems, namely, a novel family of graphical password systems integrating Captcha technology, called CaRP (Captcha as gRaphical Passwords) [8]. CaRP is click-based graphical passwords and the sequence of clicks on an image is used to derive a password.

To avoid attackers collecting and recording the passwords, Catuogno and Galdi propose a graphical password scheme replacing static graphical challenges with on-the-fly edited videos [9]. The approach shows users a short film containing a number of pre-defined events and the users have to recognize such events, such as actions or concepts within a sequence of short videos. The graphical password utilizes human ability on recognizing the “meaning” of an object instead of its shape in a video.

Yang et al. focus on the ability to solve games, which is also one of the most advanced human cognitive process abilities. They propose GISCHA, a new way to create CAPTCHA using game-based image semantics [10]. The GISCHA challenge

can be easily operated by using only simple arrow keys, mouse movements, gestures and accelerometer. Figure 3 shows a GISCHA using the simple rolling ball game. The GISCHA challenge is to move the ball to the destination hole shaped as a circle.

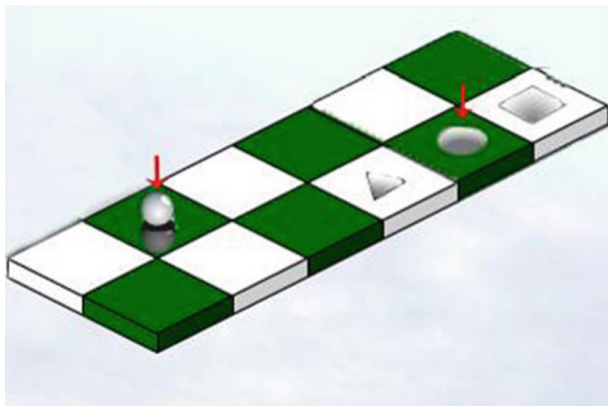


Fig. 3. A rolling ball game in GISCHA

Please select the pictures related with ~~travelling~~

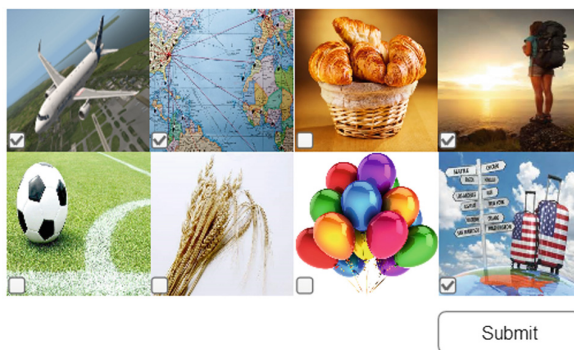


Fig. 4. An example of Relacha.

In this paper, we use the correlation other than equality, matching or similarity to improve the strength of image CAPTCHA, and design a novel CAPTCHA solution based on the relevance of associative meanings in images and the sequence of user clicks on images.



### 3 Relacha System

Now we present the overall system design of the Relacha CAPTCHA with detailed description about the Relacha construction and validation schemes.

A Relacha test consists of four procedures. The first procedure is to create a lexicon of image tags retrieved by web crawlers from hot words lists online. The second procedure generates a semantic relevance graph which depicts the correlation of all image tags provided by the lexicons retrieved. The third procedure produces and presents the test challenge by randomly walking through the semantic relevance graph. The last procedure is to receive and evaluate a user's choices according to an image associative meaning scoring algorithm, and to decide whether this user passes the Relacha test or not. Each of the four procedures of is designed with a specific goal that makes the entire framework easy to implement and use, and guarantees the generated CAPTCHA challenge robust against random attacks and exhaustive attacks.

#### 3.1 Relacha Formation

Each Relacha challenge is composed of a question text and a grid of 8 choice images, as the example shown in Fig. 4, in which a user is challenged to choose images with content meanings associated with the question keyword "traveling".

#### 3.2 Lexicon Creation

We first create a lexicon of image tag seeds, which can be automatically obtained, maintained, and refreshed. The lexicon consists of hot words online. First, we grab hot words as seeds from the Internet using web crawlers. Then we crawl and identify words correlated to those seed words according to the query suggestions from the public search engines (such as Google and Bing). Next, we filter these words to add to the lexicon. We drop those less-frequently used words based on the number of results that the search engine returns when a word is retrieved. If the number of retrieval results is below a predefined threshold (depending on the search engine used), then the word is excluded from the lexicon. We tag every remained word with a proper PoS (Part of Speech). We keep words having actual semantics, such as nouns, verbs, adjectives, and append them to the lexicon.

#### 3.3 Graph Generation

We quantize the correlation of word semantics to generate a semantic relevance graph. The correlation of any two words in the lexicon is measured by mutual information (MI) [11]. Mutual information is the correlation between variable's values and is a measure of how well a given variable can be predicted using a linear function of a set of other variables.

With two words  $W_i$  and  $W_j$ , we retrieve them individually to get online statistics denoted as  $c(W_i)$  and  $c(W_j)$  from the search engine, then search  $W_i + W_j$  in order to get retrieval results denoted as  $c(W_{i,j})$ , search  $W_j + W_i$  to get retrieval results denoted as  $c(W_{j,i})$ . Then we infer

$$c(W_i, W_j) = (c(W_{i,j}) + c(W_{j,i}))/2 \tag{1}$$

We calculate the mutual information of  $W_i$  and  $W_j$  with  $c(W_i, W_j)$ ,  $c(W_i)$  and  $c(W_j)$ , the mutual information is represented by  $MI(W_i, W_j)$  and is calculated as

$$MI(W_i, W_j) = \log_2 \frac{c(W_i, W_j) \times N}{c(W_i) \times c(W_j)} \tag{2}$$

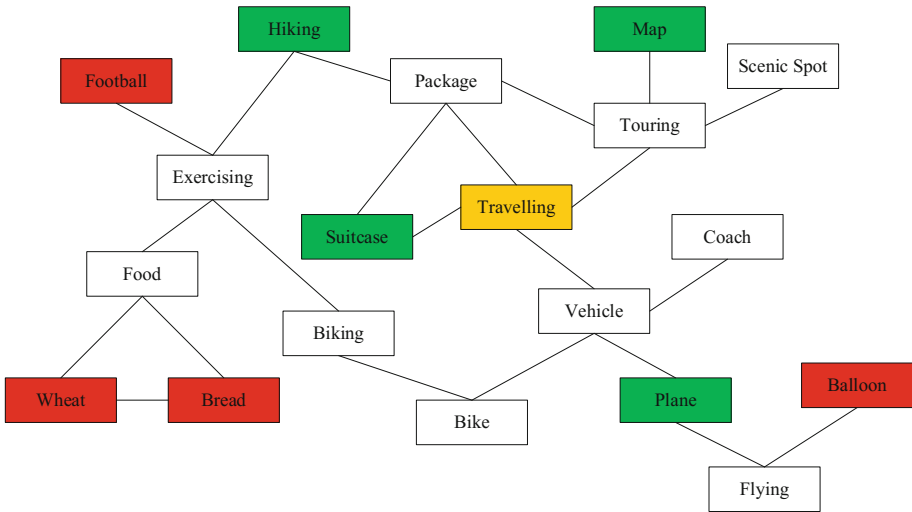
where  $N$  represents the maximum of the retrieved results.

The correlation of each word pair is represented in a tuple of three attributes ( $word_1, word_2, corr(word_1, word_2)$ ). The storage is organized as in Table 1.

**Table 1.** Correlation storage format.

$Word_1$	$Word_2$	Degree of correlation
$W_i$	$W_j$	$MI(W_i, W_j)$
.....	.....	.....

A connected undirected graph  $G = (V, E)$  is further constructed with  $n$  nodes and  $m$  edges. The nodes  $V$  are the collection of word tags of CAPTCHA images in the lexicon. Each edge  $e$  corresponds to relevance tags, with weights given by degrees of correlation denoted as  $C_{ij}$ , associating node  $v_i$  with  $v_j$ .



**Fig. 5.** An example of subgraph of the semantic relevance graph. “Traveling” is the keyword, “hiking”, “suitcase”, “map”, “plane” are the closely-related answers, “football”, “wheat”, “bread”, “balloon” are the interference answers.

If a degree of inter-word correlation is greater than a predefined threshold, the two nodes in the graph are connected with an edge. The weight value on each undirected edge is just the correlation degree of the word pair. The final graph  $G$  depicts the correlation of words in the lexicon. Figure 5 shows an example of the subgraph of the semantic relevance graph corresponding to the sample in Fig. 4.

### 3.4 Choice Generation

Given a graph and a node as the starting point, we first select a neighbor of it at random, and move to this neighbor node as current. Next random move happens subsequently from every current node. The random sequence of nodes selected in this procedure is a random walk in the graph [12]. To generate a CAPTCHA challenge randomly and dynamically, we adopt the random walk model for Relacha.

With a random walk on the connected undirected graph  $G = (V, E)$ , we define transition probabilities  $p_{t+1|t}(j|i)$  from  $v_i$  to  $v_j$  by normalizing the correlation degree out of node  $v_i$ , so  $p_{t+1|t}(j|i) = C_{i,j} / \sum_k C_{i,k}$ , where  $v_k$  ranges over all nodes connected with  $v_i$ . The notation  $p_{t+1|t}(j|i)$  denotes the transition probability from node  $v_i$  at step  $t$  to node  $v_j$  at time step  $t + 1$ .

Supposing we start at a node  $v_i$ ; if at the  $t$ -th step we are at a node  $v_j$  with probability  $p_{t|0}(j|i)$ . Clearly, the sequence of random nodes ( $v_t$ :  $t = 0, 1, 2, \dots$ ) is a Markov chain. We denote by  $M = [m_{i,j}]$ ,  $v_i, v_j \in V$ , the matrix of transition probabilities of this Markov chain. So

$$m_{i,j} = \frac{C_{i,j}}{\sum_k C_{i,k}}, v_i, v_j, v_k \in V, C_{i,k} \in E \quad (3)$$

$$P_{t+1|0} = M^T P_{t|0} \quad (4)$$

Hence

$$P_{t|0} = (M)^t \quad (5)$$

The random walk model takes the semantic relevance graph created in last step as input. The vertices are tags of images, and each edge weight is the degree of correlation between tags. The distance between two vertices is the reciprocal of the edge weight. These random walks do not have restarts (i.e. a teleport probability of returning back to their root) to avoid loop where the content of CAPTCHA answers equals to that of the question in semantics. Because by using image annotation services and libraries, an attack system is able to identify the content of images and to select those depicting similar objects [7].

We sample uniformly a random vertex as the question keyword of a Relacha task, the vertex is also the root of the random walk. For each Relacha, eight tags are needed including both answer tags and interference tags. We first generate the list of answer tags positively correlated with the root as follows:

- Get the sum of degrees of correlation of those vertices connected to the root vertex  $v_i$  based on the diagonal matrix  $D = [d_{i,i}]$ , so the sum is  $d_{i,i} = \sum_k C_{i,k}$ .
- Conduct several times of  $\alpha$ -steps walk (walking  $\alpha$  steps from the starting vertex on the graph, such as 2-steps walk) and record the vertices visited. At the same time, sum the  $d_{i,i}$  of each recorded vertex and the sum is denoted as  $S_i$ . Stop randomly walking when the following inequality is satisfied.

$$S_i \geq D_{ii}/2 \tag{6}$$

- Count the number of the generated answer tags as  $\beta$ .

Then we generate the list of confusing tags negatively or little correlated with the root as follows:

- Calculate the number of interference tags as  $\gamma$ ,  $C$  denotes the number of the choices of a Relacha test.

$$\gamma = C - \beta \tag{7}$$

- Carry out  $\gamma$  times random walk and each walk contains at least 2 steps. For each walk, sum the distance of each step as total distance  $L_i$ . Stop the walking when  $L_i$  is greater than a *min\_distance* threshold and record the end vertex.
- Repeat  $\gamma$  times such random walk and we get a list of interference tags.

Then we get the images corresponding to the list of tags from the Internet and present these images to the end-users in web to construct a Relacha challenge.

### 3.5 Correctness Evaluation

We use a result scoring mechanism to evaluate whether a user has solved the Relacha or not. We first calculate the maximal sum of the degrees of correlation of the images as an optimal value  $ov$

$$ov = \max_i \sum_i degree_i, i \in \{answers\} \tag{8}$$

Then we define a threshold based on the optimal value as passing mark  $pm$

$$pm = conf \times ov \tag{9}$$

where  $conf$  is a constant value in  $(0, 1)$  of desired confidence.

Then we calculate the sum of the degrees of correlation of the images selected by users, and the answers are given weights from high to low according to the selection sequence. Supposing a user has submitted  $n$  answers. We get user's score  $us$  as

$$us = \sum_j (\mu - (\alpha \div n) \times l_j) \times degree_j \tag{10}$$

where  $j \in \{user\_answers\}$ ,  $l_j \in \{1, 2, \dots, n\}$ ,  $\alpha$  and  $\mu$  are parameters set based on an actual lexicon.

At last, the user score is benchmarked against the passing mark. If  $us$  is beyond the passing mark, then this user passes the test, otherwise it's a failure. The scoring mechanism uses a threshold of optimal value to tolerate minor mistakes of users and increase the diversity of answers in Relacha tasks by avoiding standard answers.

## 4 Experiments and Analysis

### 4.1 Experiments

First we have created a lexicon containing about 200 words from top trending searches within the past five months. We construct the semantic relevance graph of these words. Then we built a website which would present users with the Relacha task. The participants of experiment needed to select the images that they thought correlated to the question.

We use a metric to measure CAPTCHA efficacy with respect to the number of rounds [13]. We define the number of rounds that compose a single CAPTCHA, and the minimum (threshold) number of rounds that a human user must pass to the CAPTCHA. We consider several factors in choosing optimal values for the number of rounds. First, human subjects have limits of how many rounds they are willing to tolerate. A human subject may find 5 rounds acceptable but is unlikely to agree to 500 rounds or more. Second, computers have a speed advantage over humans. A computer can guess more quickly than a human can take a test. Below, we assume that within the time it takes for a human to complete one round, a computer program is capable of completing  $n$  rounds.

The CAPTCHA efficacy metric is the probability that in the time it takes a human to take a CAPTCHA, the human will pass and a computer will not. Let  $p$  be the probability that a human user will pass a round,  $q$  be the probability that a computer will pass a round,  $n$  be the number of times faster a computer than a human,  $m$  be the number of rounds, and  $k$  be the threshold number of rounds. Then the efficacy metric  $EM$  is

$$EM = \sum_{i=k}^m \binom{m}{i} p^i (1-p)^{m-i} \times \left[ 1 - \sum_{i=k}^m \binom{m}{i} q^i (1-q)^{m-i} \right]^n \quad (11)$$

We conducted 100 rounds tests by 20 volunteer participants from our university using PCs, or smart phones, with average age of 22. Another 1,500 rounds testing were conducted by a robot program selecting answers randomly. During the experiment, we set the pass threshold  $conf$  in (9) as 75% of the optimal value due to our previous work.

We found the optimal  $m$  and  $k$  for the experimentally determined values of  $p$ . By the second testing, we considered  $q = 0.03$ . We let  $n = 100$  and searched exhaustively over of  $m$  and  $k$  until  $EM \geq 95\%$  and  $m$  was minimized. We set the number of steps denoted as  $\alpha$  of the random walk in right tags generation as 1, 2 and 3 individually and compare the experiment results in the above three settings.

Figure 6 shows the relation the percent of 100 rounds pass rate by users. The abscissa corresponds to the percent of users, and the ordinate corresponds to the percent of rounds they passed for each type of test. For example, 90% of the users passed 86% of the 2-step walk rounds. There is an obvious decrease of human pass rate when  $\alpha$  value increases, because higher  $\alpha$  values lead to weaker correlation between the questions and answers, and the questions become more difficult for users to solve.

A robot program was written to attack the Relacha that could select the images presented in web pages randomly. And the program was set to select 1, 3, 5, 8 options once individually. Figure 7 shows the attack pass rate of a robot program in 1-step walk rounds. Figure 8 shows the attack pass rate of a robot program in 2-step walk rounds. Figure 9 shows the attack pass rate of a robot program in 3-step walk rounds. We observe that the pass rate of the robot program with 1,500 tries is below 3.5%, a lot lower compared to humans. When the bot program selects 1 or 8 options once, the attack pass rate is very close to 0. When the bot program selects 3 options once, the attack pass rate is relatively higher.

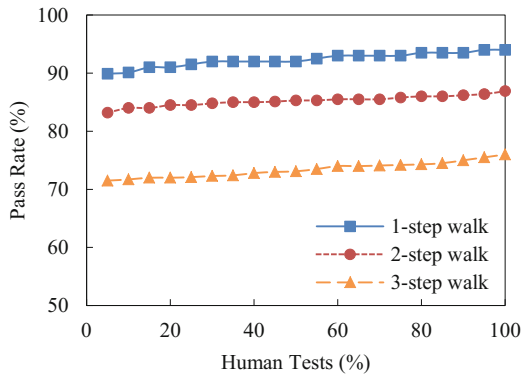


Fig. 6. Round pass rate.

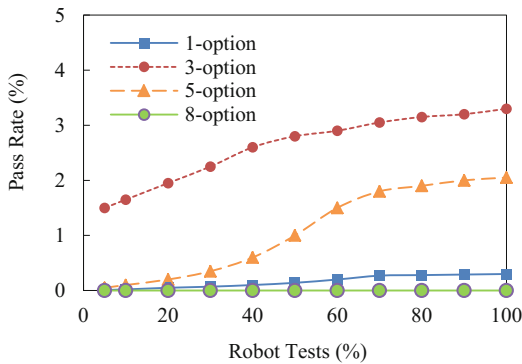


Fig. 7. Attack on 1-step walk.

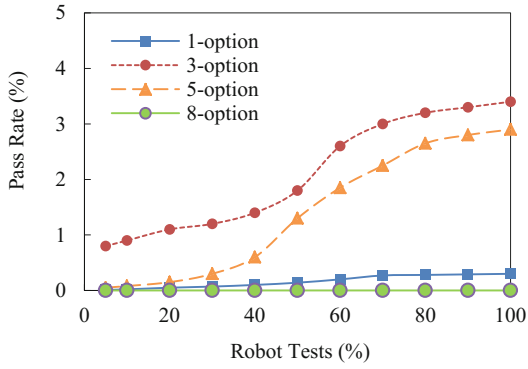


Fig. 8. Attack on 2-step walk.

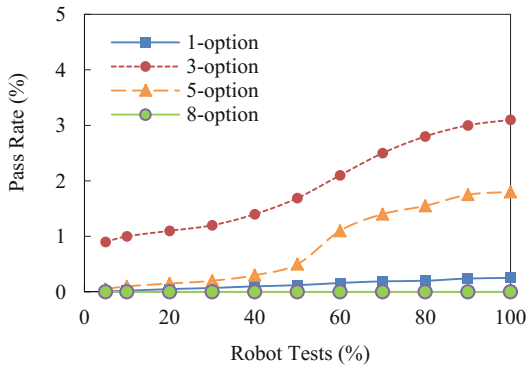


Fig. 9. Attack on 3-step walk.

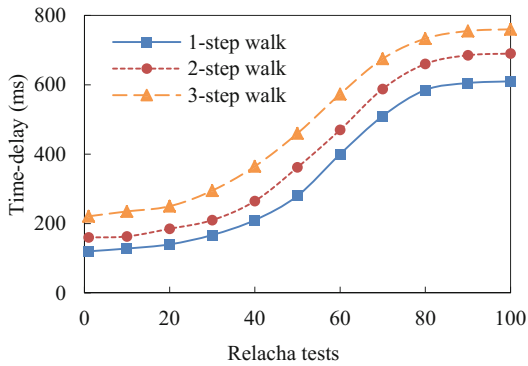


Fig. 10. Time-delays in Relacha.

Figure 10 shows the time-delays in Relacha, consisting of the execution time of Relacha generation and the load time of images. We can see that the maximum delay to present a Relacha challenge is about 760 ms and the minimum delay is about 150 ms. It is obvious that 3-step walk takes the most time to generate a Relacha test and 1-step walk takes the least time. Generally, the time-delays of Relacha stay within users' tolerance.

Table 2 gives the percentile of 100 rounds passed by each user for testing rounds. The upper and lower bounds denote the 95% confidence limits. It also shows the optimal values of  $m$  and  $k$  and the average time users spent to pass the test. To compare, we also listed the experiment results of reCAPTCHA in [10], including the pass rate of first time and the average time users took to complete a reCAPTCHA challenge. Considering both pass rate and average time, Relacha is user-friendly and easy-using.

**Table 2.** Experiment results

Test type	Percentile	Optimal $m$	Optimal $k$	Authentication time (s)
1-step walk	92.3%	4	2	4.6
2-step walk	85.2%	8	5	5.9
3-step walk	73.4%	10	7	6.1
reCAPTCHA	69%	–	–	17.5 [10]

In general, our experiments show a high performance and satisfaction from human users, and constant failure from conventional machine algorithms.

## 4.2 Discussion

As the results of experiments shown, Relacha can resist against a robot program that select the options randomly. In addition, it is promising to protect websites from machine learning-based attacks. Relacha strengthens CAPTCHA reliability against machine programs in two folds. First it avoids the dependency of measuring exact-match between CAPTCHA challenge and answers, which is vulnerable either in text or images under today's natural language processing and image pattern recognition technical level. Second, it relieves from the necessity of maintaining a local constant database for text and image material by retrieving contents and measuring their correlations dynamically from public search engine.

## 5 Conclusion

We propose a new CAPTCHA design called Relacha to distinguish humans from bots by an image correlation test, one that is promising to improve computer and information security. The image tags library used in Relacha system is made up of hot words on Internet which also makes the CAPTCHA challenge more attractive and



meaningful. A scoring mechanism is used for answer evaluation to tolerant small mistakes from different cultural background and knowledge levels of users.

Relacha uses a dynamic lexicon database based on online search engine content. Since public search engines such as Google update their indexes of images frequently [13], attackers as bot programs are unlikely to be able to exhaust all the CAPTCHA images and words. Human users with their rapid and reliable image correlation recognition and comparison (arise from years of experience with the cultural and information environment) can solve Relacha instantly. Relacha takes account of the sequence of clicks on images in correctness evaluation. Because human users tend to click the image they think most related to the question first, while bots start with the images in front.

Relacha was evaluated only against a robot that select the images presented in web pages randomly, namely the weakest possible adversary. As future work, we plan to see how Relacha behaves against more performing adversaries.

**Acknowledgments.** This material is based upon work supported by the China NSF grant No. 61472189, the CERNET Innovation Project under contract No. NGII20160601, the State Key Laboratory of Air Traffic Management System and Technology No. SKLATM201703, and the Innovation Projects of Beijing Engineering Research Center of Next Generation Internet and Applications. Opinions and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

## References

1. Yan, J., Ahmad, A.S.E.: Captcha robustness: a security engineering perspective. *Computer* **44**(2), 54–60 (2011)
2. Bursztein, E., Martin, M., Mitchell, J.: Text-based CAPTCHA strengths and weaknesses. In: *ACM Conference on Computer and Communications Security, CCS 2011, Chicago, Illinois, USA*, pp. 125–138. DBLP, October 2011
3. Smeulders, A.W.M., et al.: Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(12), 1349–1380 (2000)
4. He, K., et al.: Delving deep into rectifiers: surpassing human-level performance on ImageNet classification, pp. 1026–1034 (2015)
5. Datta, R., Li, J., Wang, J.Z.: Exploiting the human–machine gap in image recognition for designing CAPTCHAs. *IEEE Trans. Inf. Forensics Secur.* **4**(3), 504–518 (2009)
6. Goswami, G., et al.: FaceDCAPTCHA: face detection based color image CAPTCHA. *Future Gener. Comput. Syst.* **31**(1), 59–68 (2014)
7. Sivakorn, S., Polakis, I., Keromytis, A.D.: I am robot: (deep) learning to break semantic image CAPTCHAs. In: *IEEE European Symposium on Security and Privacy*, pp. 388–403. IEEE (2016)
8. Zhu, B.B., Yan, J., Bao, G., Yang, M., Xu, N.: Captcha as graphical passwords—a new security primitive based on hard AI problems. *IEEE Trans. Inf. Forensics Secur.* **9**(6), 891–904 (2014)
9. Catuogno, L., Galdi, C.: On user authentication by means of video events recognition. *J. Ambient Intell. Humaniz. Comput.* **5**(6), 909–918 (2014)
10. Yang, T.I., Koong, C.S., Tseng, C.C.: Game-based image semantic CAPTCHA on handset devices. *Multimedia Tools Appl.* **74**(14), 1–16 (2013)

11. Lopezpaz, D., Hennig, P., Schölkopf, B.: The randomized dependence coefficient. In: *Advances in Neural Information Processing Systems*, pp. 1–9 (2013)
12. Fouss, F., Pirotte, A., Renders, J.M., Saerens, M.: Random walk computation of similarities between nodes of a graph with application to collaborative recommendation. *IEEE Trans. Knowl. Data Eng.* **19**(3), 355–369 (2007)
13. Chew, M., Tygar, J.D.: Image recognition CAPTCHAs. In: Zhang, K., Zheng, Y. (eds.) *ISC 2004*. LNCS, vol. 3225, pp. 268–279. Springer, Heidelberg (2004). [https://doi.org/10.1007/978-3-540-30144-8\\_23](https://doi.org/10.1007/978-3-540-30144-8_23)

# Identification of Natural Images and Computer Generated Graphics Based on Multiple LBPs in Multicolor Spaces

Fei Peng<sup>1</sup>(✉), Xiao-hua Hu<sup>1</sup>, and Min Long<sup>2</sup>

<sup>1</sup> College of Computer Science and Electronic Engineering,  
Hunan University, Changsha 410082, China  
eepengf@gmail.com

<sup>2</sup> College of Computer and Communication Engineering,  
Changsha University of Science and Technology, Changsha 410014, China

**Abstract.** In this paper, a digital image forensics scheme based on multiple local binary patterns (LBP) in multicolor spaces is proposed for distinguishing natural images (NI) from computer generated graphics (CG). Based on the fact that the texture of natural images is more complex than that of the computer generated graphics, the diversity of their texture features are used for identification. But typical LBP only consider a single color space and a single color channel, it cannot represent the differences between NI and CG exactly. Here, we introduced two new LBP descriptors called median robust extended local binary pattern (MRELBP) and chromatic co-occurrence of adjacent local binary patterns (CCoALBP) to extract features for image identification. In our method, features of these two LBP descriptors are cascaded and they are used as the input of SVM classifier. Experimental results and analysis indicate that it can effectively identify NI and CG, and it also has a good robustness against JPEG compression, resizing, rotation and adding noise.

**Keywords:** Digital image forensics · Texture feature · Natural images  
Computer generated graphics · Local binary patterns

## 1 Introduction

Image processing techniques are becoming very popular with the quick development of information technologies. People can use computer software such as Photoshop to generate an image in an easy way, these fake images are difficult to distinguish by human eyes. Beyond that, from judicial point of view, CG may become a tool of criminals, because the criminal could make a fake image and digital images are no longer be possible to distinguish whether it is the original one or the fake one, as a result, digital image cannot be used as evidence in court of law [1], and the main purpose of image identification is to assess the authenticity of the original image [2]. Thus, the research of the identification of NI and CG received much attention in the past decade, and it has become an important research field of digital image forensics.

As we all know, NI is taken from real scenes, and NI is usually produced by cameras or cell phones. CG is made by computer software, such photoshop. So there

must have some differences between them. In recent years, some methods have been proposed to identify NI and CG. In our survey, the existing schemes can be classified into two categories, one is based on the difference of image features such as statistical features, geometric features and texture features, while the other one is based on the difference of image acquisition process.

In this paper, the diversity of image texture between NI and CG is used for identification, and two kinds of LBPs are utilized. One is MRELBP [3], which is extracted from prediction-error images (PEI), the other is called CCoALBP [4], which is extracted from original images. By combining these two kinds of LBP features, a hybrid feature vector with 616 dimensions is obtained. The main contributions of this paper include:

- Analysis about the influence of the texture features of the original image and the PEI on the identification of NI and CG is made, and as a result, we find that the texture feature of PEI is more effective for identification.
- The combination of MRELBP and CCoALBP provides good compliment for each other in the identification, and it is helpful for the improvement of the identification accuracy.
- The proposed scheme is robust against some image post-treatments, including JPEG compression, resizing, rotation and adding noise.

The rest of the paper is organized as follows. The related work is introduced in Sect. 2. The proposed scheme is described in Sect. 3. Experimental results and analyses are provided in Sect. 4. Finally, some conclusions are drawn in Sect. 5.

## 2 Related Work

### 2.1 Digital Image Forensics Based on Image Features

As mentioned above, image features include statistical features, geometric features and texture features. In 2005, Lyu and Farid et al. found that the distribution of wavelet decomposition coefficients of NI follows Laplacian distribution [5], and this characteristic is utilized for digital image forensics. 4-level wavelet decomposition is applied to images and the mean value, variance, skewness and peak value of each wavelet sub-band of the original image and its predicted version are obtained. With a total of 216 dimensions features, the identification is accomplished by using a SVM classifier. In 2014, an image forensics scheme based on multi-scale LBP was put forward by Li et al. [6]. Multiple LBP features are extracted from the original image and its corresponding PEI of Y and Cr components, where LBP uniform pattern is adopted. With 354 dimensions of features, LIBSVM is adopted for classification. An average identification accuracy of 95.10% is obtained. The experimental results of [6] proved that it is effective to use LBP for identification of NI and CG. Fractal dimension is another kind of effective descriptor for texture features, and it can be implemented for image forensics. In [7], fractal dimension is extracted from HSV color space, and the fractal dimension of PEI and the global image fractal dimension are calculated to identify NI and CG. The average identification accuracy reaches 96.00%. After that, an identification method based on multiple linear regressions is proposed in [8]. The residual

images are first extracted by using multiple linear regressions, then the fitting degree of the regression model is investigated. Image noise and fractal features are taken into consideration, and a good identification performance is achieved. Meanwhile, the feature dimension and the computational complexity are reduced.

## 2.2 Digital Image Forensics Based on Device Features

The image acquisition process of the natural images can be influenced by some non-ideal situations of the camera components. Among them, PRNU is a fixed and unique pattern noise in digital cameras, which is resulted from the imperfection of the manufacturing of CCD. So PRNU is often used in digital image forensics. In 2006, an image source identification scheme based on PRNU is proposed by Lukas et al. [9]. A PRNU model is introduced to identify the source of images. Experimental results enhance the effectiveness of using PRNU for image source identification, and it is robust against image compression. Besides, the image taken from digital cameras contains traces of resampling, which results from using CFA interpolation with demosaicing algorithms [10]. Therefore, the characteristics of CFA interpolation can be also used for image identification. As PRNU can be influenced by some factors such as the details of imaging scenes, in order to attenuate these effects, 6 mathematical models are introduced by Li [11]. These models are available for the improvement of the accuracy of source image identification.

As described above, almost all existing image source identification methods extract features from a single color channel of images. However, these descriptors cannot completely reflect the relationship between neighboring color channels. Motivated by this, a novel NI and CG identification scheme based on multiple LBPs in multicolor spaces is proposed in this paper.

## 3 Description of the Proposed Scheme

Because the texture feature of NI and CG are different in texture structure, the texture features of images are extracted for identification. Different from color features, a majority of texture features are obtained from pixels neighborhood, and it will not be affected by local deviation. Some texture features are invariant to rotation, and it is robust against image noise. Among them, LBP has been proved to be an ideal descriptor of image texture. In this paper, MRELBP is extracted from a single color channel of images, the micro and macro texture information can be revealed with MRELBP. Meanwhile, CCoALBP is calculated from two neighboring color channels, so it can reflect the relationship between neighboring color channels. Thus, MRELBP and CCoALBP are both utilized for the identification of NI and CG, The framework of the proposed scheme is illustrated in Fig. 1.

As shown in Fig. 1, a central block image  $I_c$  is first extracted from the original image  $I$ . After that, the process of color conversion is made to  $I_c$ , and the features for identification are extracted. Finally, these features are used as the input of a SVM classifier for classification. In order to avoid the effect of different sizes of images on the experiments, the size of  $I_c$  is set to  $256 \times 256$ .

### 3.1 Color Space Conversion

For color images, there are many kinds of color spaces such as RGB, HSV, nRGB and YCbCr. Different color space has different representation capability. Generally speaking, NI has a higher redundancy than CG, because image edges of CG are more obvious than that of NI. As seen in Fig. 2a(2) and b(2), the difference between NI and CG in redundancy is be more apparent in YCbCr color space. As mentioned before, MRELBP is an ideal descriptor which can reveal micro and macro texture information in an image, so it is suitable to extract MRELBP in YCbCr color space. Figure 2a(3) and b(3) illustrate the images in nRGB color space. The reason of choosing CCoALBP in nRGB color space will be presented in experiment-3 in Sect. 4.1. Due to the normalization in nRGB color space, r and g channels are scale-invariant, which make them invariant to slight intensity change. In the following, the features in nRGB color space and YCbCr clolor space are extracted for NI and CG identification.

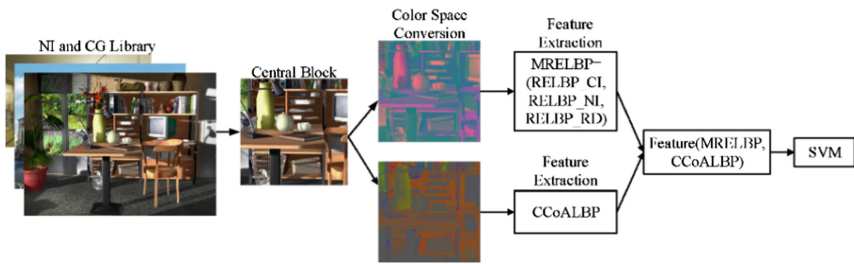


Fig. 1. Diagram of the proposed method.

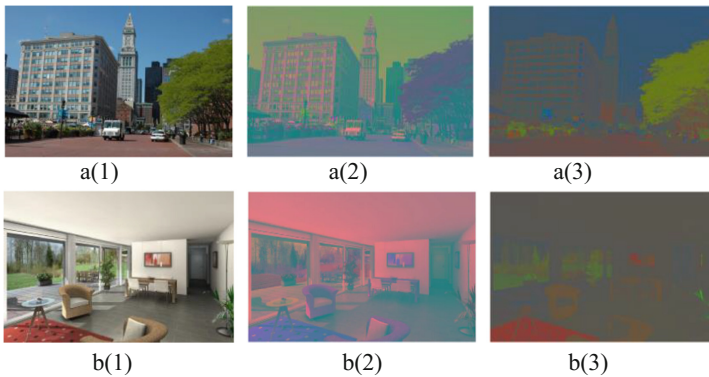


Fig. 2. Images in different color spaces: a(1) natural image; a(2) image a(1) in YCbCr color space; a(3) image a(1) in nRGB color space; b(1) computer generated graphic; b(2) image b(1) in YCbCr color space; b(3) image b(1) in nRGB color space. (Color figure online)

### 3.2 Extraction of MRELBP

Median robust extended local binary pattern is a variant of LBP [3], and it is widely used for texture classification. MRELBP has a better robustness than LBP, and it can reveal the texture macrostructure of images effectively. The procedure of MRELBP extraction is as follows.

- (1) *Pretreatment*: convert  $I$  from RGB color space to YCbCr color space, where Y is luminance component of  $I$  and Cb, Cr represent the blue-difference and red-difference chrominance components respectively. The converting procedure is defined as

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.2126 & 0.7152 & 0.0722 \\ -0.1146 & -0.3854 & 0.5 \\ 0.5 & -0.4542 & -0.0458 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

Because PEI can reduce the influence of image content on statistical characteristics of the image [12], in the first step, PEI is computed. Denoting EY, ECb and ECr as the PEI of Y, Cb and Cr channels of  $I$ . The prediction-error image can be calculated from

$$\hat{x} = \begin{cases} \max(a, b), & c \leq \min(a, b) \\ \min(a, b), & c \geq \max(a, b) \\ a + b - c, & \text{otherwise} \end{cases} \quad (2)$$

where  $x$  is a pixel value of  $I$  and  $a, b, c$  represent its neighboring right pixel, bottom pixel and bottom-right pixel, respectively.  $\hat{x}$  is the prediction-error pixel value of  $x$ . The relation of them can be seen from Fig. 3.

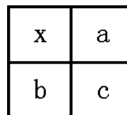
- (2) *Extraction*: as shown in Fig. 4, the procedure of MRELBP extraction is as follows.

*Step 1*: For an image  $I$  with a size of  $M \times N$ , the central block of size  $256 \times 256$  is cropped from  $I$ , denoted it by  $I_c$ . After that, convert  $I_c$  from RGB color space to YCbCr color space according to Eq. (1).

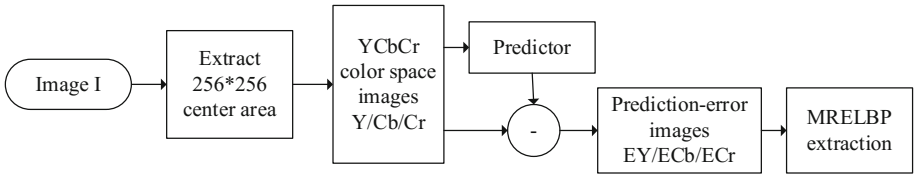
*Step 2*: Calculate the prediction-error image EY, ECb and ECr from Y, Cb and Cr according to Eq. (2).

*Step 3*: Extract MRELBP from EY, ECb and ECr, respectively.

The extraction procedure of MRELBP is shown in Fig. 4.

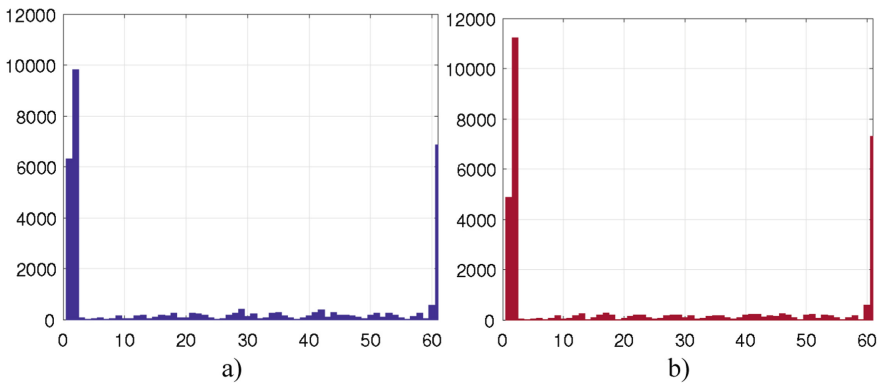


**Fig. 3.** Pixel  $x$  and its neighboring pixels.



**Fig. 4.** Diagram of the MRELBP extraction.

Median robust extended local binary pattern (MRELBP) includes three kinds of descriptors: RELBP\_CI, RELBP\_NI and RELBP\_RD. For each channel, 120 dimensions MRELBP features can be obtained, they include 2 dimensions RELBP\_CI, 59 dimensions RELBP\_NI and 59 dimensions RELBP\_RD. Therefore, there are 360 dimensions features for three color channels. Here, analysis is done to the luminance component Y to analyze the difference between NI and CG in MRELBP. Figure 5(a) represents the histogram of RELBP\_CI and RELBP\_NI of Figs. 2a(1), and 5(b) represents the histogram of RELBP\_CI and RELBP\_NI of Fig. 2b(1).



**Fig. 5.** Histogram of MRELBP: (a) histogram of NI in Fig. 2a(1); (b) histogram of CG in Fig. 2b(1).

As seen from Fig. 5, it can be found that the peak value of NI is less than that of CG. It also illustrates that the features value of NI are less than the features value of CG in low dimension. So, MRELBP can be utilized for the identification of NI and CG.

### 3.3 Extraction of CCoALBP

As MRELBP is a feature that extracted from single channel of an image, it cannot reflect the relationship of neighboring color channels. In order to obtain more correlation information between neighboring color channels of images, CCoALBP is chosen for extracting features of neighboring color channels.



Firstly, LBP is calculated from each channel [13], and it can be calculated from

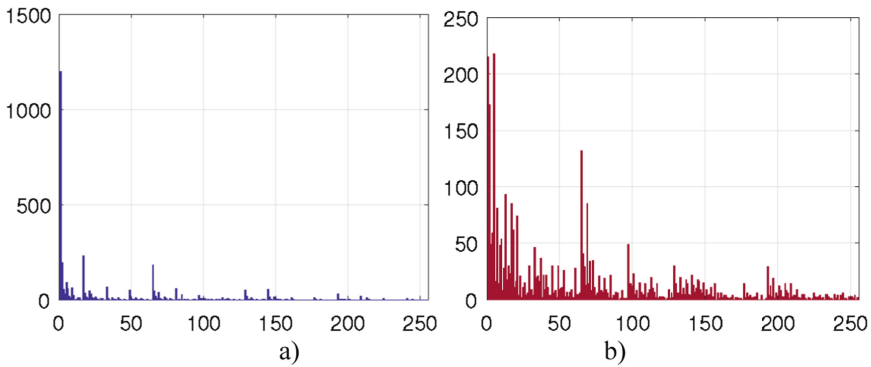
$$LBP_{P,R}^{\mu 2}(x, y) = \begin{cases} \sum_{i=0}^{P-1} s(g_i - g_c) \cdot 2^i, & \text{if } U \leq 2 \\ P \cdot (P - 1) + 3, & \text{otherwise} \end{cases} \quad (3)$$

where  $s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$ .

Then, two sparser configurations are considered, namely, LBP(+) and LBP( $\times$ ) [4]. Finally, CCoALBP feature can be calculated from color co-occurrence matrixes (CCMs) of angles  $0^\circ, 45^\circ, 90^\circ$  and  $135^\circ$ , where CCMs is defined as

$$CCM_{P,Q}^{LBP}(x, y) = \sum_{a,b \in S(I^{LBP}, B)} \begin{cases} 1, & \text{if } I_i^{LBP}(a_i, b_i) = x \\ & \text{and } I_j^{LBP}(a_j, b_j) = y, \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where  $P, Q$  and  $B$  represent the encoding strategy, directions and neighborhood of different color channels,  $S(I^{LBP})$  is the size of  $I^{LBP}$ ,  $I_i^{LBP}$  and  $I_j^{LBP}$ . Here,  $i$  and  $j$  represent the color channel, respectively. For example, for the color channel G, which is adjacent to color channel R,  $256 \times 2 \times 4 = 2048$  dimensions CCoALBP is obtained. Totally, 6144 dimensions CCoALBP can be extracted from  $I$ . Here, take color channel R for example, Fig. 6 shows the difference of CCoALBP between NI and CG.



**Fig. 6.** Histogram of CCoALBP: (a) histogram of NI in Fig. 2a(1); (b) histogram of CG in Fig. 2b(1).

It can be seen that the distribution of them are discrepant. Above analyses indicate that it is reasonable to choose CCoALBP for image identification. The procedure of the extraction of CCoALBP is as follows.

- (1) *Pretreatment*: according to the nRGB color space introduced in [14], luminance component keeps unchanged. Here, an image is first converted from RGB color space to nRGB color space according to

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \begin{pmatrix} R/(R+G+B) \\ G/(R+G+B) \\ B/(R+G+B) \end{pmatrix}. \tag{5}$$

- (2) *Extraction*: as shown in Fig. 7, the procedure of the extraction of CCoALBP is:

*Step 1*: For an image  $I$  with a size of  $M \times N$ , the central block with a size of  $256 \times 256$  is cropped from  $I$ , and it is represented as  $I_c$ . After that, convert  $I_c$  from RGB color space to nRGB color space according to Eq. (5).

*Step 2*: Calculate CCoALBP from  $I_c$  according to Eqs. (3) and (4).

The extraction of CCoALBP is shown in Fig. 7.

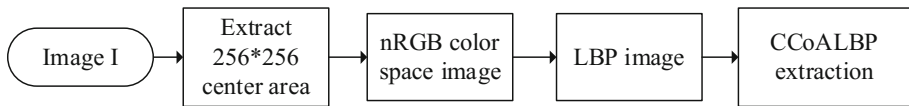


Fig. 7. Diagram of the MRELBP extraction.

### 3.4 Feature Fusion and Classification

After MRELBP and CCoALBP are calculated from the above steps, they are combined as one feature vector and used as the input of SVM classifier for classification. The image dataset is partitioned into training images and testing images. The former is used for training and the latter is used for testing. LIBSVM classifier [15] is adopted in all experiments.

## 4 Experimental Results and Analysis

The experiments are done with Matlab 2012b on a desktop computer with 2.00 GB RAM. The image dataset contains 6000 images in JPEG format. There are 3000 NIs and 3000 CGs. Among them, 800 NIs and 800 CGs are obtained from Columbia photographic images and photorealistic computer graphics dataset [16], 2200 NI from Dresden image database [17], and 2200 CG from the Internet, the URLs are <http://www.cgsociety.org>, <http://www.3dtotal.com>. The size of all central block is  $256 \times 256$ , the proportion of the training images and the testing images is 4:1 in all experiments.

### 4.1 Experimental Results

- (1) *Experiment 1*: Here, an investigation is made to test the performance of MRELBP in different color channels and different encoding methods. 800 NIs and 800 CGs

from Columbia University dataset are taken as experiment databases. Features are extracted from Y, Cb, Cr and its corresponding PEI. Here, these features and their combination are represented by  $MRELBP_i$ ,  $MRELBP_{i,j}$  and  $MRELBP_{i,j,k}$ , where  $MRELBP_i$  represents features extracted from the  $i$ -th channel,  $MRELBP_{i,j}$  and  $MRELBP_{i,j,k}$  are features that combine two and three kinds of  $MRELBP$ , respectively, where  $i, j, k \in \{Y, Cb, Cr, EY, ECb, ECr\}$ .

The experimental results are shown in Table 1. It can be found that the result of PEI received a higher classification accuracy in uniform patterns (denoted by u2), and the combination of three features can achieve a better result than only using one or two features. Hence, in the following experiments,  $MRELBP_{EY,ECb,ECr}$  in u2 encoding is chosen as classification features.

- (2) *Experiment 2*: In order to compare the performance of CCoALBP in different color spaces, this experiment evaluate the performance by using CCoALBP. The experimental results are shown in Table 2, where Mar (Marginal) represents feature vector that combines all features, Vec (Vector) represents the situation using only three channels and SVec (Simple Vector) represents the combination of three channels, two coding models and four directions. It can be found that the CCoALBP extracted from nRGB color space is more effective for image identification. In order to reduce the feature dimension and classification time,  $CCoALBP_{SVec}$  is chosen as the identification features in the following experiments.
- (3) *Experiment 3*: In above all experiments,  $MRELBP$  is applied in YCbCr color space, while CCoALBP is applied in nRGB color space. In order to test the influence of feature dimension on the two kinds of feature descriptors in the same color space, features are extracted with CCoALBP in YCbCr color space, and then we combine it with features extracted with  $MRELBP$  in YCbCr color space. From Tables 1 and 3 we can see that the accuracy achieves 98.44% with method of  $MRELBP_{EY,ECb,ECr}$ , and the feature dimension is 360. When we combine the features that extracted with CCoALBP from nRGB color space, the accuracy reaches 99.06%, and the feature dimension is 616. When we combine the features of CCoALBP that extracted from YCbCr color space, the feature dimension keeps unchanged, is 616, but the accuracy is still 98.44%. This fact means that the feature dimension has almost no influence on the accuracy of identification.

**Table 1.** Classification accuracy of different kinds of  $MRELBP$  features on single channel

Descriptor	Accuracy (%)		Method	Accuracy (%)	
	riu2	u2		riu2	u2
$MRELBP_Y$	94.50	91.56	$MRELBP_{EY}$	94.06	93.13
$MRELBP_{Cb}$	91.25	95.63	$MRELBP_{ECb}$	90.00	95.00
$MRELBP_{Cr}$	90.31	95.00	$MRELBP_{ECr}$	91.56	94.69
$MRELBP_{Y,Cb}$	96.25	96.25	$MRELBP_{EY,ECb}$	95.56	97.19
$MRELBP_{Y,Cr}$	94.06	96.25	$MRELBP_{EY,ECr}$	97.19	96.88
$MRELBP_{Cb,Cr}$	96.56	97.50	$MRELBP_{ECb,ECr}$	91.56	96.25
$MRELBP_{Y,Cb,Cr}$	96.88	97.19	$MRELBP_{EY,ECb,ECr}$	97.50	<b>98.44</b>

**Table 2.** Classification accuracy of CCoALBP features on multicolor spaces

Descriptor	Feature dimensions	Accuracy (%)		
		RGB	nRGB	YCbCr
CCoALBP <sub>Mar</sub>	6144	97.19	98.44	98.43
CCoALBP <sub>Vec</sub>	2048	96.88	98.13	97.50
CCoALBP <sub>SVec</sub>	256	95.00	97.19	96.88

**Table 3.** Comparison results of different methods with 1600 Nis and CGs.

Method	Feature dimensions	Accuracy (%)	AUC
Li <i>et al.</i> [6]	354	98.44	0.996
Lv <i>et al.</i> [7]	41	83.44	0.911
Peng <i>et al.</i> [8]	24	83.75	0.924
MRELBP <sub>EY,ECb,ECr</sub>	360	98.44	0.995
CCoALBP <sub>SVec</sub>	256	97.18	0.984
MRELBP <sub>EY,ECb,ECr</sub> + CCoALBP <sub>SVec</sub>	616	<b>99.06</b>	0.998

## 4.2 Performance Analyses

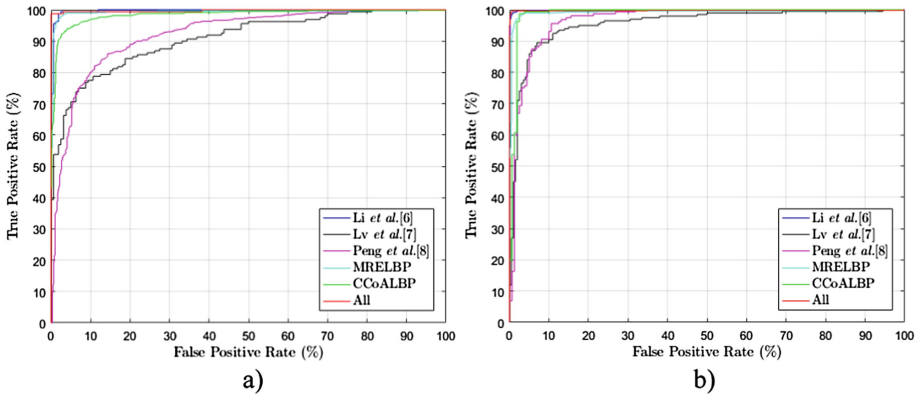
- (1) *Performance comparison:* The performance of the proposed method is compared with three methods in [6–8]. Columbia University dataset is used as experiment dataset and the experiment results are shown in Table 3. It indicates that the two texture features can identify NI and CG efficiently. With the increase of the feature dimension, the identification accuracy is also increased.

Furthermore, the comparison is also performed to another group of images of the same image dataset above, including 3000 NI and 3000 CG. The experiment results are listed in Table 4, and the results indicate that the proposed method outperforms the others.

**Table 4.** Comparison results of different methods with 6000 NIs and CGs

Method	Feature dimensions	Accuracy (%)	AUC
Li <i>et al.</i> [6]	354	99.67	1.0
Lv <i>et al.</i> [7]	41	89.25	0.945
Peng <i>et al.</i> [8]	24	90.25	0.953
MRELBP <sub>EY,ECb,ECr</sub>	360	99.00	0.995
CCoALBP <sub>SVec</sub>	256	98.92	0.992
MRELBP <sub>EY,ECb,ECr</sub> + CCoALBP <sub>SVec</sub>	616	<b>99.73</b>	1.0

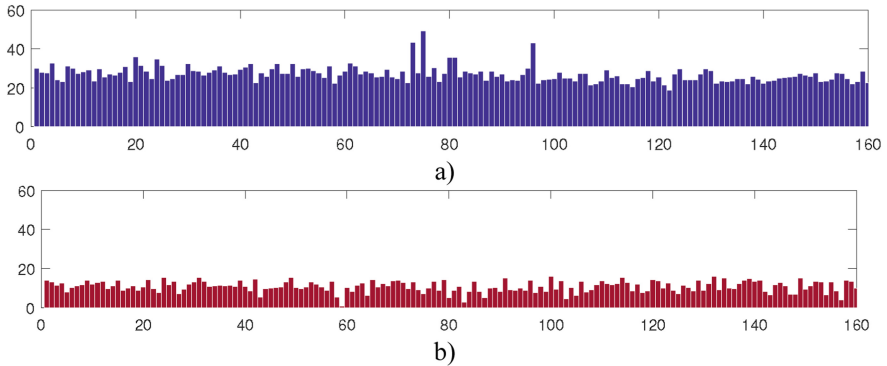
The ROCs of different methods are shown in Fig. 8. It also can be found that the proposed method has the largest area under curve (AUC). Because MRELBP is calculated from single channel of the image, it can only describe the information of textures in single channel but lacks of the relationship between different color channels, while CCoALBP is obtained from two neighboring channels. The combination of two features provides a comprehensive description of the texture features. Thus, the identification accuracy is improved.



**Fig. 8.** Comparison of the ROC curves: (a) ROC curve with 800 NI and 800 CG; (b) ROC curve with 3000 NI and 3000 CG.

It is worth noting that the method in [6] is similar to our method, both of these two schemes use variant LBP. But the method in [6] is a descriptor which only considers a single color channel, like MRELBP. Our method take not only features in a single channel, but also the features in two neighboring channels. So our method can obtain more useful information in color image than method in [6].

- (2) *Feature effectiveness*: As shown in Fig. 9, the decision value of features classification is significantly different for NI and CG. It also illustrates that it is effective to identify NI and CG with MRELBP and CCoALBP.
- (3) *Analysis of robustness*: In order to analyze the performance of robustness of the proposed method, some post-treatments like JPEG compression, resizing, rotation and adding noise are made to the images. As seen from Table 5, the JPEG quality factor has some impacts on the identification accuracy, but the accuracy is still acceptable. From the results listed in Table 5, the proposed scheme has good capability in resisting JPEG compression, resizing, rotation and adding noise.



**Fig. 9.** Comparison of decision value of feature classification: (a) decision value of NI; (b) decision value of CG.

**Table 5.** Experimental results of the robustness analysis

JPEG compression		Resizing		Rotation		Adding noise	
Quality factor	Accuracy (%)	Scaling factor	Accuracy (%)	Angle	Accuracy (%)	SNR	Accuracy (%)
20	93.50	0.5	98.75	45°	97.92	20	95.33
40	94.25	0.8	98.17	90°	97.17	40	97.00
60	95.08	1.2	97.58	135°	97.33	60	97.50
80	97.25	1.5	97.50	180°	97.83	80	97.50

## 5 Conclusion

In this paper, a novel NI and CG identification method based on multiple LBPs in multicolor spaces is proposed. As NI and CG are generated with different scenarios, there exist texture difference between NI and CG, especially in multicolor space. The combination of texture descriptor of MRELBP and CCoALBP are used to describe the texture difference between NI and CG comprehensively. By using MRELBP extracted from YCbCr color space and CCoALBP extracted from nRGB color space, 616 dimensions features are obtained for identification. Experimental results indicate that the proposed method can achieve a good identification accuracy and has a good robustness against some image post-treatments.

**Acknowledgments.** This work was supported in part by project supported by National Natural Science Foundation of China (Grant Nos. 61572182, 61370225), project supported by Hunan Provincial Natural Science Foundation of China (Grant No. 15JJ2007).

## References

1. Bayram, S., Sencar, H., Memon, N., Avcibas, I.: Source camera identification based on CFA interpolation. In: IEEE International Conference on Image Processing 2005, vol. 3, pp. III-69–72 (2005)

2. Ng, T.T., Chang, S.F., Lin, C.Y., Sun, Q.: Passive-blind image forensics. *Multimed. Secur. Technol. Digit. Rights* **15**, 383–412 (2006)
3. Liu, L., Lao, S.Y., Fieguth, P.W., Guo, Y.L., Wang, X.G., Pietikainen, M.: Median robust extended local binary pattern for texture classification. *IEEE Trans. Image Process.* **25**(3), 1368–1381 (2016)
4. Nosaka, R., Ohkawa, Y., Fukui, K.: Feature extraction based on co-occurrence of adjacent local binary patterns. In: *Proceedings of the 5th Pacific Rim Conference on Advances in Image and Video Technology*, pp. 82–91 (2012)
5. Lyu, S., Farid, H.: How realistic is photorealistic? *IEEE Trans. Sig. Process.* **53**(2), 845–850 (2005)
6. Li, Z.H., Zhang, Z.Z., Shi, Y.Q.: Distinguishing computer graphics from photographic images using a multiresolution approach based on local binary patterns. *Secur. Commun. Netw.* **7**(11), 2153–2159 (2014)
7. Lv, Y., Wan, G., Shen, X.J., Chen, H.P.: Blind identification of photorealistic computer graphics based on fractal dimensions. In: *Proceedings of the 2014 International Conference on Computer, Communications and Information Technology*, pp. 257–260 (2014)
8. Peng, F., Zhou, D.L., Long, M., Sun, X.M.: Discrimination of natural images and computer generated graphics based on multi-fractal and regression analysis. *AEU-Int. J. Electron. Commun.* **71**, 72–81 (2017)
9. Lukas, J., Fridrich, J., Goljan, M.: Digital camera identification from sensor pattern noise. *IEEE Trans. Inf. Forensics Secur.* **1**(2), 205–214 (2006)
10. Gallagher, A.C., Chen, T.: Image authentication by detecting traces of demosaicing. In: *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR workshops)*, pp. 1–8 (2008)
11. Li, C.-T.: Source camera identification using enhanced sensor pattern noise. *IEEE Trans. Inf. Forensics Secur.* **5**(2), 280–287 (2010)
12. Sutthiwan, P., Cai, X., Shi, Y., Zhang, H.: Computer graphics classification based on markov process model and boosting feature selection technique. In: *Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP 2009)*, pp. 2913–2916 (2009)
13. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
14. Zhu, C., Bichot, C.-E., Chen, L.: Multi-scale color local binary patterns for visual object classes recognition. In: *Proceedings of the 2010 20th International Conference on Pattern Recognition (ICPR 2010)*, pp. 3065–3068 (2010)
15. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**(3), 27:1–27:27 (2011)
16. Ng, T.T., Chang, S.F., Hsu, J., Pepeljugoski, M.: Columbia photographic images and photorealistic computer graphics dataset, ADVENT. Columbia University, Technical report, 205-2004-5 (2004)
17. Gloe, T., Böhme, R.: The dresden image database for benchmarking digital image forensics. In: *Proceedings of the 2010 ACM Symposium on Applied Computing*, pp. 1584–1590 (2010)

# A Formal Android Permission Model Based on the B Method

Lu Ren, Rui Chang<sup>(✉)</sup>, Qing Yin, and Yujia Man

State Key Laboratory of Mathematical Engineering and Advanced Computing,  
Zhengzhou 450001, China  
crix1021@163.com

**Abstract.** The rapid development of Android devices brings the increase of security requirements, especially for access control. Recently, many enhancements have been put forward towards the Android permission mechanism. However, few researches focus on the formalization and verification of security schemes. In this paper, we propose a formal Android permission model based on the B method, describing mechanism specifications and proving security properties. All model components are type checked by AtelierB, with 87% (154 out of 178) of generated proof obligations proved yet. The model is fully animated and checked by ProB. The results show that all specifications are well-defined without any deadlock and invariant violation. The proposed B model is for not only security analysis, but also system animation and extension. It presents a feasible approach to specify and verify the security scheme in the embedded system, which is able to translate into executable codes and implement practical module as well.

**Keywords:** Android · Permission mechanism · B method · Formal model  
Model verification

## 1 Introduction

Recently, the security of access control on mobile devices has been a research hotspot. Android is a widespread embedded operating system designed with multi-layer access control protections, among which the permission mechanism in middleware framework is most vulnerable due to the coarse-grained policy, incomplete administration and other inherent vulnerabilities [1]. With numerous mechanism enhancements applied to different Android systems, the security analysis should be explicit and complete enough to assure the correct execution of practical programs. A formal access control model can specify the proper mechanism and provide a theoretical framework for security study. However, there are only a few researches targeting at mechanism formalization and verification in Android devices, especially code generation consistent with the formal model. The B method is an application-oriented formal method for describing, designing and developing the essential module. The B model can be checked and proved, thus mathematically guarantee the high security. It has been utilized in many safety-critical fields, but rarely in Android system.



Our main contribution is the formalization of Android permission mechanism based on the B method. We construct a formal access control framework, which describes the mechanism specification and simulates inter component communication in the system. Compared with existing permission models [2–4], the proposed B model adds more implementable details, such as the permission revocation. It can be utilized in the analysis and proof of scheme security, and also as a fundamental framework to support extensions and ensure the consistency. By theorem proving, we can find the incompleteness and analyze the feasibility of mechanism enhancement, which contributes to guide the security design. By model checking, we can find the concrete deadlock and invariant violation in the model. Furthermore, this paper lays a foundation for the whole system formalization. The model specifications can be animated, tested, and then translated into executable codes after proving. The developed formal model consists of three abstract machines, which are type checked and proved by AtelierB. 87% of proof obligations are demonstrated yet. ProB is utilized to check the model and explore the state space. The checker results show that all formal specifications are well-defined. The B model is animated for the application case as well. The results show the feasibility of mechanism formalization and verification.

The remainder is organized as follows: Sect. 2 provides overview of the Android system and B method. The permission mechanism is then analyzed in Sect. 3. Section 4 constructs two elementary abstract machines, followed by the comprehensive permission model. The designed B model is evaluated in Sect. 5. Finally, we discuss the related and future work in Sect. 6, and conclude in Sect. 7.

## 2 Background

### 2.1 Android Overview

Android is an open source software stack including the Linux kernel, middleware and application layer, with diverse access control mechanisms enforced on each tier [5].

**Application.** An Android application consists of various components, which are typically divided into four types: Activity, Service, Broadcast Receiver, and Content Provider. Every application has an `AndroidManifest.xml` file, providing the system with basic information, such as package attributes, contained components and relevant permissions.

**Component and ICC.** As the basic unit of application, a component should be instantiated by the system at runtime. Android kernel sandbox isolates one application execution from others. Application framework accomplishes permission based access control on the component. By default, the components of same application run in the same process, and components running in different processes communicate under control of ICC (Inter-Component Communication) mechanism. Successful ICC is the key to trouble-free running of apps.

**Permission.** Each Android permission is defined as a unique string, used for limiting access to certain codes or data. Each permission has its identifier, permission group and protection level (including Normal, Dangerous, Signature and SignatureOrSystem).

Apart from system permissions, the custom permission of an application can be defined in the app package. The following categories of essential permissions are clarified in `AndroidManifest.xml` file:

- A new permission declared by application to limit access to specific component, defined in `<permission>` element.
- A permission requested by application for the normal operation, defined in `<use-permission>` element.
- A permission enforced to restrict interaction with all sub-components of the application, defined in `<application>/android:permission` attribute.
- A permission for protecting specific component, which overwrites the third type of permission, defined in `android:permission` attribute of an individual component.

## 2.2 Android Permission Mechanism

The permission mechanism is reinforced at both the application installation time and running time. When installed, an application requests a set of dangerous level permissions to be granted by user. Typically, users are only informed of the permission group. At runtime, system performs permission check on instantiated components and controls the ICC, making sure that a component cannot be accessed unless its permission requirements are satisfied. In Android 5.0 (or lower), all permissions must be granted at installation time, and cannot be revoked until the application is uninstalled. Since 6.0, Android supports dynamic permission revocation, i.e., the user can revoke certain permissions at run time [6].

The Android application framework layer is the most vulnerable part for access control. Many security threats occur in the permission mechanism due to coarse-grained policy, incomplete administration and other defects. In [1], Davi et al. illustrated the security issues in Android permission mechanism and proposed the privilege escalation attack. A malicious application could exploit the vulnerability in application interfaces and ICC mechanism, causing unauthorized operations and permission leakages.

## 2.3 B Method

B method is a state-based formal method depending on the Zermelo-Fraenkel set theory and first order logic [7]. The B development comprises two closely linked activities: writing formal specifications and proving them. The first stage is to build abstract machines which contain all defined requirements. A formal model specification is composed of data, relative invariable properties and corresponding operations. Then, the abstract machines are refined with more details and finally transformed into fault-free concrete models. During the modeling process, the second activity performs numerous type checks and demonstrations of theorems in order to prove the correctness of specifications and conservation of invariants. After proved, the model can be coded into C or Ada language [8]. Many structures are defined for development, like MACHINE, REFINEMENT, IMPLEMENTATION, etc. There are lots of mature tools supporting the B method, such as AtelierB and ProB.

### 3 Preliminary Analysis

The features of access control mechanism are analyzed informally at first, such as basic system elements, authorization rules, and security properties.

#### 3.1 Model Element and Property

Based on realistic Android application framework, Fig. 1 presents the formal model structure, where the permission and component abstract machine are in base layer, with the system abstract machine above. In every entity, we extract necessary elements and take them as variables in corresponding models.

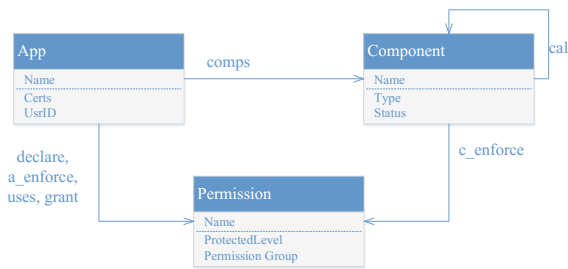


Fig. 1. Relations among the model entities.

Figure 1 also indicates specific relationships among the application, component and permission entities, which are consistent with the security properties summarized in Table 1. Moreover, ten basic authorization rules in Android middleware are obtained according to the official document [6], compatible with the permission mechanism. They should hold the security properties as well.

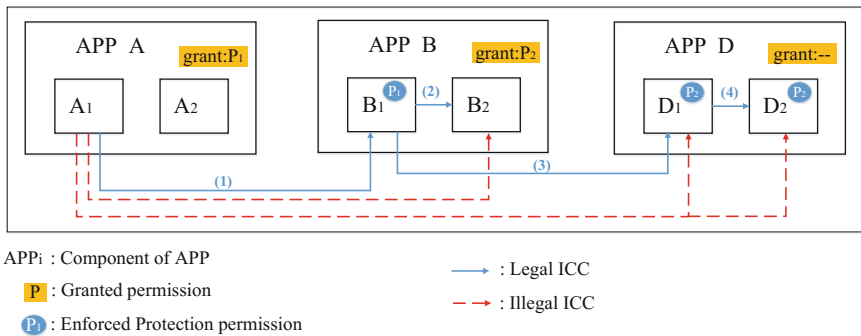
The properties in Table 1 should be translated into either invariant statements or operation conditions in the formal model. While invariants guarantee that the system always satisfies static properties, operation specifications are used to maintain dynamic attributes of the system.

In terms of dynamic operation, possible inter-component communication scenes are depicted in Fig. 2. Normally, component  $A_1$  and  $A_2$  are granted to permission  $P_1$ , thus  $A_1$  has legal privileges to invoke  $B_1$  which is enforced by  $P_1$ . Similarly,  $B_1$  can communicate with  $D_1$ . The communication between  $B_1$  and  $B_2$  is legal because they belong to the same application. However, Fig. 2 also highlights possible privilege escalation attacks. Component  $A_1$  should not have legal privileges to invoke  $D_1$ , but it can directly send messages to  $B_1$ . When  $B_1$  has the permissions to call  $D_1$ ,  $A_1$  can be re-delegated to invoke  $D_1$  through (1)(3). Similarly,  $A_1$  can access  $B_2$  illegally through (1)(2), causing the permission leakage. Consequently, the general situation of this kind of attack could be demonstrated as follows:

If the component  $C$  cannot be authorized to access  $C'$ , and it is able to start  $C''$ , which can perform the certain operation on  $C$ , then  $C$  can access  $C'$  via  $C''$ .

**Table 1.** Security properties.

Property	Description
Uniqueness	Each installed application has a unique identifier
	Each component contained in an application has a unique identifier
	Each permission in the system has a unique identifier
	Each component belongs to only one installed application
	Each permission belongs to only one permission group
	Resource in the system has unique value, held by installed applications
Existence	Each component is defined in an installed application
	All permissions are declared in the system or applications
	All the parts (e.g. components) involved in authorization activities are installed in the system
	All the components involved in communications are instantiated at run time
Least-privilege principle	Each application, by default, can only access the components that it requires to do its work and no more



**Fig. 2.** ICC call in Android middleware.

### 3.2 Component Status

An application invocation in Android middleware can be regarded as the communication process with state migration under a control network. To describe communicating states of all components, the following six state constants are defined: free (*fr*), requiring (*rq*), waiting (*wt*), verifying (*vf*), calling (*call*), called (*called*) and unavailable (*un*). Accordingly, a function for status expression is set up, with all components initialized to free (*fr*)<sup>1</sup>.

Components should be instantiated at runtime. In general, an activity can be instantiated for multiple times while other types of components for once. Suppose that

<sup>1</sup> The value of a state variable is displayed in italic front in this paper.

each component in the system only has one running instance. Without considering invocation on a running component, we use a 2-tuple (caller, callee) to indicate the status of communicating pairs. The status transition process is as follows.

- $(fr, fr) \rightarrow (rq, vf) \rightarrow (call, called)$   
 When both components are not instantiated and in state *fr*, as component  $A_1$  requests to communicate with  $B_1$  in Fig. 2. The state of  $(A_1, B_1)$  is set to  $(rq, vf)$  at the time of condition check, and then converted to  $(call, called)$  after a successful communication.
- $(called, fr) \rightarrow (wt, vf) \rightarrow (call, called)$   
 Already called by component  $B_1$ ,  $A_1$  needs to invoke another component  $B_2$ . Then, set  $B_1$ 's state as *wt*,  $B_2$ 's as *vf*. The state of  $(A_1, B_1)$  turns to  $(call, called)$  after communicating.
- $(call, fr) \rightarrow (wt, vf) \rightarrow (call, called)$   
 It suggests that a component requires invoking several components. As displayed in Fig. 2,  $B_1$  needs to call  $B_2$  and  $D_1$ . A communicating link has been made between  $B_1$  and  $B_2$ , and then  $B_1$  and  $D_1$ 's states become *wt* and *vf*. The state of  $(B_1, D_1)$  comes into  $(call, called)$  when the call succeeds.

## 4 Android Permission Model with B Method

Based on preliminary analyses, the Android permission mechanism is formalized with B method. The system model *Perm\_Sys* is established on two base abstract machines: *Permission* and *Component*, specifying permission and component items respectively.

### 4.1 Two Base Abstract Machine

The basic permission abstract machine is shown in Table 2. An object consists of the identifier, corresponding group and protection level. The unique identifier is represented by a distinct element of the set. Model operations include adding an item, changing permission group and so on.

**Table 2.** Permission abstract machine.

MACHINE	<i>Permission</i>
SETS	<i>Permission</i> ; $PL = \{Normal, Dangerous, Signature, SigOrSys\}$
VARIABLES	<i>perms, perm_PL, perm_group</i>
INVARIANT	$perms \subseteq PERMISSION \wedge$ $perm\_PL \in perms \rightarrow PL \wedge$ $perm\_group \in perms \rightarrow INT$
OPERATIONS	add_perms, modify_group, delete_one_perm, delete_set_perms

The abstract machine, constant and variable names are displayed in italics in this paper.

**Table 3.** Component abstract machine.

MACHINE	<i>Component</i>
USES	<i>Permission</i>
SETS	<i>Component</i> ; <i>Type</i> = { <i>activity</i> , <i>service</i> , <i>c_provider</i> , <i>b_receiver</i> }; <i>STATUS</i> = { <i>fr</i> , <i>rq</i> , <i>call</i> , <i>vf</i> , <i>called</i> , <i>wt</i> , <i>un</i> }
VARIABLES	<i>cmps</i> , <i>c_enforce</i> , <i>type</i> , <i>exported</i> , <i>c_status</i>
INVARIANT	$cmps \subseteq COMPONENT \wedge type \in cmps \rightarrow TYPE \wedge$ $c\_enforce \in perms \rightarrow cmps \wedge$ $exported \in cmps \rightarrow BOOL \wedge$ $c\_status \in cmps \rightarrow STATUS$
OPERATIONS	add_cmps, add_enforce, change_one/pair_status, delete_cmps, clear_all_status

In model *Permission*, the set *PERMISSION* contains all of the permissions. For simplicity, an integer set denotes the permission group, and the permissions in different groups are mapped to distinct integers. The set *PL* defines four protection levels, among which we mainly consider the *Dangerous* category, while other types of permissions are granted normally by default.

Table 3 shows the basic component abstract machine, which uses *Permission* in order to read relevant variables. Static information of each item contains an identifier, the type, exported attribute and protecting permissions. Introducing a status function into abstract machine, the model operations include adding a component, adding enforced permissions, changing its state, etc.

As shown in Table 3, all components are contained in the set *COMPONENT*, and the set *TYPE* defines four types of components. The variable *exported* and *c\_status* represent the exported attribute and component status respectively. Explicit specifications on status transition are interpreted through model operations. The invariants define every variable type and preserve the security attributes.

## 4.2 Permission System Model

**Abstract Entity and Relation.** The permission system model *Perm\_Sys* includes abstract machine *Component* and *Permission*, utilizing variables and operations in them.

The entities variables in *Perm\_Sys* include the current permission and component items defined in two base abstract machines (i.e., *perms* and *cmps*), and a new variable *app* is set to represent the existing application set in the system. It is related to both the *perms* and *cmps*. Table 4 lists their relations in detail.

The relation variables in Table 4 can be divided into two classes: static permission attribute and dynamic information at runtime. The permission attribute of a component can be fetched from the *AndroidManifest.xml* file. For instance, the variable *uses* records the permissions need to be used, which are clarified in `<uses-permission>` tag. Dynamic information includes authorization and invocation relations. The variable *grant* denotes the permissions currently delegated, and *call\_pair* contains several pairs of communicating components dynamically.

**Table 4.** Relations between the model entities.

Code	Description
<i>a_enforce</i>	A relation between a permission and an application; where each permission might be used to protect more than one application; and that an application can be enforced by many permissions
<i>c_enforce</i>	A relation between a permission and a component; where each permission might be used to protect more than one component; and that a component can be enforced by many permissions
<i>uses</i>	A relation between a permission and an application. An application usually needs to use lots of permissions; In addition, each permission might be utilized by more than one application, such as the network access ability
<i>grant</i>	A relation between a permission and an application. An app can be granted many permissions either by system or user, depending on the need-to-use permission set. Similarly, each permission also might be assigned to more than one application
<i>declare</i>	A functional relation between a permission and an application. An application can define its own permissions. Thus, each custom permission is declared only by one application
<i>comps</i>	A functional relation between a component and an application. An app should define its own components, and each unique component is contained only in one application
<i>call_pair</i>	A binary relation on the component set, recording the communication requests of component pairs. Each component can be either a caller or callee for only once
<i>gain</i>	A relation between a permission and a component. Each component may gain extra permissions through successive invocations, which are inherited from callee components. One permission can be gained by many components as well

**Formal Model Operation.** Overall operations in *Perm\_Sys*, consist not only of manipulations on the application, such as installation, startup and permission assignment, but also of ICC simulations. There are ten crucial access-related operations defined in the model: four kinds of requests (*Out\_Req\_1*, *In\_Req\_1*, *Out\_Req\_2*, *In\_Req\_2*), a successful access (*Succ\_Access*), a finish action (*EndAccess*), a dynamic permission revocation (*Revoke\_ac*), status modification and restoration (*Modify\_ac*, *Restore\_st*, *Clr\_voidcall*). These operations also concern the state transitions of a component.

In *Perm\_Sys*, a component's state is migrated only if the invocation context satisfies verification conditions. And once permission check fails, the component remains in its original state. Figure 3 presents the complete state transition process. States of both caller and callee component start from *fr*, and they're changed under different operations. All these access-related operations have synchronization relationships according to the state transition of components.

As expressed, the revocation action is take into account in the model. Users may revoke an application's permission in any possible locations in the whole communicating process, resulting in diverse responses. Note that our model aims at specifying access control mechanism regardless of engineering project details, while the

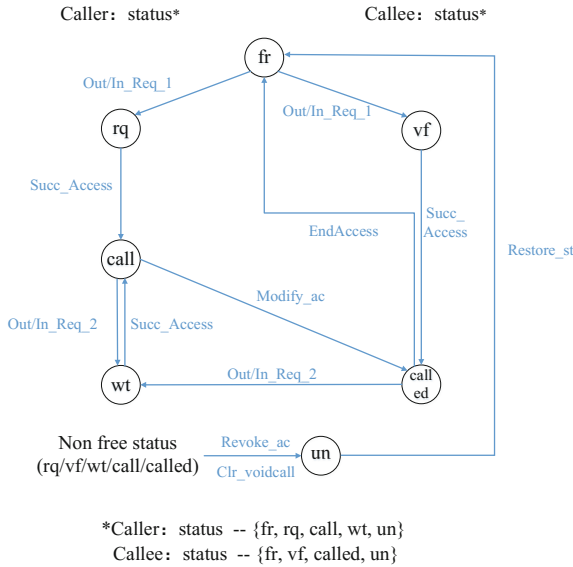


Fig. 3. State transition of a component in *Perm\_Sys*

synchronizations describe all possible calling situations theoretically. However, the practical Android permission framework only realizes some of cases due to their complexities, which may compromise the completeness as well.

## 5 Model Evaluation

### 5.1 Property Specification

Invariants and operation conditions of the model express and guarantee critical security properties. Take the following statements as examples:

**Uniqueness.** The mathematical set keeps every entity and relation unique. For example, the invariant  $app \subseteq APP \wedge comps \in cmps \rightarrow app$  in model *Perm\_Sys*, is to assure that “each installed application has unique identifier” and “each component belongs to only one installed application”.

**Existence.** Different types of variables record various existing entities. All model operations are performed on recording variables. e.g.

In *Perm\_Sys*, the variable *perms* is the set consisting of all permissions declared in the system or applications, and in every permission-related operation, there is a precondition  $pp \in perms$  to verify the existence of input parameter *pp*.

**The principle of least privilege.** This principle is observed by invariants and operations at the same time, such as the invariant  $grant \subseteq uses$  and operation *Out\_Req\_1* ( $cc_1, cc_2$ ) in *Perm\_Sys*. If a component  $cc_1$  requires to invoke  $cc_2$ , the communication



succeeds only when the  $cc_2$ 's enforced permission set is a subset of  $cc_1$ 's granted permission. i.e.,  $c\_enforce[\{cc_2\}] \subseteq grant[\{comps(cc_1)\}]$ .

**Privilege escalation threat.** Besides security attributes above, the component re-delegation behavior can be described as well. The major reason for privilege escalation is that the component callee's permissions will be assigned automatically to the caller, causing that caller obtains illegal permissions during communication. Therefore, a variable *gain* is designated to sum up all inherited permissions from both directly and indirectly called component. Accordingly, the threat can be examined by traversing all components. If there is one component  $cc$  with  $gain(cc) - grant(cc) \neq \emptyset$ , the system is subject to the privilege escalation attack.

## 5.2 Model Proof

Our model development is based on AtelierB 4.3, which supports the syntax and type check, POs (proof obligations) generation, automatic and interactive demonstration. Different levels of force are provided by the automatic prover to balance the proof efficiency and computing time [8].

In terms of typing, all model components are checked and shown to be well formed and semantically correct. In terms of proofs, Table 5 gives the number of POs and the rate of automatic proofs in the model (under the force 2 of prover). PO is a logical expression generated through predicate calculus. It utilized to find errors in specifications and guide the correction. Typically, the high automatic proof rate contributes to reduce the complexity of model, indicating the high quality.

**Table 5.** The number of POs in the formal permission model.

Machine	Code size	POs	Proved	Unproved	Proof rate
<i>Permission</i>	72	12	10	2	83%
<i>Component</i>	84	19	16	3	84%
<i>Perm_Sys</i>	245	147	91	56	62%

The prover results show that 117 out of 178 POs are proved automatically. We are now working on demonstrating remaining lemmas manually. Actually, some comprehensible POs can be easily proved by interactive prover, which are conducive to simplify the proof task. For example, in abstract machine *Permission*, when the preserved invariant ( $perm\_PL \in perms \rightarrow PL$ ) is checked by interactive prover, the following initial PO is given to be demonstrated:  $Sysperms \times \{Dangerous\} \in Sysperms \rightarrow PL$ .

By launching the mini proof command, prove command and predicate prover on reduced hypotheses in order, the original goal is reduced to be TRUE. Similarly, we have already proved 37 POs by utilizing the interactive prover, and the current proof rate is 87% (154 out 178), above average in the B development. As the first B method based access control model in the embedded system, the prover results are promising. The abstract model can be transformed into executable codes after proved.

### 5.3 Model Check

Apart from theorem proving, model checking is applied to the permission model. ProB is an animator and model checker for B method, which supports fully automatic animation of B specification, test case generation and so on [9]. Compared with the AtelierB type check, ProB checker can discover a wide range of errors in specifications by exploring all possible states, each of which is determined by all variable values in the model. We can also test system model by manually animating abstract machines. The checked state space of our model is infinite because sets are defined without certain elements. ProB assigns random values to these sets and ensures to cover all operations.

For abstract machine *Permission*, the number of variables and operations is small, thus we choose the mixed breadth first and depth first search strategy to explore the state space. The model checker result shows that it has traversed 16028 distinct states. The graphical view cannot be displayed due to the large scale. Typically, we obtain the current state space graph when all the operations are covered once, as shown in Fig. 4.

There are totally 567 distinct states and 1132 transitions in Fig. 4, and there is not any deadlock and invariant violation. Every independent branch corresponds to certain assigned values for constants and sets. All outer clusters start from the center one, which animated many initial scenes. They are different in shape because of random animation.

Since the abstract machine *Component* uses *Permission*, the animator should involve all permission-related operations as well, greatly increasing the complexity. After we apply the breadth first strategy until all operations are covered once, the size of current state space (containing 5231 distinct states and 13607 total transitions) is too large to display in the paper. The model check result shows that there is no deadlock and invariant violation in the specification as well.



**Fig. 4.** The checked state space graph of *Permission*

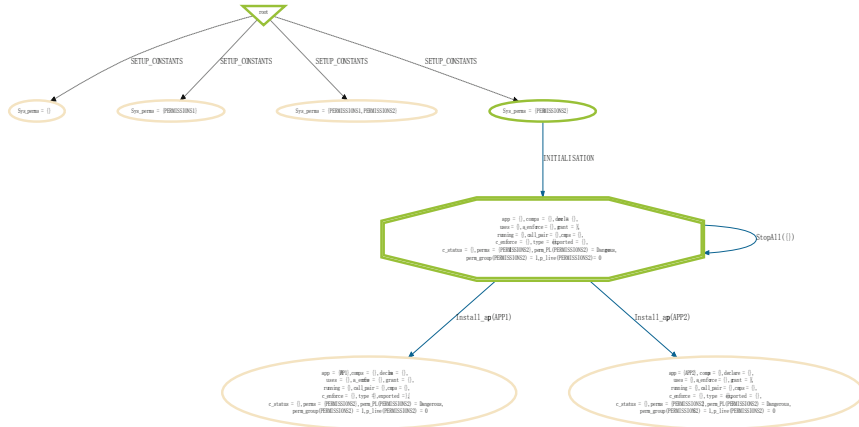


Fig. 5. A random animation view of *Perm\_Sys*

*Perm\_Sys* is the largest abstract machine and its animation is most complicated. Similarly, we gain the current explored state space view containing 77681 distinct states and 219325 total transitions. Both deadlock and invariant violation are not found. Furthermore, we take advantage of ProB to random animate the system model and test a variety of running cases. Due to the scale restriction, here gives a 2-step random animation sample.

Figure 5 draws an animated scene in the permission system model. As shown in the flowchart, starting from the root node, the animator first sets up one system permission constant, and then determines the initialization. Then it can choose to install one of two different applications, leading to different state. For example, if APP1 is installed in the system, APP1 will be added in the app set by enabling installation operation in the model. Further, we increase more applications and test various access cases until cover all system operations, with the animation view becoming larger in size. The animation of abstract machine closes to implementation of actual system, which is the advantage of the B method based model. According to the model check results so far, our model specification is well-defined and suitable for Android permission mechanism.

## 6 Related Work and Discussion

Formalization of Android permission mechanism was pioneered by Shin et al. [2]. They built a Coq model and assessed its safety and completeness. Similarly, Betarte et al. [3] enriched the model and formally proved the principle of least privilege and presence of privilege escalation in the specifications. Inspired by those model, our model can be used for mechanism analyses as well. However, different from those modeling thoughts, our model takes advantages of B method and targets at building a specification which guarantees the security property and provides extensible framework for attack defense scheme. We have extended our fundamental model in three ways to enhance permission mechanism.

In the study of Fragkaki et al. [4], an abstract permission model was proposed with desired attributes, and a system enhancement is constructed based on it. The work process is similar to our work, but the model is too abstract to keep practical solution in accordance with model specifications completely. Since the B method is advantaged in developing the safety-critical software, our model is promising to realize an executable module with consistency. Besides, we can apply the model proving and checking to find the incompleteness and analyze the feasibility of enhancement scheme, conducive to security design.

The formalization can also be used for the system modeling. Hoffmann et al. [10] used Event-B method to establish a formal model of the API in micro-kernel L4 and evaluate the technical feasibility. fmC/OS [11] is a formal operating system model with the B method, including memory management, task management and intertask communication, etc. According to research techniques, we believe that our framework is suitable for embedded system modeling. The specifications can be implemented as a verifiable secure module after proved.

However, our framework is still in the design phase and it is too simple to describe access control policy completely. It only considers the ICC call among components without other approaches. We are already trying to enrich and refine the model by adding more actual structures and actions. We also work on the model checking to test and verify the consistency of the model specifications.

## 7 Conclusion

This paper proposes a comprehensive Android permission model using the B language, which contains three abstract machines and supports verification and extension. We take advantage of the B method to specify the Android permission, component and application entities, ensuring that access control operations always satisfy invariant security properties. All model components are type checked by AtelierB, and 154 out of 178 proof obligations are proved yet. The model is checked and tested by ProB. The results show that there is no deadlock and invariant violation in all model specifications. Compared with other formal models, the B method based permission model is closer to actual implementation, which can be translated into executable code after being fully proved. It is now used not only for security analysis, but also as a fundamental framework to support mechanism extensions. Furthermore, our formal model requires proofs and refinements.

**Acknowledgments.** Thanks to project supported by the National Natural Science Foundation of China (No. 61572516).

## References

1. Davi, L., Dmitrienko, A., Sadeghi, A.-R., Winandy, M.: Privilege escalation attacks on Android. In: Burmester, M., Tsudik, G., Magliveras, S., Ilić, I. (eds.) ISC 2010. LNCS, vol. 6531, pp. 346–360. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-18178-8\\_30](https://doi.org/10.1007/978-3-642-18178-8_30)

2. Shin, W., Kiyomoto, S., Fukushima, K., Tanaka, T.: A formal model to analyze the permission authorization and enforcement in the Android framework. In: IEEE Second International Conference on Social Computing, pp. 944–951. IEEE Computer Society, Washington, DC, USA (2010). <https://doi.org/10.1109/SocialCom.2010.140>
3. Betarte, G., Campo, J.D., Luna, C., Romano, A.: Verifying Android’s permission model. In: Leucker, M., Rueda, C., Valencia, F.D. (eds.) ICTAC 2015. LNCS, vol. 9399, pp. 485–504. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-25150-9\\_28](https://doi.org/10.1007/978-3-319-25150-9_28)
4. Fragkaki, E., Bauer, L., Jia, L., Swasey, D.: Modeling and enhancing Android’s permission system. In: Foresti, S., Yung, M., Martinelli, F. (eds.) ESORICS 2012. LNCS, vol. 7459, pp. 1–18. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-33167-1\\_1](https://doi.org/10.1007/978-3-642-33167-1_1)
5. Android Developers: <https://developer.android.com/guide/platform/index.html>. Accessed 27 June 2017
6. Android Permission: <https://developer.android.com/guide/topics/security/permissions.html>. Accessed 27 June 2017
7. Abrial, J.R.: The B-book: Assigning Programs to Meanings. Cambridge University Press, Cambridge (1996)
8. Presentation of the B method | Méthode B: <http://www.methode-b.com/en/b-method/>. Accessed 01 Jul 2017
9. The ProB Animator and Modelchecker: <https://www3.hhu.de/stups/prob/>. Accessed 01 Jul 2017
10. Hoffmann, S., Haugou, G., Gabriele, S., Burdy, L.: The B-Method for the construction of microkernel-based systems. In: Julliand, J., Kouchnarenko, O. (eds.) B 2007. LNCS, vol. 4355, pp. 257–259. Springer, Heidelberg (2006). [https://doi.org/10.1007/11955757\\_23](https://doi.org/10.1007/11955757_23)
11. Chen, D., Sun, Y., Chen, Z.: A Formal Specification in B of an Operating System. Open Cybern. Syst. J. 9(1), 1125–1129 (2015). <https://doi.org/10.2174/1874110X01509011125>

# S-SurF: An Enhanced Secure Bulk Data Dissemination in Wireless Sensor Networks

Jian Shen<sup>1,2</sup>, Tiantian Miao<sup>1</sup>, Qi Liu<sup>1</sup>, Sai Ji<sup>1(✉)</sup>, Chen Wang<sup>1</sup>,  
and Dengzhi Liu<sup>1</sup>

<sup>1</sup> Jiangsu Engineering Center of Network Monitoring,  
Nanjing University of Information Science and Technology,  
Nanjing 210044, China

s.shenjian@126.com, 18362086690@163.com, qrankl@163.com,  
jisai@nuist.edu.cn, wangchennuist@126.com, liudzdzh@126.com

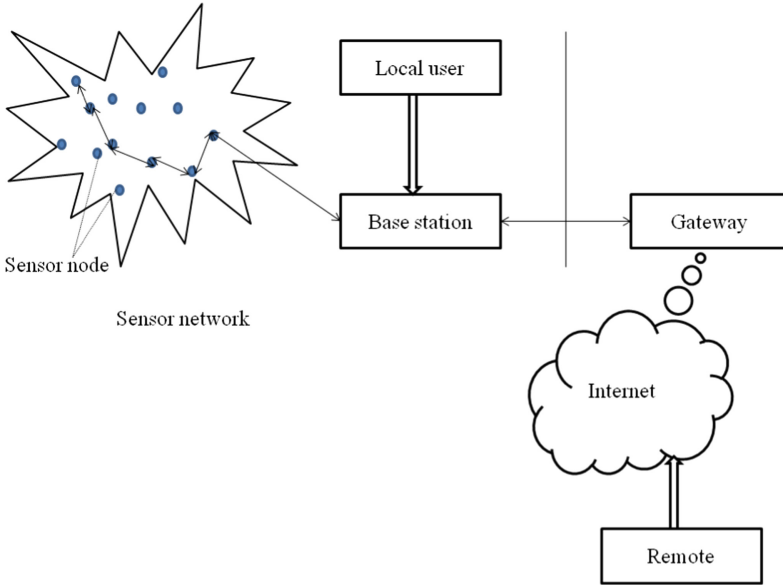
<sup>2</sup> State Key Laboratory of Information Security,  
Institute of Information Engineering, Chinese Academy of Sciences,  
Beijing 100093, China

**Abstract.** Wireless sensor networks (WSNs) have recently gained a lot of attentions as a hot topic of research, with extensive applications being explored. In WSNs, bulk data dissemination protocols are responsible for reprogramming, which have been proposed for efficiency and security. However, few of them can simultaneously achieve both reliability and security. To address this problem, we propose an enhanced protocol entitled Secure Survival of the Fittest (S-SurF) based on SurF [6] in this paper. The proposed protocol is composed of four main phases: packet preprocessing, flooding, negotiation and data verification. Moreover, S-SurF incorporates a time-reliability model to predict the minimum completion time and hence seizes the most opportune moment to transit between flooding and negotiation schemes. In addition, extensive analysis proves the efficiency and security of S-SurF.

**Keywords:** Bulk data dissemination protocol · Data authentication  
Time-reliability model · S-SurF · WSNs

## 1 Introduction

As one of the major milestones in the field of communication, wireless sensor networks (WSNs) [1] have been taken general attention. With the advantages of dynamic performance, reliability, integration and self-organization, WSNs have been applied in many fields in the past decades. These applications include environmental monitoring, structural protection, military surveillance information collection etc. Typical WSNs are composed of massive small-sized sensor nodes with communication, sensing and wireless computing capabilities [2]. These nodes are often deployed randomly in the sensor fields, forming an infrastructure-less network and performing tasks in a collaborative manner. An example of WSNs used for monitoring and analysis purposes is shown in Fig. 1.



**Fig. 1.** An example of wireless sensor networks

In many situations, WSNs are expected to operate without manual intervention for a long period. During the lifetime of WSNs, it is required to update their code image, upgrade the software and maintain over-the-air firmware to cope with the changes of the internal and external environment. Considering the large number of deployed sensor nodes and the complex network environment, it is infeasible to physically collect previously deployed nodes. In other words, the process of updating messages need to be completed “over-the-air”, using the existing network itself [5]. To satisfy the above purposes, an effective bulk data dissemination protocol is required, which can reliably disseminate new binaries to all nodes in the network.

Two most commonly used schemes in data dissemination are flooding and negotiation, but none of them performs well during the whole dissemination process. In flooding methods [19–21], each node directly broadcasts the received messages, and thus the flooding process can be quite quick. However, since broadcasting has no ACK or NACK, the senders have no idea of which packets is lost. To improve the reliability, all packets should be retransmitted. Once flooding scheme is applied into dense networks, it results in redundancy, contention and collision problems so as to extend the completion time. Negotiation-based protocols, like Deluge [25], CORD [26] and CoCo+ [27], adopt control messages acting as the NACK to avoid those problems in flooding and guarantee the reliability. Nevertheless, it inclines to prolong completion time since the additional control process postpones data distributions.

To take full advantages of the characteristics of flooding and negotiation, Zheng *et al.* propose SurF in [6], which selectively utilizes negotiation and

leverages flooding opportunistically through a time-reliability model. The model can predict the time efficiency of negotiation and flooding schemes. SurF integrates the reliable model to dynamically seize the optimal transition point to switch over between those two schemes and to utilize the better one to shorten completion time of the dissemination process. Moreover, Zheng *et al.* prove that SurF can improve the efficiency without sacrificing reliability. Note that the working environment in SurF is supposed to be trustworthy. However, the network environment is cluttered with various hazards in reality.

Opponents are real and they threaten the normal operation of WSNs all the time. These security vulnerabilities may cause different attacks like eavesdropping [34], pollution attacks [30], wormhole attacks [35] etc. Hence, in this paper, we improve the SurF by taking the process of packet preprocessing and verification into account. In addition, we also enhance the time-reliability accordingly.

### 1.1 Our Contributions

The contributions of this paper are summarized as follows.

- SurF and DiDrip are reviewed in this paper. We find that (1) SurF is efficient but not secure enough; (2) DiDrip improves the security of data dissemination protocols but it prolongs the completion time.
- To address those problems, we propose a secure enhanced protocol called S-SurF. It integrates the process of data authentication of DiDrip into SurF, and hence improves the security of SurF.
- We optimize the time-reliability model proposed in SurF by considering the time used for data preprocessing and verification. Through the model, the most opportune moment to put negotiation into use is found.

### 1.2 Organization

The remainder of this paper is organized as follows. In the following section, related works are presented. Two protocols named SurF and DiDrip are reviewed in Sect. 3. In Sect. 4, the analytical model and the design of S-SurF are proposed. In Sect. 5, the evaluations of efficiency and security are presented. Finally, the conclusion of this paper is covered.

## 2 Related Works

In the literature, many bulk data dissemination protocols [9, 16–27, 37] have been proposed for WSNs. However, due to the limitation of WSNs, the most perfect scheme is not easy to be designed. In [20], the authors describe the flooding scheme, which broadcasts received messages immediately, but may result in broadcast storm problem. To alleviate this problem, Kyasanur *et al.* propose an adaptive probabilistic flooding protocol in Smart Gossip [37], it reduces redundancy by probabilistic rebroadcasting. In addition, Lou and Wu present DCB in



[17], which avoid redundant dissemination by adopting sender selection scheme. However, due to the process of selecting sender is highly energy consumed, Feng *et al.* suggest to leverage beamforming and allow multiple senders to distribute the same packet with minimal transmission power in [18]. Later, more and more researchers like Hui note that those methods mentioned before can only alleviate broadcast storm problem, rather than eliminate it. Therefore, methods based on negotiation [9, 22–27, 36] are put forward. For example, Hui and Culler introduce Deluge in [25], Deluge works in three way handshake mechanism which largely reduces the number of distributed data items. The disadvantage of Deluge is that it follows a centralized reprogramming. Compared with Deluge, CORD [26] follows a distributed reprogramming approach where multiple users are involved in data dissemination. Recently, Zheng *et al.* find that negotiation is not always needed in the whole data distribution process, and propose SurF [6] which takes flooding as a substitute.

### 3 Review of SurF and DiDrip

In this section, we briefly review SurF and the building process of time-reliability model [6] firstly. Then, we analyze the security vulnerabilities in SurF. Finally, to enhance the security of SurF, we learn the DiDrip [7, 8].

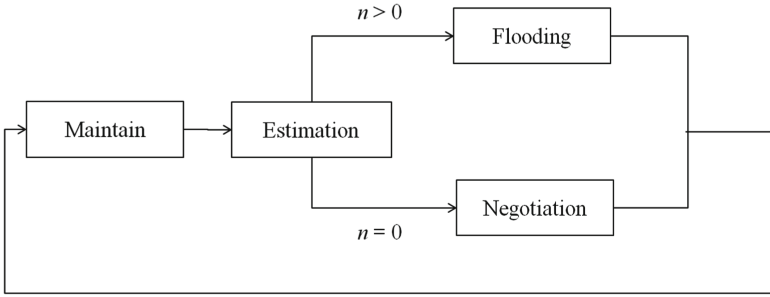
#### 3.1 Survival of the Fittest (SurF)

To some extent, SurF is more efficient and reliable than other bulk data dissemination protocols. Because it is not only dependent on flooding, but also related to negotiation. Here SurF is briefly described.

**The mechanics of SurF.** The superiority of SurF is reflected in dynamic usage of flooding and negotiation. To improve efficiency, a time-reliability model is integrated with SurF. The objective function of this model is  $T_{ij}^H(n, \phi, q_{ij})$ , the completion time of the dissemination task of sensor node  $i$  in SurF, where  $n$  is the round of flooding,  $\phi$  is the required reliability and  $q_{ij}$  is the link quality from node  $i$  to its neighbor  $j$ . We minimize the completion time by finding the  $n'$ , in which  $T_{ij}^H(n', \phi, q_{ij}) = \min \{T_{ij}^H(n, \phi, q_{ij})\}$ .

Based on  $n'$  figured out before, we further derive the best opportunity to put negotiation into use. As shown in Fig. 2, If  $n = 0$ , SurF switches to negotiation; if  $n > 0$ , switches to flooding. Note that  $n$  is not a fixed value, and it changes with the network environment. More details about SurF and its model please refer to [6].

**Security vulnerabilities of SurF.** It is worth noting that the operating environment is assumed to be trustworthy in SurF and no adversary is existed in WSNs. However, opponents are real in reality, and they threaten the normal operations of WSNs [12].



**Fig. 2.** Schematic representation of SurF

Attacks considered in this paper only aim to corrupt the data integrity and drain the limited resources. In WSNs, opponents are easily access to those data packets broadcasted in dissemination process, and those packets may well be modified or erased by them. However, without mechanism to authenticate the received data items, receivers can not affirm whether the messages are what they really want. Furthermore, since bulk data dissemination asks each node to distribute its new data packets, adversaries may well exploit this inherent feature to launch DoS attacks to exhaust limited battery energy of sensor nodes. For example, adversaries can maliciously inject fake data to the network, and force all nodes waste energy in constantly distributing/updating fake data items.

We admit that most of the bulk data dissemination protocols are not secure enough, not excluding SurF. Moreover, SurF may suffer more serious consequences for that it is quicker than other protocols. To improve the security of SurF, we have learned DiDrip, a secure protocol used for detecting whether messages has been maliciously modified.

### 3.2 DiDrip

Those years, security plays an increasing important role in the design of data dissemination protocols, and many schemes about data authentication are presented in [3–11, 13–15, 28–33]. In this paper, we select DiDrip into use.

DiDrip allows network owners and authorized users with different privileges to disseminate data items without relying on the base station. The essential idea of DiDrip is to make full use of Merkle hash tree [13] to reduce the number of public key operations. It consists of four phases: system initialization, user joining, packet preprocessing and packet verification.

In system initialization phase, the network owner derives private key  $x$  and corresponding public parameter  $y$ . After that, the public parameter  $y$  is preloaded in all nodes of the network.

In user joining phase, if a user with the identity  $UID_j$  hopes to obtain privilege level, he should register with the base station by giving his identity and other details. Firstly, user  $U_j$  chooses the private key  $SK_j$  and derives the public key  $PK_j = SK_j \cdot Q$ . Secondly,  $U_j$  submits a 3-tuple  $\langle UID_j, Pri_j, PK_j \rangle$  given  $Pri_j$

is the privilege level of  $U_j$ . At last, the network owner generates the certificate  $Cert_j = \{UID_j, PK_j, Pri_j, SIG_x \{h(UID_j || PK_j || Pri_j)\}\}$ .

In packet preprocessing phase, we have two methods (hash chain and hash tree) to construct the packets of the respective data. Take Merkle hash tree for example, all the hash values of data items are viewed as the leaves of the tree, and each internal node is computed as the hash value of the concatenation of two child nodes. The process is continued until the tree is created. After that,  $U_j$  signs the root node and transmits the advertisement packet  $P_0$ . Finally, he distributes data packets to other nodes.

In verification phase, the received packets are divided into two categories: advertisement packets and data packets. If the received packet is an advertisement packet  $P_0 = \{Cert_j, root, SIG_{SK_i} \{root\}\}$ , receiver  $S_j$  then sequentially authenticates the dissemination privilege  $Pri_j$ , certificate, signature; Otherwise, it is a data packet,  $S_j$  checks the authenticity and integrity of  $P_i$  through the already identified root node. Please refer to [7] for more details.

## 4 S-SurF: A Novel Bulk Data Dissemination Protocol

Zheng *et al.* illustrate the superiority of SurF in efficiency theoretically and experimentally in [6], but they ignore the importance of security. Without authentication, the network is vulnerable to various hazards like DoS attacks. To enhance the security of SurF, we propose a novel protocol named S-SurF in which the advantages of SurF and DiDrip are combined. Moreover, an enhanced time-reliability model is presented.

### 4.1 An Enhanced Time-Reliability Model

To seize the fittest transition point, Zheng *et al.* introduce a reliable model without taking potential hazards of the network into account. To compensate for the limitation, we propose S-SurF and optimize the time-reliability model in this paper. In the model, we simply regard the local optimality as the objective of optimization. Therefore, we predict  $T_{ij}^{SH}(n, \phi, q_{ij})$ , the completion time of node  $i$  in S-SurF.  $T_{ij}^{SH}$  can be estimated as:

$$T_{ij}^{SH}(n, \phi, q_{ij}) = T_{prep}^{DATA} + T_{ij}^{SF}(n) + T_{ij}^{SN}(\phi, R_j, q_{ij}). \quad (1)$$

Notations used in S-SurF are listed below.

- $N$ , the number of packets of one page.
- $T_{pkt} + T_{back}$ , the average transmission time per packet, given the expected back-off time  $T_{back}$ .
- $N_{supp}$ , the number of suppressed ADVs in negotiation scheme.
- $K$ : the depth of Merkle hash tree.
- $plr$ : the packet loss rate.
- $\tau_l$ : the expected time between two successive ADVs.

- $T_{h(\cdot)}$ : the time used for once of hash operation.
- $R_j(n, R_j^0)$ : the expected reliability of  $j$  after node  $i$  flooding the data  $n$  times, given that  $j$  already has the reliability of  $R_j^0$ .

$$R_j(n, R_j^0) = 1 - (1 - R_j^0) \cdot (1 - p_i q_{ij})^n. \quad (2)$$

- $a_i$ : the number of nodes in  $i$ -th level ( $1 \leq i \leq K$ ) of the Merkle hash tree. According to the characteristics of Merkle hash tree, we conclude that:

$$a_1 = 1, a_i = \left\lceil \frac{a_{i+1}}{2} \right\rceil, a_k = N. \quad (3)$$

- $T_{prep}^{DATA}$ : the time used in packet preprocessing phrase, it can also be defined as the time needed to create a Merkle hash tree. Considering that all packets disseminated should be preprocessed,  $T_{prep}^{DATA}$  can be evaluated as:

$$T_{prep}^{DATA} = \left( \sum_{i=1}^K a_i \right) \cdot T_{h(\cdot)}. \quad (4)$$

- $T_{verify}^{DATA}$ : the expected verification time of per packet. As shown in Fig. 4, if a node wants to verify  $d_1$  contained in  $P_1$ , it just need to check if  $h(h(h(d_1)||e_2)||e_{3-4}) = e_{1-4}$ . So can be defined as:

$$T_{verify}^{DATA} = K \cdot T_{h(\cdot)}. \quad (5)$$

- $T_{ij}^{SF}(n)$ : the time needed in flooding phrase in S-SurF, which consists of the time used for flooding and verification. For simplicity, we assume the verification results of privilege, certificate and signature are all positive. Given that only these packets received by receiver ought to be authenticated,  $T_{ij}^{SF}(n)$  can be computed as:

$$T_{ij}^{SF}(n) = n \cdot N \cdot (T_{pkt} + T_{back}) + n \cdot N \cdot (1 - plr) \cdot T_{verify}^{DATA}. \quad (6)$$

- $T_{ij}^{SN}(\phi, R_j, q_{ij})$ : the time used for negotiation from  $i$  to  $j$  in S-SurF, given the already achieved reliability  $R_j$ . As shown in Fig. 3, the time of the negotiation phrase  $T_{ij}^{SN}$  comprises three parts:  $T_{ij}^{SADV}$ ,  $n_2 \cdot T_{ij}^{SREQ}$ ,  $T_{ij}^{SDATA}$ . Note that  $n_2$  presents the REQ transmission rounds in negotiation.

$$T_{ij}^{SN}(\phi, R_j, q_{ij}) = T_{ij}^{SADV} + n_2 \cdot T_{ij}^{SREQ} + T_{ij}^{SDATA}. \quad (7)$$

- $T_{ij}^{SADV}$ : the time used for ADV transmission. Compared with  $T_{ij}^{ADV}$  of SurF, we take the time of packet preprocessing and verification into account. It is:

$$T_{ij}^{SADV} = \tau_l \cdot \left( \frac{1}{q_{ij}} - 1 + N_{supp} \right) + T_{prep}^{ADV} + (n_{ADV} - N_{supp}) \cdot (1 - plr) \cdot T_{verify}^{ADV}, \quad (8)$$

where  $T_{prep}^{ADV}$  is the time of packets (ADV) preprocessing.

- $n_{ADV}$ : the number of ADV that should have been sent by the sender, which is:

$$n_{ADV} = \frac{T_{ij}^{ADV}}{\tau} = \frac{1}{q_{ij}} - 1 + N_{supp}. \tag{9}$$

- $T_{verify}^{ADV}$ : the verification time of per ADV, which can be described as:

$$T_{verify}^{ADV} = T_{check}^{Pri} + T_{check}^{Cret} + T_{check}^{SIG}, \tag{10}$$

where  $T_{check}^{Pri}$ ,  $T_{check}^{Cret}$  and  $T_{check}^{SIG}$  represent the time of checking privilege  $Pri_j$ , certificate, signature respectively.

- $T_{ij}^{SREQ}$ : the time of per round REQ transmission. Since REQ is not discussed in DiDrip, we simply assume that the value of  $T_{ij}^{SREQ}$  is equal to  $T_{ij}^{REQ}$ .
- $T_{ij}^{SDATA}$ : the time used for DATA dissemination. Even though the process of DATA transmission scatters in different rounds, the time can be computed by the expected number of transmitted packets. That is:

$$T_{ij}^{SDATA} = \frac{1 - R_j(n, 0)}{q_{ij}} \cdot N \cdot (T_{pkt} + T_{back}) + n_{DATA} \cdot (1 - plr) \cdot T_{verify}^{DATA}. \tag{11}$$

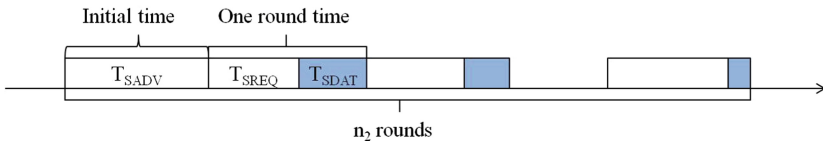
- $n_{DATA}$ : the number of DATA sent by node  $i$ , which is:

$$n_{DATA} = \frac{T_{ij}^{DATA}}{T_{pkt} + T_{back}} = \frac{1 - R_j(n, 0)}{q_{ij}} \cdot N. \tag{12}$$

- Note that we define the completion time above where  $n_2 \neq 0$ . Here we redefine  $T_{ij}^{SH}$  in a general form. That is:

$$T_{ij}^{SH}(n, \phi, q_{ij}) = T_{prep}^{DATA} + \begin{cases} T_{ij}^{SN}(\phi, 0, q_{ij}) & n = 0 \\ T_{ij}^{SF}(n) + T_{ij}^{SN}(\phi, R_j, q_{ij}) & 0 < n < N_F \\ T_{ij}^{SF}(n) & n = N_F \end{cases} \tag{13}$$

given  $N_F$  is the upper bound of floodings. It is obvious that the optimal transition point between flooding and negotiation can be represented by  $n'$ , where  $T_{ij}^{SH}(n', \phi, q_{ij}) = \min \{T_{ij}^{SH}(n, \phi, q_{ij})\}$ .



**Fig. 3.** Composition of the completion time of the negotiation scheme

### 4.2 Design of S-SurF

A novel bulk data dissemination protocol should be quick, reliable and secure. To satisfy these requirements, we integrate DiDrip into SurF and further propose S-SurF. Referring to Fig. 4, S-SurF is composed of four phrases: packet preprocessing, flooding, negotiation and data authentication. In packet preprocessing phrase, if a valid user wants to disseminate data items, he will need to construct the data dissemination packets first. In flooding and negotiation phrase, users distribute data packets by ways of flooding and negotiation accordingly. In data authentication phrase, each node verifies received packets. If the result is positive, the node updates the data based on received items. In the following subsections, each phase is described in detail.

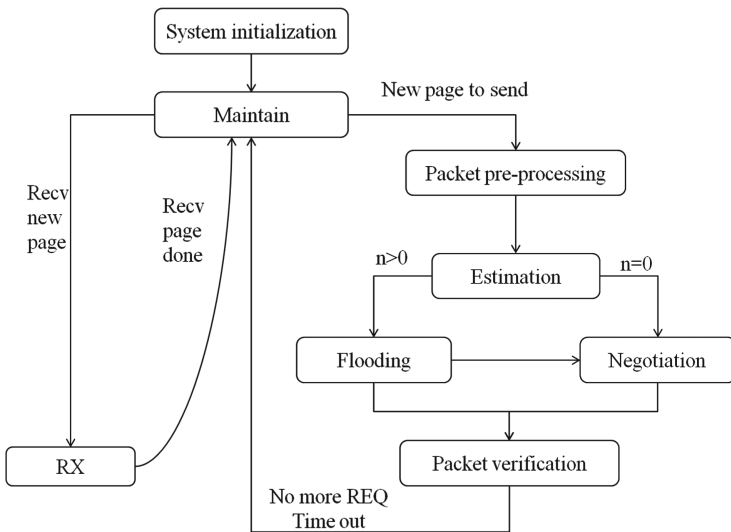
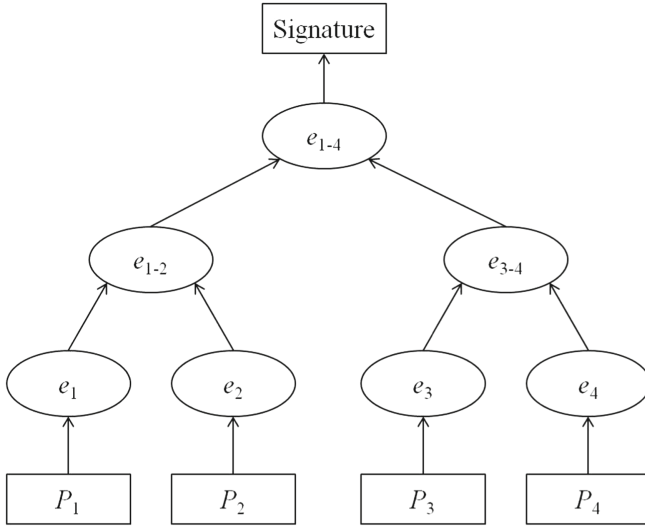


Fig. 4. Example of Merkle hash tree

**Packet preprocessing.** Suppose that a user with  $UID_j$  wants to distribute  $n$  data items:  $d_i = \{key_i, version_i, data_i\} (i = 1, 2, 3, \dots, n)$ , he should first construct those packets by creating a hash chain or a Merkle hash tree. Merkle hash tree method is chosen in this paper since it allows each packet to be immediately authenticated upon its arrival.

We illustrate the building process of Merkle hash tree by considering  $n = 4$  as an example. All the hash values of packets  $e_i = h(P_i)$  are used as the leaves of the Merkle hash tree, as depicted in Fig. 5. Here  $h(\cdot)$  presents a public one-way cryptographic hash function (e.g. SHA-1). Each internal node is computed as the hash value of two adjacent leaf nodes, for example  $e_{1-2} = h(e_1 || e_2)$ . The process continues until the root node is formed. Note that before disseminating data items,  $U_j$  signs the root node and transmits the advertisement packet  $P_0 = \{Cert_j, root, SIG_{SK_i} \{root\}\}$ .



**Fig. 5.** State transition diagram of S-SurF

**Flooding.** Based on the result of the enhanced time-reliability model, if  $n > 0$ , S-SurF switches into flooding state. In flooding phase, we use the probabilistic flooding [6, 16, 20] as the dissemination scheme. When sensor node receives a new message, the node will then rebroadcast it with the probability  $P$ . To alleviate the broadcast storm problem, the node should wait for a random amount of time before rebroadcasting the received messages. Moreover,  $P$  can be derived based on the following rules:

- If  $\exists j, I_j = 1$ , then set  $P = 1$ ;
- If  $\forall j, I_j = 0$ , then set  $P = 0$ ;

Otherwise, set  $P = \alpha P' \bar{I}_j + (1 - \alpha) P'$ , where  $I$  is the dependence indicators,  $P'$  is the original rebroadcasting probability and  $\alpha$  is the coefficient which represents the weight.

**Negotiation.** If the value of  $n$  of the enhanced time-reliability model is equal to 0, S-SurF switches into negotiation state. In this phrase, we put Deluge [6, 22, 25], a standard dissemination protocol of TinyOS into use. Three kinds of packets and operations are involved in Deluge: ADV, the advertisement packets which are used to periodically broadcast the version of possessed data; REQ, after hearing the ADV messages, neighboring nodes send REQ messages to ADV messages sender if a newer version is detected; DATA, the requested data will be packeted into DATA messages and sent to DATA messages sender.

**Data verification.** In this phase, each received packet is verified by receiver. For advertisement packet  $P_0$ , receiver  $S_j$  first check the legality of dissemination privilege  $Pri_j$  and further estimate whether he possesses the access right of

forthcoming ADVs. If the result is positive,  $S_j$  uses the public key of the network owner to run an *ECDSA verify* operation to authenticate the certificate  $Cert_j$ . If  $Cert_j$  is valid,  $S_j$  then authenticates the signature. For data packets  $P_i$  (such as  $P_1$ ), it can be verified immediately by checking if  $h(h(h(d_1)||e_2)||e_{3-4}) = e_{1-4}$ . Finally, if the verification result is positive and the version number is new, receiver updates the old data according  $P_i$ ; Otherwise,  $P_i$  is discarded.

## 5 Evaluation

Considering that the WSNs are limited in energy and vulnerable to interference from external environments, a remarkable bulk data dissemination protocol should at least be efficient and secure. To prove our model and protocol is novel, the efficiency and security of them are analyzed in this section.

### 5.1 Performance Analysis

We notice that both flooding and negotiation are not always needed in the whole process of data dissemination, because flooding induces blind retransmission/broadcast storm problems and negotiation prolongs the completion time. In S-SurF, we integrate those two schemes and always put the more suitable one into use. For example, S-SurF leverages negotiation in time when flooding is considered as inefficiency. Thus, compared with S-SurF, neither flooding nor negotiation used alone is efficient enough.

In S-SurF, we can seize the best transition point through the time-reliability model. We also find the optimal flooding rounds  $n'$  and predict the completion time, that is:  $T_{ij}^{SH}(n', \phi, q_{ij}) = \min\{T_{ij}^{SN}(\phi, 0, q_{ij}), T_{ij}^{SF}(n) + T_{ij}^{SN}(\phi, R_j, q_{ij}), T_{ij}^{SF}(N_F)\} + T_{prep}^{DATA}$ . In any way,  $T_{ij}^{SH}(n', \phi, q_{ij}) \leq \min\{T_{ij}^{SN}(\phi, 0, q_{ij}), T_{ij}^{SF}(N_F)\} + T_{prep}^{DATA}$ .

### 5.2 Security Analysis

In this paper, S-SurF are used to resist problems mentioned in Sect. 3.1.

**Data Integrity.** In S-SurF, authenticated users sign the root of the Merkle hash tree with their private key. Moreover, each node possesses public keys of other nodes and can verify the certification of senders in terms of network owners public key. Then, receiver authenticates the root of the hash tree and data packets orderly. Hence, once the forged data items appear, it can be easily detected.

**Resistance Against DoS Attack.** Based on symmetric decryption and hash function operations, packet verification is a very efficient method to resist against DoS attack. Firstly, receiver immediately authenticates received packets through a few hash operations. Secondly, it rejects/filters any bogus data item from the adversary. As a result, even if adversaries ood the nodes with a great deal of packets, not much computation resource is depleted to verify the packets. And hence, S-SurF can prevent from the DoS attacks to some extend.



## 6 Conclusion

In WSNs, bulk data dissemination is used for system updating. To satisfy the requirements of efficiency and security, we propose an enhanced bulk data dissemination protocol named S-SurF in this paper. In S-SurF, the efficiency of the protocol is guaranteed by the combination of flooding and negotiation, and the security is actualized by the integration of SurF and DiDrip. Moreover, to seize the optimal transition point between two schemes, an optimized time-reliability model is proposed. S-SurF is safer than SurF and quicker than DiDrip. In the further, we plan to prove the superiority of S-SurF experimentally and further study the potential performance improvements.

**Acknowledgments.** This work is supported by the National Natural Science Foundation of China under Grant Nos. 61672295, 61672290 and U1405254, the State Key Laboratory of Information Security under Grant No. 2017-MS-10, the 2015 Project of six personnel in Jiangsu Province under Grant No. R2015L06, the CICAET fund, and the PAPD fund.

## References

1. Akyildiz, I., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. *Comput. Netw. Int. J. Comput. Telecommun. Netw.* **38**(4), 393–422 (2002)
2. Marco, A., Casas, R., Ramos, J.L.S., Coarasa, V., Asensio, A., Obaidat, M.S.: Synchronization of multihop wireless sensor networks at the application layer. *IEEE Wirel. Commun.* **18**(1), 82–88 (2011)
3. Shen, J., Shen, J., Chen, X., Huang, X., Susilo, W.: An efficient public auditing protocol with novel dynamic structure for cloud data. *IEEE Trans. Inf. Forensics Secur.* (2017). <https://doi.org/10.1109/TIFS.2017.2705620>
4. Yin, J., Madria, S.: SecRout: a secure routing protocol for sensor networks. In: *International Conference on Advanced Information Networking and Applications*, pp. 393–398 (2006)
5. Huang, L., Setia, S.: *Reliable Bulk Data Dissemination in Sensor Networks*, vol. 14, no. 3, pp. 115–124. George Mason University (2007)
6. Zheng, X., Wang, J., Dong, W., He, Y., Liu, Y.: Bulk data dissemination in wireless sensor networks: analysis, implications and improvement. *IEEE Trans. Comput.* **65**(5), 1428–1439 (2016)
7. He, D., Chan, S., Guizani, M., Yang, H., Zhou, B.: Secure and distributed data discovery and dissemination in wireless sensor networks. *IEEE Trans. Parallel Distrib. Syst.* **26**(4), 1129–1139 (2015)
8. He, D., Chan, S., Guizani, M.: Small data dissemination for wireless sensor networks: the security aspect. *IEEE Wirel. Commun.* **21**(3), 110–116 (2014)
9. Iabbassen, D., Moussaoui, S.: Data dissemination protocols in wireless sensor networks client/server versus mobile agent paradigms. In: *International Conference on Innovative Computing Technology*, pp. 45–50 (2015)
10. Shen, J., Zhou, T., He, D., Zhang, Y., Sun, X., Xiang, Y.: Block design-based key agreement for group data sharing in cloud computing. *IEEE Trans. Dependable Secure Comput.* (2017). <https://doi.org/10.1109/TDSC.2017.2725953>

11. He, D., Chan, S., Zhang, Y., Yang, H.: Lightweight and confidential data discovery and dissemination for wireless body area networks. *IEEE J. Biomed. Health Inform.* **18**(2), 440–448 (2014)
12. Mavoungou, S., Kaddoum, J., Taha, M., Matar, G.: Survey on threats and attacks on mobile networks. *IEEE Access* **4**, 4543–4572 (2016)
13. Merkle, R.: Protocols for public key cryptosystems. In: *IEEE Security Privacy*, pp. 122–134 (1980)
14. Jose, J.: Secure data dissemination protocol in wireless sensor networks using XOR network coding. *Int. J. Adv. Res. Comput. Commun. Eng.* **3**(27), 57–61 (2014)
15. Shen, J., Chang, S., Shen, J., Liu, Q., Sun, X.: A lightweight multi-layer authentication protocol for wireless body area networks. *Future Gener. Comput. Syst.* (2016). <https://doi.org/10.1016/j.future.2016.11.033>
16. Sasson, Y., Cavin, D., Schiper, A.: Probabilistic broadcast for flooding in wireless mobile ad hoc networks. *Wirel. Commun. Netw.* **2**, 1124–1130 (2013)
17. Lou, W., Wu, J.: Toward broadcast reliability in mobile ad hoc networks with double coverage. *IEEE Trans. Mob. Comput.* **6**(2), 148–163 (2007)
18. Feng, J., Lu, Y., Jung, B., Peroulis, D., Hu, Y.: Energy-efficient data dissemination using beamforming in wireless sensor networks. *ACM Trans. Sens. Netw.* **9**(3), 31 (2013)
19. Li, Y., Bartos, R.: Efficient regional information dissemination protocol for intermittently connected mobile wireless sensor networks. In: *Southeastcon*, pp. 1–8 (2016)
20. Ni, S., Tseng, Y., Chen, Y., Sheu, J.: The broadcast storm problem in a mobile ad hoc network. In: *Proceedings of ACM MobiCom* (1999)
21. Sabbineni, H., Chakrabarty, K.: Location-aided flooding: an energy-efficient data dissemination protocol for wireless-sensor networks. *IEEE Trans. Comput.* **54**(1), 36–46 (2005)
22. Kifah, S., Abdullah, S.: An adaptive non-linear great deluge algorithm for the patient-admission problem. *Inf. Sci.* **295**, 573–585 (2015)
23. Perrig, A., Canetti, R., Tygar, J., Song, D.: Efficient authentication and signing of multicast streams over lossy channels. In: *IEEE Symposium on Security and Privacy*, p. 56 (2000)
24. Liu, A., Ning, P.: TinyECC: a configurable library for elliptic curve cryptography in wireless sensor networks. In: *International Conference on Information Processing in Sensor Networks*, pp. 245–256 (2008)
25. Hui, J., Culler, D.: The dynamic behavior of a data dissemination protocol for network programming at scale. In: *The ACM Conference on Embedded Networked Sensor Systems*, pp. 81–94 (2004)
26. Huang, L., Setia, S.: Cord: energy-efficient reliable bulk data dissemination in sensor networks. In: *Proceedings of IEEE FOCOM*, vol. 14, no. 3, pp. 115–124 (2008)
27. Zhao, Z., Bu, J., Dong, W., Gu, T., Xu, X.: CoCo+: exploiting correlated core for energy efficient dissemination in wireless sensor networks. *Ad Hoc Netw.* **37**, 404–417 (2016)
28. Shen, J., Wang, C., Castiglione, A., Liu, D., Esposito, C.: Trustworthiness evaluation-based routing protocol for incompletely predictable vehicular ad hoc networks. *IEEE Trans. Big Data* (2017). <https://doi.org/10.1109/TBDDATA.2017.2710347>
29. Bagga, N., Sharma, S., Jain, S., Sahoo, T.: A cluster-tree based data dissemination routing protocol. *Proc. Comput. Sci.* **54**, 7–13 (2015)

30. Yu, Z., Wei, Y., Ramkumar, B., Guan, Y.: An efficient signature-based scheme for securing network coding against pollution attacks. In: IEEE International Conference on Computer, pp. 1409–1417 (2008)
31. Murugadoss, G., Sivakumar, P., Manikandan, R.: Secure and scattered data sighting and dissemination in wireless sensor networks. *Aust. J. Basic Appl. Sci.* **10**(2), 114–124 (2016)
32. Shen, J., Liu, D., Liu, Q., Sun, X., Zhang, Y.: Secure authentication in cloud big data with hierarchical attribute authorization structure. *IEEE Trans. Big Data* (2017). <https://doi.org/10.1109/TBDATA.2017.2705048>
33. Farooq, M., Zhu, Q.: Secure and reconfigurable network design for critical information dissemination in the internet of battlefield things (IoBT). In: International Symposium on Modeling and Optimization in Mobile (2017)
34. Li, X., Xu, J., Dai, H., Zhao, Q., Cheang, C., Wang, Q.: On modeling eavesdropping attacks in wireless networks. *J. Comput. Sci.* **11**, 196–204 (2015)
35. Shrivastava, A., Dubey, R.: Wormhole attack in mobile ad-hoc network: a survey. *Int. J. Secur. Appl.* **9**(7), 293–298 (2015)
36. Kulik, J., Heinzelman, W., Balakrishnan, H.: Negotiation-based protocols for disseminating information in wireless sensor networks. *Wirel. Netw.* **8**(2–3), 169–185 (2002)
37. Kyasanur, P., Choudhury, R., Gupta, I.: Smart gossip: an adaptive gossip-based broadcasting service for sensor networks. In: IEEE International Conference on Mobile AdHoc and Sensor Systems, pp. 91–100 (2006)

# MCloud: Efficient Monitoring Framework for Cloud Computing Platforms

Jijun Zeng<sup>1,2</sup>, Zhenyue Long<sup>1,2</sup>, Guiquan Shen<sup>1,2</sup>, Lihao Wei<sup>1,2</sup>,  
and Yunkui Song<sup>3</sup>(✉)

<sup>1</sup> Information Center, Guangdong Power Grid Co. Ltd.,  
Guangzhou 510000, China

<sup>2</sup> CSG-Key Laboratory of Information Technology Testing,  
Guangzhou 510000, China  
{zengjijun, longzhenyue,  
shenguiquan, weilihao}@gdxx.csg.cn

<sup>3</sup> Institute of Software, Chinese Academy of Sciences, Beijing 100190, China  
songyk@otcaix.iscas.ac.cn

**Abstract.** Cloud computing platforms have the characteristics of large scale, complex interaction and dynamic environment, which cause it more and more serious to software and system security, and the monitoring technology is one of the key technologies to ensure the reliability. However, the spatial topology of a cloud computing platform often changes, so static monitoring can bring huge network overhead when collecting multi-node and multi-level monitoring data. To address the above problems, this paper proposes MCloud, an efficient monitoring framework for cloud computing platform. We first propose an organization model for monitoring objects with a tree-linked list. Then, we optimize the indexing mechanism for the efficient retrieval of monitoring data. Finally, we design and implement a monitoring framework MCloud integrated in cloud computing platforms, and the experimental results show that MCloud can effectively reduce the monitoring performance of the cloud platform by more than 20%.

**Keywords:** Runtime monitoring · Monitoring framework · Cloud computing  
System reliability · Monitoring period

## 1 Introduction

Cloud computing has become a mainstream computing paradigm to construct new information systems, and has the characteristics of large scale, complex composition, dynamic environment with continuous evolution, which makes service failures more and more serious, and then how to guarantee the continued reliable operations of cloud computing platforms has become a real challenge. The distributed monitoring system provides the data source for analyzing platforms' running state, and is the premise to guarantee the reliable operations of the platforms, so gradually becomes the focus of industries and the hotspot of academic researches. At present, cloud monitoring systems mainly have centralized and distributed models. The centralized scheme has the

advantages of easy deployment and small delay, but it has high overhead and poor scalability. The hierarchy distributed scheme is scalable and suitable for the large-scale monitoring demand, but it has a long usually delay. The size of the cloud computing platform often reaches hundreds of physical servers and tens of thousands of virtual machines, the spatial topology continues to change, operations (e.g., increase, delete, and migrate virtual machines) make monitoring data model change, and thus accurately locating monitoring data is difficult. At the same time, collecting monitoring data of multi-node and multi-level can bring huge network overhead affecting network performance.

To address the above problem, this paper proposes a low-cost monitoring framework MCloud for cloud computing platforms, which aims to the monitoring requirements of physical and virtual machines, adopts improved hierarchy architectures, and groups according to physical servers which virtual machines belong to. Physical servers as collector nodes collect and perpetuate the monitoring data running on it. Main management node is only responsible for forwarding requests, transmitting monitoring data between nodes and clients, and thus can efficiently reduce the network overhead. Furthermore, we propose a hybrid object organization model based on tree-linked list, which contains the historical relationship of mapping physical machines to virtual machines, and combines the monitoring data organization model based on a tree, to provide efficient requests within the time window of monitoring data. Finally, based on a real cloud computing platform, we design and implement a monitoring framework, and the experimental results show that the framework can efficiently reduce the overload of monitoring performance of the cloud platform by more than 20%.

## 2 MCloud Model

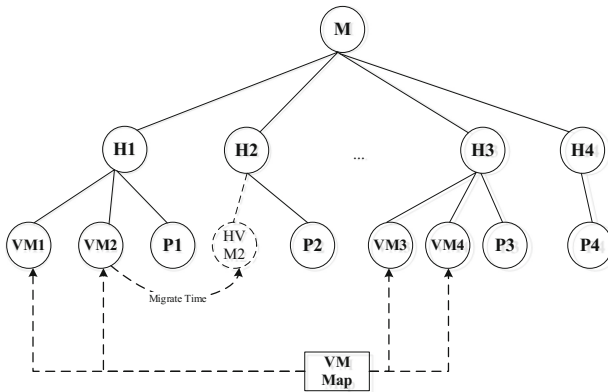
### 2.1 MCloud for Monitoring Cloud

Monitoring cloud platforms deployed in virtual environments collects the information of physical resources, virtualization resources in physical layer and applications running on virtual resources. When the master node receives requests of monitoring data from clients, it locates the physical server of the monitoring data and forward the requests. Therefore, master nodes must maintain the organization model of monitored objects.

The main management node in the cloud platform, the managed physical server nodes and the virtual machines (VM) adopt one-to-many mapping mode. This paper uses a tree structure to describe the monitoring objects of a cloud computing platform. Since the tree structure has the ability to depict layers and the relations between hierarchical entities, it can well meet the characteristics of cloud monitoring systems. However, the migration operations of virtual machines may change the virtual machine's physical server in the operation process, because each physical server assembles and persists the data of all virtual machines on the physical server. Thus, we store the historical data of individual virtual machines on multiple physical servers. When the historical data of monitored nodes is checked at the current gathering node,

the monitoring data before the change will loss. Therefore, in addition to storing the tree structure of the architecture with three layers, we also need to store the change history of mapping relation between virtual machines and physical servers, because the chain table structure has the ability to describe state evolution. We use a linked list with a time label to depict the history of transformation.

When a client requests to the master node within a specific time window for monitoring specific objects, the master node will first locate the object's current position, and check whether objects the user request within the time window to provide monitoring data. However, when the object position changes frequently, the length of the linked list can make maintenance more difficult. Meanwhile, pulling the data from several different Collector nodes also brings additional performance overhead, so we provide compression operations for long linked lists.



**Fig. 1.** Architecture of collecting monitoring data

As shown in Fig. 1, the components of the MCloud include two classes of monitored nodes, namely Host and VM. The performance represents collected performance data, VMMMap is the data structure recording the migration historical information of virtual machines, the Key Value is the identity of a virtual machine, the Value is the linked list recording the history of VM location changes with a time label. So, we describe the model MCloud = {Master, Host, VM, Performance, VMMMap} as follows.

- Master (M): is a logical concept to interact with the server side. The Master node extends capabilities on the basis of the Host node. Master nodes can be considered as special Host nodes, i.e., Master = union {Hosts};
- Host (H): is a physical server with the installed virtualized software, has the life cycle management of the VM lifecycle. We present it as Host = {ID} ∪ {VMlifecycle}, and the ID is the unique identification of the Host nodes;

- VM: is essentially a computer implemented by software, but it can have the same capabilities as a real physical server. We describe it as  $VM = \{ID\} \cup \{\text{Functions provides}\}$ , where ID is the unique identification of VM. We use UUID [1] to represent it, which is the unique identification to identify VM with virtualization;
- Performance (P): refers to the monitoring data of resource consumption. We will discuss its organizational structure in detail in the next section;
- HVM: records the historical information of a VM and has the same label with VM, but also needs to record the migration time. This makes it possible to identify whether the historical data is the users' needs by comparing the migration time and the client request time window. We characterize it as  $HVM = \{VM\} + \{\text{MigrateTime}\}$ ;
- VMMap: records the change history of VM locations, Key Value is VM, and Value is a linked list that records the location of VM's history. The elements of the linked list are HVM with the same ID, which depict  $VMMMap = \{VM\} + \{HVMs\}$ .

## 2.2 Operations in MCloud

MCloud model includes operations on tree structure nodes, such as creation, insertion, migration, query, deletion, and entity type (Host, VM) decision as follows.

- Host creation and insertion: are effective when the size of the cloud platform needs to increase the scale. The new Host node of MCloud needs to consider the status of the VM running on the Host node, otherwise it will cause the loss of VM monitoring data, as shown in line 3–8 of the algorithm.

---

### Algorithm 1: Host Insertion

---

Input: Identification of Host *host\_ID* ;

Output: MCloud *MCloud*.

1. Init Entity[] VMSet
  2. Host newHost = Create(host\_ID);
  3. IF there are some VMs running on the Host
  4.     FOR VM\_ID=UUID(VM)
  5.         VMSet[i++]=new VM(VM\_ID);
  6.     END FOR
  7. ENDIF
  8. Children(MCloud).add(newHost)
- 

- VM migration: changes the upper Collector node of the VM, meanwhile the historical information of the VM remains in the old host. To search the historical data of virtual machines during the history retrieval, MCloud model updates in time and increase corresponding records for the VM in the VMMap, when the VM instance is migrated.

---

**Algorithm 2: VM Migration**

---

Input: Identification of target Host *Thost\_ID*, Identification of VM *vm\_ID*;

Output: MCloud.

1. VM *vm* = search(*vm\_ID*);
  2. HOST *FHost* = Parent(*vm*);
  3. HOST *THost* = search(*THOST\_ID*);
  4. HVM *hvm* = Create(*vm*, *currentTime*);
  5. Children(*FHost*).remove(*vm*);
  6. Children(*FHost*).add(*hvm*);
  7. Children(*THost*).add(*vm*);
  8. Array *VMArray* = *VMap*.get(*vm*);
  9. IF *VMArray* == null
  10.     Create(*VMArray*);
  11. ENDIF
  12. *VMArray*.add(*hvm*);
- 

- Searching resident Hosts: is for the client requests to locate physical node list of the monitoring data. For a VM object, it returns the physical nodes of all monitoring data of VMs in the time window, and reads the historical change information of the VM from *VMap*.

---

**Algorithm 3: Searching resident Hosts**

---

Input: Identification of Target Entity *ID*, Target time *lasttime* ;Output: Resident Hosts List *hostList*.

1. Entity *entity* = search(*ID*);
  2. IF isHost(*entity*)
  3.     Return *entity*
  4. END IF
  5. ELSE
  6.     *hostList*.add(*host*);
  7.     *VMArray* = *VMap*.get(*vm*);
  8.     IF *VMArray* != null
  9.         FOR HVM *hvm* in *VMArray*
  10.             IF *hvm*.migratetime > *lasttime*
  11.                 *hostList*.add(Parent(*hvm*))
  12.             END IF
  13.             ELSE
  14.                 BREAK
  15.             END ELSE
  16.         END FOR
  17.     END IF
  18. END ELSE
  19. RETURN ERROR
-



### 3 MCloud Architecture

Cloud monitoring framework MCloud adopts a hierarchical architecture as shown in Fig. 2. Master is responsible for collecting performance data of all physical servers and virtual machines and present to the user by many ways, each physical server is responsible for collecting and perpetuating all the performance data of their own and of the virtual machine running on them. The physical server and the virtual machine need to run the monitoring Agent to collect performance data. The monitoring collector module (Collector) is also required on the physical server to collect and perpetuate the performance data of the virtual machines running on it.

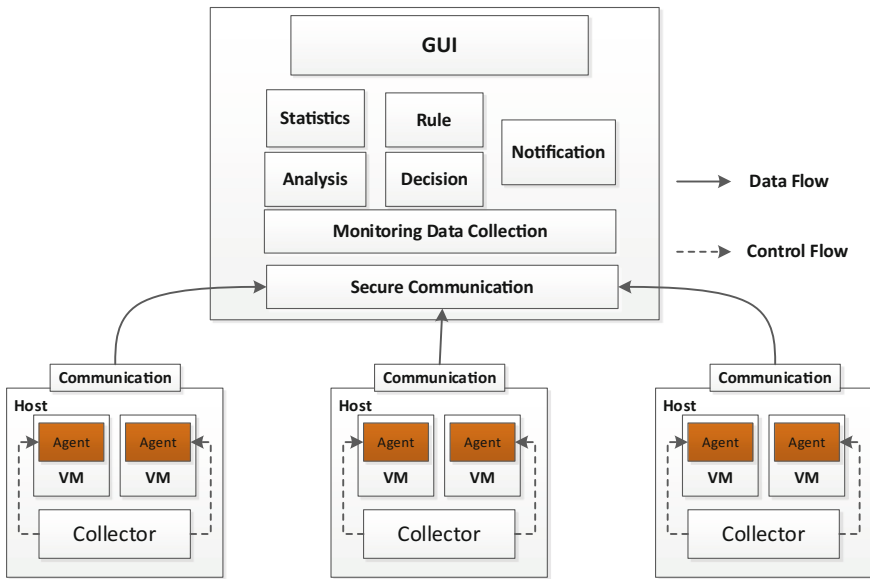


Fig. 2. Monitoring system architecture

It is important to note that, due to the advantages of virtualization, the CPU usage of the virtual machine can be directly captured by physical server [2]. Agents and collectors interact with data on a regular basis. With each fixed time slice based on users' requirements, Master collects the resource monitoring data of each node, makes statistical analysis and policy recommendations for the collected data, and then feed the results back to users.

The functional structure of the monitoring system is: monitoring agent module, monitoring collector module, master control module and data analysis and display module. These modules are deployed on different types of nodes of the monitoring system and are described in detail.

- Agent: is deployed in virtual machines, and is responsible for collecting monitoring data in the virtual machine, interacting with monitoring collector module by the way of XML-RPC [3] to, and transferring the monitoring data from inside the virtual machine to a physical server.
- Collector: monitoring collector module is deployed on a physical server, is responsible for receiving a virtual machine monitor information, and indexes the monitoring collected, and then transfer to the XML file. When receiving a data transfer request, a copy of the user's required data is generated for the user's requirements, and the monitoring data is sent to the client module through the SSH protocol.
- Control module: is deployed on the master node, responsible for putting forward the data request for related monitoring collector module according to customer's specific requirements, and lead the monitoring collector module to sends the user required monitoring data generated to the client module.
- Client: is responsible for monitoring data acquisition, analysis and display. The client is divided into statistical analysis module, decision-making recommendations module and display module. The statistical analysis module is responsible for the separation of metadata and monitoring information. The decision and suggestion module is responsible for locating the resources that reach the bottleneck on the basis of statistical and analytical data. The display module presents the required performance data to the user with charts and alerts according to the decisions and recommendations made by the decision and recommendation module.

## 4 Evaluation

MCloud manages the monitoring data, which is the basis of the following series of network transmission optimizations including the communication mechanism combined based-on-demand transmission optimization with variable cycle - event-driven. In this section, we will first test the additional performance overhead from using MCloud model, and look to whether within an acceptable scope, then examine whether the network transmission optimization algorithms can effectively reduce network overhead of the monitoring system.

The additional performance overhead of MCloud model mainly includes two parts, the first is the performance overhead of Host node, the second is performance overhead of the Master node responsible for collecting all the Host node monitoring information, and providing external service interface. This section will test the costs of these two parts separately.

### 4.1 Host Node Performance Overhead

A physical server is configured to 24-core 2.4 GHz Intel Xeon CPU, 24 GB memory, and we set the standard configuration of the virtual machine for 1-core and 1 GB memory, reserve 4 GB memory for dom0 [4] each physical server, then a single physical server run 20 virtual machines for full workload. This section tests the

overhead of the physical server using the MCloud to manage performance data. Test results are shown in Fig. 3, and the performance overhead of a physical server to collect 20 virtual machines on it is about 10%. The overhead is trivial and the performance overhead is acceptable.

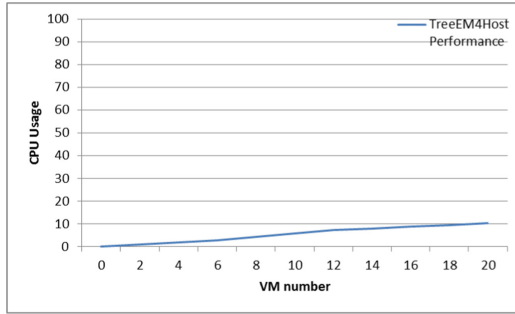


Fig. 3. Host node performance overhead

### 4.2 Master Node Performance Overhead

We discuss the performance overhead of the MCloud model on the Master node by comparing the performance overhead of Master nodes to verify that the MCloud model will not be affected by the change of the size of managed monitored data. To describe the busy or idle state of a physical server, the experiment gives two different conditions: All Hosts idle and All Hosts busy, and their description is as follows:

- All Hosts are idle: Host node except Master node has no VM running;
- All Hosts are busy: All nodes are fully loaded, which means running 20 VMS.

Figure 4 shows the compared experimental results of the MCloud performance overhead on Master in “All Hosts idle” and “All Hosts busy”.

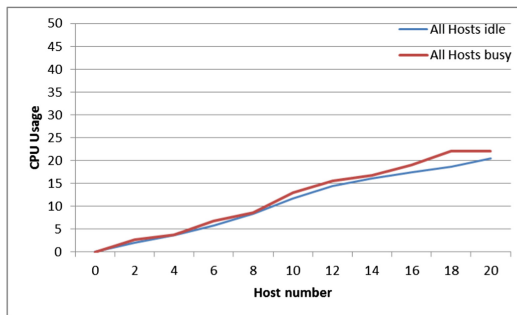


Fig. 4. Overhead comparison with workload and no workload

Thus, the performance overhead of MCloud on Master has nothing to do with the state of idle or busy Host. The performance overhead of MCloud does not fluctuate with the size of the managed performance data, and does not bring too much extra overhead in heavy workload.

### 4.3 MCloud Replica Mechanism

On-demand transmission optimization is implemented with the MCloud replica mechanism. This experiment verifies that the MCloud replica mechanism can greatly reduce the amount of network transmission between the physical server and client.

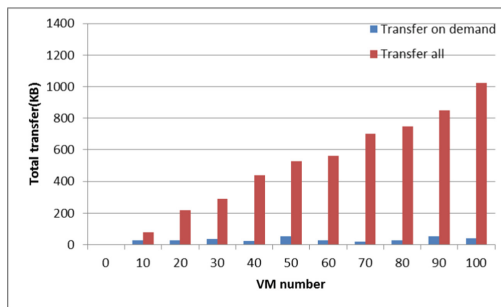


Fig. 5. Transmission network overhead comparison

The standard configuration of the virtual machine is 6-core and 2.5 GB memory on each physical server, 4G memory for dom0, and 8 virtual machines on a single physical server. We build a cloud platform for testing with 13 physical servers and 100 virtual machines. The experiment compares the growth rates between global transmission and on-demand transmission of network cost with the increasing number of virtual machines. Figure 5 shows the experimental results, which show that the transmission cost of global transmission is increasing with the increasing number of virtual machines, when the number of virtual machines on the cloud platform is more than 100 and a single transmission reached 1 MB. The on-demand transport mechanism has nothing to do with the size of the cloud platform, so a single transmission between a physical server and a client remains about 10 KB.

## 5 Related Work

Current monitoring systems for cloud monitoring platforms in a virtualization environment aim at the problems of monitoring large-scale resources for cloud computing. We introduce solutions from industries and academia for weighing both the effectiveness and overhead of monitoring systems from architecture, communication strategies and persistence. When there are a huge number of monitored nodes, the effectiveness of centralized architectures will decrease due to low response time, so it is

difficult to meet the system requirements. In hierarchy architectures, monitoring agents distributed on different nodes are logically divided into several groups with a hierarchy, and each monitoring agent belongs to the child branch of collector nodes in the higher level. Collectors group internal affairs, collect monitoring data of this group, and then provide the group to the higher level. At the top level, the root node collects the monitoring data passed from the lower collectors, and then sends control commands to the specific monitoring agent. Ganglia, a well-known distributed monitoring system, uses a hierarchy architecture with an agent Gmond collecting data and a collector Gmetad to merge data. Compared with centralized architectures, the aggregation node in the hierarchy architecture takes a part of a collector's tasks to reduce the burden of monitoring servers. When the number of nodes increases, the hierarchy architecture can ensure the normal operations. However, it also has disadvantages, such as the deployment is much more complex to determine the domain of monitored nodes in the system. In addition, because of the method of hierarchical transmission information, the monitoring information should be integrated and analyzed in several layers of collectors, which makes the system have more time delay, and thus reduces the effectiveness. The monitoring systems for cloud computing can be divided into systems by the cloud platform provider (such as System Status Dashboard [5], CloudWatch [6], etc.) and third-party monitoring products (such as CloudStatus [7], vFogLight [8], etc.). Because of cloud computing is a further development of distributed computing, parallel computing and grid computing, most cloud monitoring systems are more or less to learn from the monitoring methods of the grid computing and typical methods (e.g., Ganglia [11], Parmon [9], Supermon [10]). In addition, server virtualization is an important technology to support cloud computing, and various virtual machine manufacturers also provides monitoring products for virtualization environment (e.g., VMware vCenter [12], Citrix XenCenter [13]).

## 6 Conclusion

Cloud computing platforms have the characteristics of large scale, complex interaction and dynamic environment, which cause it more and more serious to software and system security, and the monitoring technology is one of the key technologies to ensure the reliability. However, the spatial topology of cloud computing platforms often changes, so static monitoring can bring huge network overhead when collecting multi-node and multi-level monitoring data. To address the above problems, this paper proposes MCloud, an efficient monitoring framework for cloud computing platform. We first propose an organization model for monitoring objects with a tree-linked list. Then, we optimize the indexing mechanism for the efficient retrieval of monitoring data. Finally, we design and implement a monitoring framework MCloud integrated in cloud computing platforms, and the experimental results show that MCloud can effectively reduce the monitoring performance of the cloud platform by more than 20%.

## References

1. Leach, P.J., Mealling, M., Salz, R.: A Universally Unique Identifier (UUID) Namespace (2005)
2. Nguyen, H., Shen, Z., Tan, Y., et al.: FChain: toward black-box online fault localization for cloud systems. In: IEEE 33rd International Conference on Distributed Computing Systems, pp. 21–30 (2013)
3. Allen, S., Lapp, J., Merrick, P.: XML Remote Procedure Call (XML-RPC), 11 April 2006. U.S. Patent 7,028,312
4. Menasce, D.: TPC-W: a benchmark for e-commerce. *IEEE Internet Comput.* **6**(3), 83–87 (2002)
5. Zahariev, A.: Google app engine. Helsinki University of Technology (2009)
6. Amazon CloudWatch: Amazon, Inc. Accessed 7 Feb 2010
7. Cloudstatus: Hyperic Inc. <http://www.cloudstatus.com/>
8. Foglight. <http://www.quest.com/foglight-for-virtual-enterprise-edition/>
9. Massie, M.L., Chun, B.N., Culler, D.E.: The ganglia distributed monitoring system: design, implementation, and experience. *Parallel Comput.* **30**(7), 817–840 (2004)
10. Buyya, R.: PARMON: a portable and scalable monitoring system for clusters. *Softw.-Pract. Experience* **30**(7), 723–740 (2000)
11. Sottile, M.J., Minnich, R.G.: Supermon: a high-speed cluster monitoring system. In: IEEE International Conference on Cluster Computing, pp. 39–46. IEEE (2002)
12. Bunch, C.: Automating VSphere: With VMware VCenter Orchestrator. Prentice Hall Press, Upper Saddle River (2012)
13. Citrix.: XenCenter. <http://community.citrix.com/display/xs/XenCenter>

# Secure Vibration Control of Flexible Arms Based on Operators' Behaviors

Jiantao Li<sup>1</sup>, Hua Deng<sup>1(✉)</sup>, and Wenjun Jiang<sup>2</sup>

<sup>1</sup> Central South University, Changsha, China  
ljt371@sina.com, hdeng@csu.edu.cn

<sup>2</sup> Hunan University, Changsha, China  
jiangwenjun@hnu.edu.cn

**Abstract.** Data analysis technique has been applied in many fields including smart industrial manufacture, which even leads to the industry 4.0 era. In this paper, we study how to control the vibration of flexible arms by exploiting analysis on operators' behaviors, so as to guarantee the safety of the instrument and users around. Vibration problem usually happens when starting and stopping the flexible arm. There are two common vibration problems. One is the unstable rotation, i.e., the so-called "one fast-one slow" effect. The other is the inaccurate stopping position, because of the vibration after stopping. In addition, current flexible arms are usually controlled/operated by human (i.e., the operator). The starting and stopping effects are highly depended on the expertise of operators. Our work in this paper takes a novel perspective from the operator, and we strive to search the best starting and stopping approaches to minimize the vibration. To be specific, we first analyze the possible states of a moving flexible arm, and theoretically determine the strategy for keeping safe states. Next, we empirically study which operations can lead to safe motion states, by conducting various real world tests and simulations. Finally, we summarize the findings and suggest safe operations. In the integrated process, we exploit data analysis technique and it shows significant effectiveness in solving industrial safety problem of flexible arm's vibration control.

**Keywords:** Smart industrial manufacture · Data analysis  
Safe vibration control · Flexible arms · Operators' behaviors

## 1 Introduction

Data technology has been applied in many fields and promotes their developments. Successful examples include personalized recommendation system [1], smart transportation [2], smart education [3], and smart industrial manufacture [4, 5]. Application of data technology in industry is an important feature of the industry 4.0 era [6]. In this paper, we study how to control the rotation of flexible arms by exploiting analysis on operators' behaviors, so as to guarantee the safety of the engineering mechanical equipment and around users.

Flexible arm is an important type of engineering mechanical arms, which is the critical and important parts of various engineering machinery and equipments [7]. Figure 1 shows some flexible arms in real life. Flexible arm is a kind of multi-functional efficient mechanical parts. The performance of an industrial machine is highly depended on the engineering arm. Working conditions of engineering mechanical are usually serious, and the load of engineering mechanical equipments is very complicated. Therefore, it is necessary to conduct movement detection and analysis, to ensure the safety and reliability of its work, and to improve construction quality and operation efficiency.

The control for flexible arms' rotation in engineering mechanical equipments is very challenging, because they often experience unwanted vibrations at the end points typically due to elastic deflections and system disturbances. Vibration problem further leads to reduced endpoint positioning accuracy, and negatively affects the overall control performance of the flexible arm [8].



**Fig. 1.** Flexible arms in real life: hydraulic aerial cage (left) and concrete pump truck (right).

Vibration problem of flexible arms have attracted many attentions. Several methods have been proposed (e.g., [9–11]), using different techniques or models. However, there is a lack of work studying the problem from the operators' perspective. This is exactly the motivation of our work in this paper. We try to understand the vibration problem from the operators' view, and select proper operations for safe vibration control. Our contributions are threefold: (1) We exploit data analysis technique to solve industrial safety problem of flexible arm's vibration control. (2) We analyze the possible states of a moving flexible arm, and theoretically determine the strategy for keeping safe states. (3) We empirically study which operations can lead to safe states, by conducting various real world tests and simulations on operators' behaviors. Based on this, we suggest safe operations to guide future research and manufacture.



## 2 Literature Review and Our Work

We briefly review related works and identify the connections and differences from our work. Next, we introduce the overview of our work in this paper. We also provide formal definitions of commonly used concepts in this paper.

### 2.1 Related Works

Data technology enhances the development of industrial manufactures, including deep learning [12], neural network [13], evolutive algorithms [], wireless sensor [14], RFID [15], cloud computing [16, 17], and so on. [18] proposes an architecture of industry 4.0-based manufacturing system.

[9] presented a model-independent control strategy of the flexible-joint manipulator (FJM). [10] proposed an intelligent proportional-integral (IPI) control strategy of a single link FJM. Most recently, [11] studied trajectory tracking in FJM system and developed a distributed higher-order differential feedback controller (DHODFC) using the link and joint position measurement, so as to reduce joint vibration in step input response and to improve tracking behavior in reference trajectory tracking control. [19] proposed an approach to obtain the global analytical modes (GAMs) and establish discrete dynamic model with low degree-of-freedom for a three-axis attitude stabilized spacecraft installed with a pair of solar arrays. [20] proposed a method with impulse spectrum as a more general way for vibration control of the robotic arm of flexible multiple links. [8] provided a review on vibration control of flexible arms, particularly focusing on two methods: the proportional-integral-derivative (PID) methods and the robust sliding mode controller (SMC). [21] proposed a method of predictive function control of the single-Link manipulator with flexible joint.

Although many work have been done, there is a lack of work from the users' view. In real life experience, the vibration problem is highly depended on operators' operations. Hence, we take a different perspective to study vibration problem from operators' behaviors.

### 2.2 Overview of Our Work

We take the system of a flexible arm as a black-box; in which the input signal (e.g., the rotation speed or the driving torque) is the system incentive and the output signal (e.g., the vibration, or rotation angle) is the system response. We strive to search the best starting and stopping approaches to minimize the vibration. To be specific, we first analyze the possible states of a moving flexible arm, and theoretically determine the strategy for keeping safe states. Next, we empirically study which operations can lead to safe motion states, by conducting various real world tests and simulations. Finally, we summarize the findings and suggest safe operations. In the integrated process, we exploit data analysis technique and it shows significant effectiveness in solving industrial safety problem of flexible arm's vibration control.

**Table 1.** The descriptions of concepts.

Concepts	Descriptions
Input signal	The driving signal that makes flexible arms start to rotation
Output signal	The response of flexible arms
Angular displacement	The angle in radians through which a point or line has been rotated in a specified sense about a specified axis
Angular velocity	The rate of change of its angular displacement regarding time
“One fast-one slow” effect	The unstable rotation during the starting process
Ramp signal	A signal whose graph is shaped like a ramp
Harmonic step signal	A signal whose graph is shaped with multiple steps
Starting/stopping time	The time taken for starting or stopping flexible arms
$x$ s starting/stopping	The starting or stopping time length is $x$ seconds

### 3 Features of Flexible Arms

In order to better study the vibration characteristics of flexible arms, we study the features of flexible arms in this section, which consist of three steps: (1) its natural frequency; (2) possible states during movement; and (3) its vibration problems (Table 1).

#### 3.1 Natural Frequency

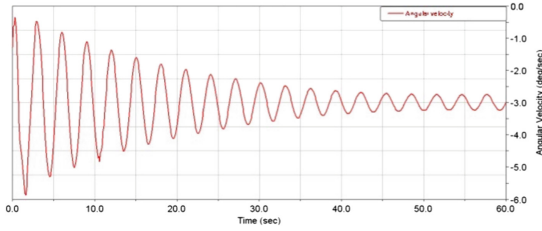
Suppose the start speed of flexible arm' rotation is  $-3$  d/s (3 degree per second, ‘-’ indicates the rotation direction is anticlockwise). The angular velocity of the end point of flexible arm is shown in Fig. 2. During the rotation, the angular velocity of the end point follows the law of single-freedom systems, as the following:

$$\omega(t) = Ae^{-\delta t} \cos(\omega_0(t) + \varphi_0) + B \quad (1)$$

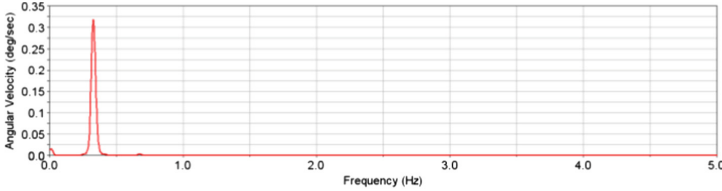
where  $A$  is the maximum angular velocity,  $B$  is the angular velocity when reaching balance,  $\omega_0$  is the natural frequency,  $\varphi_0$  is the phase angle,  $\delta$  is the damping coefficient.

After Fourier transition, the curve is converted into the figure in Fig. 3. It shows that the natural frequency of the flexible arm rotation is 0.3296 hz, the period is 3.034 s, and it rotates about 20 times in a minute.

During the rotation of the flexible arm, the changing angular velocity leads to its unstable vibration, particularly the “one slow-one fast” effect. It means the inconsistency in the speed of the flexible arms, i.e., some times the speed



**Fig. 2.** Angular velocity of the end point of flexible arms.

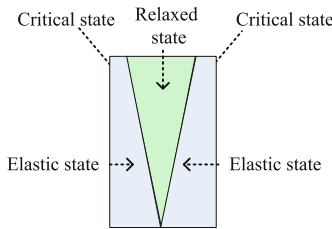


**Fig. 3.** Frequency of flexible arms.

is approaching 0 d/s, while some times the speed is approaching 6 d/s in Fig. 2. This effect further cause the large rotation distance and strong fluid pressure hit. Therefore, it is very necessary to study the rotation rule, and search for the proper starting time, and starting input signal, as well as the stopping time and stopping input signals.

### 3.2 Possible States During Rotation

During the rotation, flexible arms will have different states, namely, the relaxed state, the elastic state, and the critical state (i.e., the left or right boundary between relaxed state and elastic state), as shown in Fig. 4.



**Fig. 4.** States.

During the starting process (taking left starting for example), the flexible arm will change from relaxed state to right elastic state. After starting, i.e.,

when it reaches the maximum angular velocity, the flexible arm is in the elastic state. It will release elastic energy when rotating to the left, forming a combined motion of zuni left-right and rotation to the left, i.e., the crawling phenomenon.

During its stable working state, the flexible arm is in the critical state without elastic energy, and the combined motion.

During the stooping process (taking left stopping for example), the speed of the flexible arm will reduce. If it passes the critical state and transfer to the elastic state, the flexible arm will do free-damped vibration.

### 3.3 Vibration Problems

According to the rotation tiring test of flexible arms in a long period, long flexible arms are facing two vibration problems. One is the unstable rotation, i.e., the “one fast-one slow” effect, or the crawling phenomenon. The other is the inaccurate stop position, because of the vibration after stopping. In addition, current flexible arms are usually controlled/operated by human (i.e., the operator). The starting and stopping effects are highly depended on the expertise of operators.

Our work in this paper takes a novel perspective from the operator, and we strive to search the best starting and stopping approaches to minimize the vibration. We take the flexible arm as a black-box, and testing the output with different input signals. The inputs are predefined by the operators; while the outputs are measured by the angular distance and speed, using wireless sensors.

## 4 Solving the Starting Vibration

To solve the starting vibration of flexible arms, we mainly test two factors: the starting time length (i.e., the period from begin starting to stable working), and the proper starting signals. We test two different starting time lengths: 2s starting and 3s starting. For each starting time length, we test different input signals.

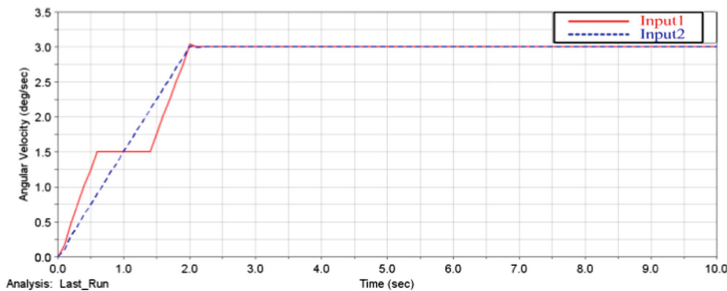


Fig. 5. Two input signals of 2s starting.

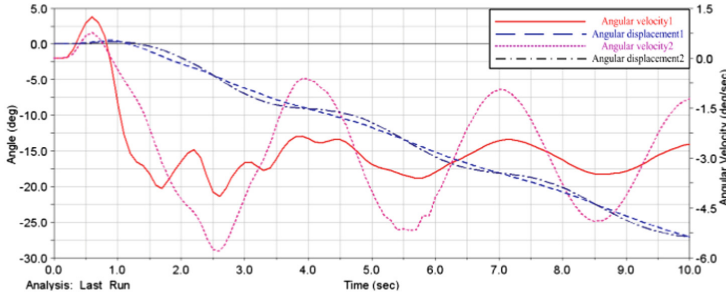


Fig. 6. Output angular displacement and velocity of 2s starting.

**2s Starting.** Firstly, we test two different input signals for 2s starting as shown in Fig. 5. In this figure, Input1 is the modulated signal which can be generated using two approaches, one is combining step signal with input shaping; the other is using the weighted harmonic signal. Meanwhile, Input2 is the ramp signal. 0–2s is the starting phase. After that, the flexible arm is in the stable working state. Figure 6 shows their corresponding angular velocity and angular displacement. Note that angular displacement is an integral of angular velocity. We can see that, the output angular displacement of Input1 is shorter than that of Input2, indicating that the flexible arm has a smaller vibration if started with the modulated signal.

**3s Starting.** Figure 7 shows three different input signals for 3s starting. In this figure, Input1 is a modified sine signal; Input2 is the modulated signal with harmonic signal and small step signal; Input 3 is a tangent signal. 0–3s is the starting phase. After that, the flexible arm is in the stable working state. Figure 8 shows their corresponding outputs. Compared to other two input signals, The output angular displacement of Input2 is the shortest, and the rotation velocity of the flexible arms is almost uniform with Input2. Again, it indicates that the modulated signal is a better choice for starting flexible arms. The fatigue life of the flexible arm is prolonged by reducing the stress variation.

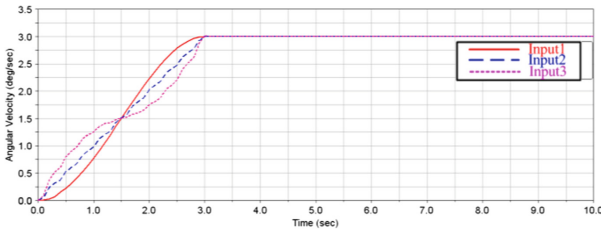


Fig. 7. Three input signals of 3s starting.

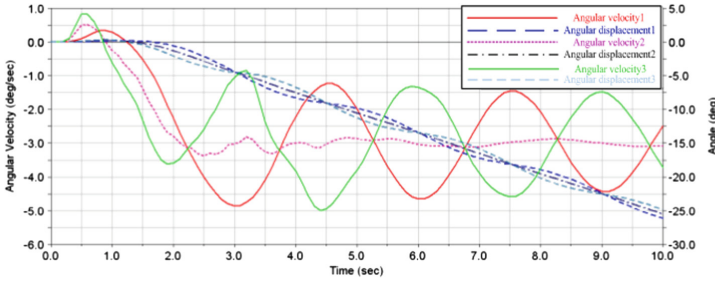


Fig. 8. Output angular displacement and velocity of 3s starting.

In addition, the output angular displacement of the harmonic step signal in 3s starting (Fig. 8) is shorter than that in 2s starting (Fig. 6). We also test the 4s starting, and the result shows that 3s starting is the best among 2s, 3s, and 4s starting. We try many times and the performance is stable. We analyze the reason and find that, this is because 3s is the most close to the flexible arm's frequency by nature.

## 5 Solving the Stopping Vibration

To solve the stopping vibration of flexible arms, we test three factors: the stopping time length (1s, 2s, 3s), different input signals, and the arm gesture before stopping.

### 5.1 Stopping Vibration Control

**1s Stopping.** Figure 9 shows three different input signals for 1s stopping. In this figure, Input1 is the ramp signal; Input2 is the modulated signal with harmonic signal and small step signal. 0–3s is the starting phase, 3–7s is the stable working phase, and 7–8s is the stopping phase. After that, there is no input signal for the flexible arms. Figure 10 shows their corresponding outputs. From the figure, we can see that the angular displacement of Input2 is shorter. The output swing range of Input1 is 15.5197–20.7403d, and the maximum swing is 5.2206d. Meanwhile, the output swing range of Input2 is 15.7808–20.3579d, and the maximum swing is 4.5771d, reduced by 12.3%. The reason for the limited improvement is that the stopping time is too short to generate enough microwave to eliminating the elastic energy.

**2s Stopping.** Figure 11 shows three different input signals for 2s stopping. In this figure, Input1 is the ramp signal; Input2 is the modulated signal with harmonic signal and small step signal; Input 3 is the modulated signal with harmonic signal and large step signal. 0–3s is the starting phase, 3–7s is the stable working phase, and 7–9s is the stopping phase. After that, there is no input signal for the flexible arms. Figure 12 shows their corresponding outputs. From the figure, we can see that the angular displacement of Input2 is the shortest.

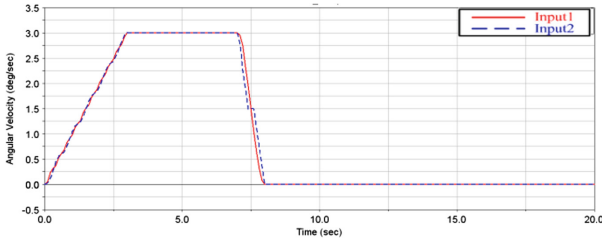


Fig. 9. Two input signals of 1s stopping.

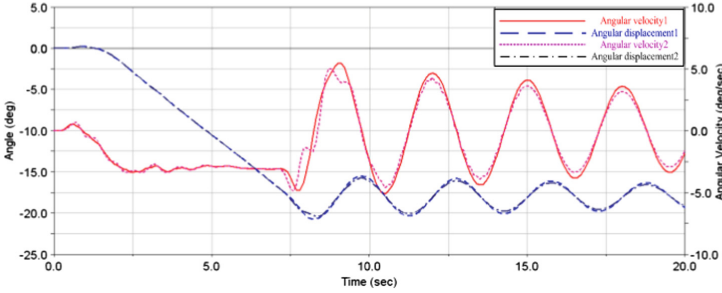


Fig. 10. Output angular displacement and velocity of 1s stopping.

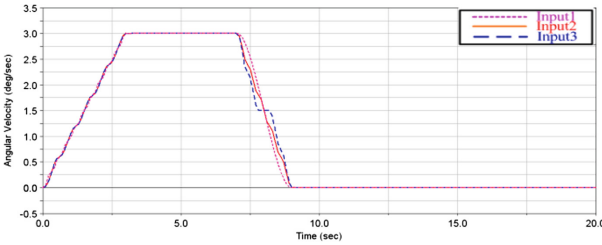


Fig. 11. Three input signals of 2s stopping.

**3s Stopping.** Figure 13 shows three different input signals for 3s stopping. In this figure, Input1 is the harmonic linear signal; Input2 is the harmonic step signal. 0–3s is the starting phase, 3–7s is the stable working phase, and 7–10s is the stopping phase. After that, there is no input signal for the flexible arms. Figure 14 shows their corresponding outputs. From the figure, we can see that the angular displacement of Input2 is shorter.

In addition, the output angular displacement of the harmonic step signal in 3s stopping (Fig. 14) is shorter than that in 1s and 2s stopping (Figs. 10 and 12). We try many times and the performance is stable. We analyze the reason and find that, this is because 3s is the most close to the flexible arm’s natural frequency. In addition, the stopping time with less than 3s is too short to eliminate the effect of elastic energy.

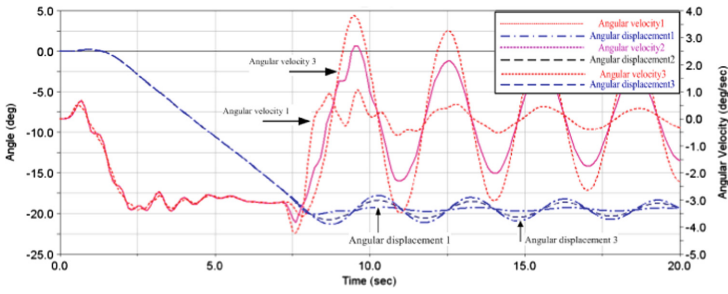


Fig. 12. Output angular displacement and velocity of 2s stopping.

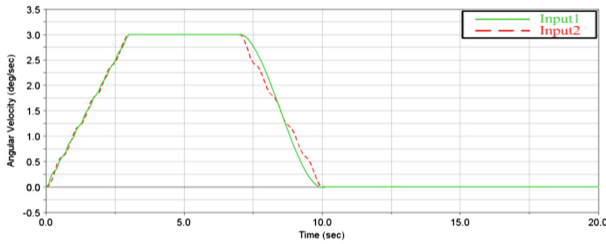


Fig. 13. Two input signals of 3s stopping.

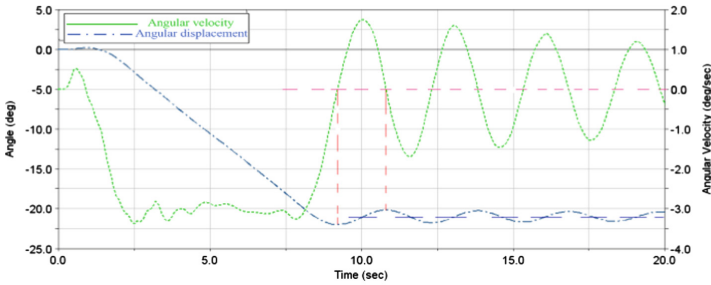


Fig. 14. Output angular velocity and displacement of 3s stopping.

## 6 Conclusion

In this paper, we study the safe rotation control of flexible arms by empirically studying operators' behaviors and corresponding effects. The novelty of our work is exploiting data technology to solve industrial safety problem. We analyze the feature of a moving flexible arm, and theoretically determine the strategy for keeping safe states. Then, we study which operations can lead to safe movement via intensive simulations and tests. According to the empirical studies in Sects. 4 and 5, we summarize the proper operations that lead to less vibration:

1. The preferred time length of starting and stopping is that close to the natural period ( $1/\text{natural-frequency}$ ) of flexible arms.



2. The preferred input signal is the harmonic modulated signal that is a harmonic signal modulated by small step signal.
3. The above two findings are suitable in any gesture (either non-horizontal or horizontal).

**Acknowledgments.** This work is supported by National Key R&D Program of China 2016YFF0203400, NSFC grants 61502161 and 61632009.

## References

1. Jiang, W., Wu, J., Wang, G., Zheng, H.: Forming opinions via trusted friends: time-evolving rating prediction using fluid dynamics. *IEEE Trans. Comput.* **65**, 1211–1224 (2015). <https://doi.org/10.1109/TC.2015.2444842>
2. Tedjasaputra, A., Sari, E.: Sharing economy in smart city transportation services. In: *Proceedings of the SEACHI 2016 on Smart Cities for Better Living with HCI and UX, SEACHI 2016*, pp. 32–35. ACM, New York (2016)
3. Kobayashi, T.: MSaaS-type smart education support system using social media. In: *2015 3rd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering*, pp. 119–127, March 2015
4. Xu, X., Hua, Q.: Industrial big data analysis in smart factory: current status and research strategies. *IEEE Access* **PP**(99), 1 (2017)
5. Park, C.Y., Laskey, K.B., Salim, S., Lee, J.Y.: Predictive situation awareness model for smart manufacturing. In: *2017 20th International Conference on Information Fusion (Fusion)*, pp. 1–8, July 2017
6. Khan, M., Wu, X., Xu, X., Dou, W.: Big data challenges and opportunities in the hype of industry 4.0. In: *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2017
7. Mardani, A., Ebrahimi, S.: Computational dynamic modeling and sequential PID controlling of a tendon-based manipulator with highly slender flexible arms. In: *2016 4th International Conference on Robotics and Mechatronics (ICROM)*, pp. 542–547, October 2016
8. Gadsden, S.A., AlShabi, M.: A comparison of vibration control strategies for a flexible-link robot arm. In: *2015 10th International Symposium on Mechatronics and its Applications (ISMA)*, pp. 1–5, December 2015
9. Chatlatanagulchai, W., Meckl, P.H.: Model-independent control of a flexible-joint robot manipulator. *ASME J. Dyn. Syst. Meas. Control* **131**(4), 041003:1–041003:9 (2009)
10. Agee, J.T., Kizir, S., Bingul, Z.: Intelligent proportional-integral (iPI) control of a single link flexible joint manipulator. *ASME J. Vib. Acoust.* **21**(11), 2273–2288 (2015)
11. Agee, J.T., Bingul, Z., Kizir, S.: Trajectory and vibration control of a single-link flexible-joint manipulator using a distributed higher-order differential feedback controller. *ASME J. Dyn. Syst. Meas. Control* **139**(5), 081006:1–081006:9 (2017)
12. Lei, Y., Jia, F., Zhou, X.: A deep learning-based method for machinery health monitoring with big data. *J. Mech. Eng.* **51**(21), 49–56 (2015)
13. Sun, W., Shao, S., Yan, R.: Induction motor fault diagnosis based on deep neural network of sparse auto-encoder. *J. Mech. Eng.* (2016)
14. Prieto, J., Mazuelas, S., Bahillo, A., Fernandez, P., Lorenzo, R.M., Abril, E.J.: Adaptive data fusion for wireless localization in harsh environments. *IEEE Trans. Signal Process.* **60**(4), 1585–1596 (2012)

15. Zhang, D., He, Z., Qian, Y., Wan, J., Li, D., Zhao, S.: Revisiting unknown RFID tag identification in large-scale internet of things. *IEEE Wirel. Commun.* **23**(5), 24–29 (2016)
16. Wang, S., Wan, J., Zhang, D., Li, D., Zhang, C.: Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. *Comput. Netw.* **101**, 158–168 (2016)
17. Shu, Z., Wan, J., Zhang, D., Li, D.: Cloud-integrated cyber-physical systems for complex industrial applications. *Mob. Netw. Appl.* **21**(5), 1–14 (2015)
18. Lee, J., Bagheri, B., Kao, H.A.: A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manuf. Lett.* **3**, 18–23 (2015)
19. Liu, L., Cao, D., Wei, J.: Rigid-flexible coupling dynamic modeling and vibration control for a three-axis stabilized spacecraft. *ASME J. Vib. Acoust.* **139**, 041006:1–041006:14 (2017)
20. Zhang, W.: Vibration control of multilink flexible robotic arm with impulse spectrum. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2622–2627, October 2016
21. Zhang, Z., Hu, C.: Predictive function control of the single-link manipulator with flexible joint. pp. 129–146 (2012)

# A New Color Image Encryption Scheme Based on Chaotic Hénon Map and Lü System

Chong Fu<sup>1(✉)</sup>, Gao-yuan Zhang<sup>1</sup>, Bei-li Gao<sup>2</sup>, Jing Sun<sup>1</sup>,  
and Xue Wang<sup>1</sup>

<sup>1</sup> School of Computer Science and Engineering, Northeastern University,  
Shenyang 110004, China

fuchong@mail.neu.edu.cn

<sup>2</sup> China Waterborne Transport Research Institute, Beijing 100088, China

**Abstract.** This paper presents a new efficient chaos-based color image cipher robust against chosen-plaintext attack. The chaotic Hénon map and Lü system are employed to produce the permutation and substitution keystream sequences for image data scrambling and mixing, respectively. In the permutation stage, the positions of colored subpixels in the input image are scrambled using a pixel-swapping mechanism. To strengthen the robustness of the substitution procedure against chosen-plaintext attack, we introduce a novel mechanism for associating the keystream sequence with the plain-image. Compared with other related mechanisms, the new mechanism is implemented during the subpixel values mixing process rather than the keystream generation process. As a result, the keystream sequence can be reused among different rounds of substitution operation, and hence the computational cost is reduced. The suggested mechanism also increases the diffusion intensity and a desired diffusion effect can be achieved with only two rounds of overall permutation-substitution operation. Experimental results demonstrate that the proposed scheme has a satisfactory level of security.

**Keywords:** Image cipher · Permutation-substitution · Hénon map  
Lü system · Plaintext-dependent keystream sequence

## 1 Introduction

Over the past decade, great concerns have been raised about the security of images transmitted over the Internet. A direct and obvious way to protect image data from unauthorized eavesdropping is to employ an encryption algorithm. Unfortunately, commonly used block ciphers, including DES and AES, characterized by high computational complexity are difficult to meet the increasing demand for real-time communications when dealing with digital images characterized by bulk data capacity. To meet this challenge, many different encryption technologies have been proposed. Among them, the chaos-based technology has suggested a promising direction. In 1998, Fridrich suggested an iterative permutation-substitution model for construction of secure image ciphers [1]. In each round of the cipher, the pixel positions are firstly scrambled in a pseudorandom manner, which leads to a great reduction in the

correlation among neighboring pixels. Then, the pixel values are altered sequentially and the influence of each pixel is diffused to all its succeeding ones during the modification process. The whole permutation-substitution operation is often iterated for several rounds to ensure the influence of each individual pixel can be spread over the entire cipher-image.

Conventionally, three area-preserving invertible chaotic maps, i.e., the cat map, the baker map, and the standard map, are widely used for image scrambling. Unfortunately, this kind of permutation strategy suffers from two main disadvantages, i.e., the periodicity of discretized version of chaotic maps and only applicable to square images [2–4]. To address these two drawbacks, Fu et al. [5] suggested an image scrambling scheme using a chaotic sequence sorting mechanism. In [6], inspired by the natural ripple-like phenomenon that distorts a reflection on a water surface, Wu et al. suggested a novel scrambling algorithm that shuffle images in an  $n$  dimensional ( $n D$ ) space using wave perturbations.

Recently, it has been reported that many existing image encryption schemes have been successfully broken by using known/chosen-plaintext attacks [7–10]. This is due to the fact that the substitution keystream sequences used in these schemes is solely determined by the secret key. That is, the same keystream sequence is used to encrypt different plain-images unless a different secret key is used. Consequently, the keystream sequence may be determined by encrypting some specially created images (e.g. an image with all pixels having the same value) and then comparing them with their corresponding outputs. Obviously, if a keystream sequence depends on both the secret key and the plaintext, then such analysis may become impractical. For instance, in [11], the keystream elements are extracted from multiple times iteration of the logistic map, and the iteration times are determined by plain-pixel values. In [12, 13], the authors introduced a mechanism that dynamically alters the value of control parameter of the chaotic map under the control of plain-pixel values. Unfortunately, the above mechanisms are implemented during the keystream generation process. This means a new keystream sequence should be produced for each round of substitution operation, thereby increasing the computational cost. To address this problem, this paper suggests a new mechanism that dynamically alters the values of the keystream elements under the control of plain-subpixel values during the subpixel values mixing process. As the keystream generation process has no dependency on the plain-image, the keystream sequence can be reused among different rounds of substitution operation. The suggested mechanism also increases the diffusion intensity and a desired diffusion effect can be obtained with only two rounds of overall permutation-substitution operation. In the permutation stage, the positions of subpixel in each color channel of the plain-image are scrambled across the entire color space using a pixel-swapping strategy under the control of a keystream sequence generated from the Hénon map. Experimental results demonstrate that the proposed scheme has a satisfactory level of security.

The remainder of this paper is organized as follows. The proposed color image encryption algorithm is described in detail in Sect. 2. In Sect. 3, the diffusion performance of the proposed cryptosystem is analyzed. In Sect. 4, we analyze the security of the proposed cryptosystem through various statistical analyses, key space analysis and key sensitivity analysis. Finally, conclusions are drawn in the last section.

## 2 The Image Encryption Scheme

As aforementioned, the chaotic Hénon map and Lü system are employed in our scheme to generate the permutation and substitution keystreams for image data scrambling and masking. A brief introduction of the two models is given below.

The Hénon map [14], introduced by Michel Hénon as a simplified model of the Poincaré section of the Lorenz model, is one of the most studied examples of dynamical systems that exhibit chaotic behavior. The Hénon map takes a point  $(x_n, y_n)$  in the plane and maps it to a new point, as described by

$$\begin{cases} x_{n+1} = y_n + 1 - ax_n^2, \\ y_{n+1} = bx_n, \end{cases} \tag{1}$$

where  $a$  and  $b$  are two parameters. The map is chaotic with classical parameter values  $a = 1.4$  and  $b = 0.3$ . Evidently, the initial conditions  $(x_0, y_0)$  of the map are the immediate candidate for the secret key for permutation, as they uniquely determine a chaotic trajectory from which the permutation keystream is extracted.

Mathematically, the Lü system [15] is described by

$$\begin{cases} \dot{x} = a(y - x), \\ \dot{y} = -xz + cy, \\ \dot{z} = xy - bz, \end{cases} \tag{2}$$

where  $a, b$  and  $c$  are real parameters. Numerical experience shows that the system exhibits chaotic behavior when  $a = 36, b = 3$  and  $c \in (12.7, 17.0) \cup (18.0, 22.0) \cup (23.0, 28.5) \cup (28.6, 29.0) \cup (29.334, 29.345)$ . Similarly, the initial conditions  $(x_0, y_0, z_0)$  of the system are used as the secret key for substitution.

Without loss of generality, a 24-bit RGB color image of size  $W \times H$  is used as an input. The detailed encryption process is described as follows:

**Step 1:** Load the subpixel data of the input image to a 2-D byte matrix

$$imgData = \begin{bmatrix} p_0 & p_1 & \cdots & p_{3 \times W - 1} \\ p_{3 \times W} & p_{3 \times W + 1} & \cdots & p_{3 \times W \times 2 - 1} \\ \cdots & \cdots & \cdots & \cdots \\ p_{3 \times W \times (H-1)} & p_{3 \times W \times (H-1) + 1} & \cdots & p_{3 \times W \times H - 1} \end{bmatrix}$$

**Step 2:** Generate a chaotic matrix  $chaoMat$  of the same size as that of  $imgData$  with each element consisting of two floating-point numbers obtained by iterating map (1).

**Step 2.1:** Pre-iterate map (1) for  $T_0$  times to avoid the harmful effect of transitional procedure, where  $T_0$  is a constant.

**Step 2.2:** Continue the iteration for  $3 \times W \times H$  times. For each iteration, the current values of the two state variables  $x$  and  $y$  are stored into a matrix  $chaoMat =$

$$\begin{bmatrix} (cm_{x(0)}, cm_{y(0)}) & (cm_{x(1)}, cm_{y(1)}) & \cdots & (cm_{x(3 \times W - 1)}, cm_{y(3 \times W - 1)}) \\ (cm_{x(3 \times W)}, cm_{y(3 \times W)}) & (cm_{x(3 \times W + 1)}, cm_{y(3 \times W + 1)}) & \cdots & (cm_{x(3 \times W \times 2 - 1)}, cm_{y(3 \times W \times 2 - 1)}) \\ \cdots & \cdots & \cdots & \cdots \\ (cm_{x[3 \times W \times (H-1)]}, cm_{y[3 \times W \times (H-1)]}) & (cm_{x[3 \times W \times (H-1) + 1]}, cm_{y[3 \times W \times (H-1) + 1]}) & \cdots & (cm_{x(3 \times W \times H - 1)}, cm_{y(3 \times W \times H - 1)}) \end{bmatrix}$$

as an element in the order from left to right, top to bottom.

**Step 3:** Extract a permutation matrix  $permMat =$

$$\begin{bmatrix} (pm_{x(0)}, pm_{y(0)}) & (pm_{x(1)}, pm_{y(1)}) & \cdots & (pm_{x(3 \times W - 1)}, pm_{y(3 \times W - 1)}) \\ (pm_{x(3 \times W)}, pm_{y(3 \times W)}) & (pm_{x(3 \times W + 1)}, pm_{y(3 \times W + 1)}) & \cdots & (pm_{x(3 \times W \times 2 - 1)}, pm_{y(3 \times W \times 2 - 1)}) \\ \cdots & \cdots & \cdots & \cdots \\ (pm_{x[3 \times W \times (H-1)]}, pm_{y[3 \times W \times (H-1)]}) & (pm_{x[3 \times W \times (H-1) + 1]}, pm_{y[3 \times W \times (H-1) + 1]}) & \cdots & (pm_{x(3 \times W \times H - 1)}, pm_{y(3 \times W \times H - 1)}) \end{bmatrix}$$

from  $chaoMat$  according to

$$\begin{cases} pm_{x(n)} = \text{mod}(\text{sig}(\text{abs}(cm_{x(n)}), \alpha), H), \\ pm_{y(n)} = \text{mod}(\text{sig}(\text{abs}(cm_{y(n)}), \alpha), 3 \times W), \end{cases} \quad (3)$$

where  $\text{abs}(x)$  returns the absolute value of  $x$ ,  $\text{sig}(x, \alpha)$  returns the  $\alpha$  most significant decimal digits of  $x$ , and  $\text{mod}(x, y)$  divides  $x$  by  $y$  and returns the remainder of the division. An  $\alpha$  value of 15 is recommended as all the state variables in our scheme are declared as double-precision type, which has 15 or 16 decimal places of accuracy.

**Step 4:** Generate a chaotic sequence of length  $L_{cs} = 3 \times W \times H$  by iterating system (2).

**Step 4.1:** Pre-iterate system (2) for  $T_0$  times for the same purpose mentioned above. The system can be numerically solved by using fourth-order Runge-Kutta method, as given by

$$\begin{cases} x_{n+1} = x_n + (h/6)(K_1 + 2K_2 + 2K_3 + K_4), \\ y_{n+1} = y_n + (h/6)(L_1 + 2L_2 + 2L_3 + L_4), \\ z_{n+1} = z_n + (h/6)(M_1 + 2M_2 + 2M_3 + M_4), \end{cases} \quad (4)$$

where

$$\begin{cases} K_j = a(y_n - x_n) \\ L_j = -x_n z_n + c y_n \\ M_j = x_n y_n - b z_n, \end{cases} \quad (\text{with } j = 1)$$

$$\begin{cases} K_j = a[(y_n + hL_{j-1}/2) - (x_n + hK_{j-1}/2)] \\ L_j = -(x_n + hK_{j-1}/2)(z_n + hM_{j-1}/2) + c(y_n + hL_{j-1}/2) \\ M_j = (x_n + hK_{j-1}/2)(y_n + hL_{j-1}/2) - b(z_n + hM_{j-1}/2), \end{cases} \quad (\text{with } j = 2, 3)$$

$$\begin{cases} K_j = a[(y_n + hL_{j-1}) - (x_n + hK_{j-1})] \\ L_j = -(x_n + hK_{j-1})(z_n + hM_{j-1}) + c(y_n + hL_{j-1}) \\ M_j = (x_n + hK_{j-1})(y_n + hL_{j-1}) - b(z_n + hM_{j-1}), \end{cases} \quad (\text{with } j = 4)$$

and the step  $h$  is chosen as 0.005.

**Step 4.2:** Continue the iteration for  $W \times H$  times. For each iteration, the current values of the three state variables  $x$ ,  $y$  and  $z$  are in turn stored into an array  $subSeq = \{ss_0, ss_1, \dots, ss_{3 \times W \times H-1}\}$ .

**Step 5:** Extract a substitution keystream  $subKstr = \{sk_0, sk_1, \dots, sk_{3 \times W \times H-1}\}$  from  $subSeq$  according to

$$sk_n = mod[sig((abs(ss_n), \alpha), G_L), \tag{5}$$

where  $G_L$  is the number of gray levels in the input image (for a 24-bit RGB image,  $G_L = 256$ ).

**Step 6:** Encipher the subpixel data of the input image, i.e. the matrix  $imgData$ , using iterative permutation-substitution operations.

**Step 6.1:** Scramble the subpixels in  $imgData$  according to the permutation matrix  $permMat$ , or more specifically, swap each subpixel  $p_n$  in  $imgData$  with another one located at  $(pm_{x(n)}, pm_{y(n)})$  in the order from left to right, top to bottom.

**Step 6.2:** Mask the values of each scrambled subpixel in  $imgData$  in the order from left to right, top to bottom.

**Step 6.2.1:** Bit-wise rotate the currently applied keystream element  $sk_n$  to the right under the control of the value of the previously operated subpixel  $p_{n-1}$ , as described by Eq. (6).

$$\begin{aligned} sk_{n(new)} &= [sk_n \ll (\log_2 G_L - mod(p_{n-1}, \log_2 G_L))] \\ & (sk_n \gg mod(p_{n-1}, \log_2 G_L)), \end{aligned} \tag{6}$$

where “ $\ll s$ ” and “ $\gg s$ ” denote a left and right shift by  $s$  bit, respectively, and “ $\oplus$ ” denotes a bit-wise OR operation. For the first subpixel  $p_0$ , the initial value  $p_{-1}$  can be set as a constant.

**Step 6.2.2:** Calculate the cipher-subpixel values according to Eq. (7).

$$c_n = sk_{n(new)} \oplus [mod((p_n + sk_{n(new)}), G_L)] \oplus c_{n-1}, \tag{7}$$

where  $p_n$  is the currently operated subpixel,  $c_n$  and  $c_{n-1}$  are the output and previous ciphered subpixels, respectively, and  $\oplus$  performs bit-wise exclusive OR operation. Similarly, one may set the initial value  $c_{-1}$  as a constant.

**Step 6.2.3:** return to **Step 6.2.1** until the values of all the subpixels in  $imgData$  are mixed.

**Step 7:** Repeat **Step 6** until the influence of each individual subpixel is spread out over the entire cipher-image.

The decryption procedure is similar to that of the encryption process except that some steps are followed in a reversed order. Particularly, the inverse of Eq. (7) is given by

$$p_n = \text{mod}[(sk_{n(\text{new})} \oplus c_n \oplus c_{n-1} + G_L - sk_{n(\text{new})}), G_L]. \quad (8)$$

As can be seen from the above description of the proposed encryption algorithm, we associate the keystream sequence with the plain-image by bitwise rotating each keystream element before it is applied to a subpixel, and the number of bits to be rotated is determined by the original value of the previously operated subpixel. As the chaotic sequence generation procedure (*Step 4*) and the subsequent keystream sequence quantification procedure (*Step 5*) have no dependencies on the plain-image, the keystream sequence can be reused among different rounds of substitution operation. Moreover, as can be seen from Eq. (7), a subpixel is mixed with the keystream element as well as previous ciphered subpixel, where the latter contains the accumulated effect of all its previous subpixels values. In our scheme, by associating the keystream sequence with the plain-image, the influence of each individual subpixel also acts on keystream elements. As a result, the diffusion intensity is increased and a desired *UACI* performance can be achieved with fewer rounds of overall permutation-substitution operation.

### 3 Analysis of the Diffusion Performance of the Proposed Scheme

As aforementioned, the substitution procedure serves to spread the influence of individual plaintext bits over as much of the ciphertext as possible. This is of much importance because otherwise the cryptosystem will be vulnerable to chosen-plaintext attack. The differential analysis is the most common way to implement the chosen-plaintext attack. To do this, an opponent may firstly create two plain-images with only one-bit difference, and then encrypt the two images using the same secret key. By observing the differences between the two resulting cipher-images, some meaningful relationship between plain-image and cipher-image could be found out, and it further facilitates in determining the keystream. Obviously, this kind of cryptanalysis may become impractical if a cryptosystem is highly sensitive to plaintext, i.e. changing one bit of the plaintext affect every bit in the ciphertext.

To measure the diffusion property of an image cryptosystem, two criteria, i.e., *NPCR* (the number of pixel change rate) and *UACI* (the unified average changing intensity) are commonly used. The *NPCR* is used to measure the percentage of different pixel numbers between two images. Let  $I_1(i, j, k)$  and  $I_2(i, j, k)$  be the  $(i, j)$ th pixel in  $k$ th color channel ( $k = 1, 2, 3$  denotes the red, green, and blue color channels, respectively) of two images  $I_1$  and  $I_2$ , the *NPCR* can be defined as:



$$NPCR = \frac{\sum_{k=1}^3 \sum_{i=1}^H \sum_{j=1}^W D(i, j, k)}{3 \times H \times W} \times 100\%, \tag{9}$$

where  $D(i, j, k)$  is defined as

$$D(i, j, k) = \begin{cases} 0 & \text{if } I_1(i, j, k) = I_2(i, j, k), \\ 1 & \text{if } I_1(i, j, k) \neq I_2(i, j, k). \end{cases} \tag{10}$$

The second criterion,  $UACI$  is used to measure the average intensity of differences between the two images. It is defined as

$$UACI = \frac{1}{3 \times H \times W} \left[ \sum_{k=1}^3 \sum_{i=1}^H \sum_{j=1}^W \frac{|I_1(i, j, k) - I_2(i, j, k)|}{G_L - 1} \right] \times 100\%. \tag{11}$$

Clearly, no matter how similar the two input images are, a good image cryptosystem should procedure outputs with  $NPCR$  and  $UACI$  values ideally being equal to that of two random images, which are given by

$$NPCR_{expected} = \left( 1 - \frac{1}{2^{\log_2 G_L}} \right) \times 100\% \tag{12}$$

and

$$UACI_{expected} = \frac{1}{G_L^2} \left( \frac{\sum_{i=1}^{G_L-1} i(i+1)}{G_L - 1} \right) \times 100\%. \tag{13}$$

For instance, the  $NPCR$  and  $UACI$  values for two random color images in 24-bit RGB format are 99.609% and 33.464%, respectively.

The  $NPCR$  and  $UACI$  of the proposed cryptosystem are evaluated using five standard 24-bit color test images of size  $512 \times 512$  taken from the USC-SIPI image database. The differential images are created by randomly changing 1-bit in the original ones, as listed in Table 1. The two images in each test pair are encrypted using the same secret key, and their  $NPCR$  and  $UACI$  values under different number of cipher rounds are calculated and compared with that of the conventional scheme, as listed in Tables 2 and 3, respectively. As can be seen from Tables 2 and 3, though both the proposed and the conventional schemes take two rounds to obtain a desired  $NPCR$  value, one more round is needed by the conventional scheme to obtain a desired  $UACI$  value. Therefore, the proposed substitution strategy provides superior computational efficiency.

## 4 Security Analysis

In this section, thorough security analysis has been carried out, including the most important ones like brute-force analysis, statistical analysis and key sensitivity analysis, to demonstrate the high security of the proposed scheme.

**Table 1.** Differential images used in *NPCR* and *UACI* tests

Test image name	Color component	Pixel position (x, y)	Pixel value	
			Original	Modified
F16	R	(160, 271)	193	192
House	B	(43, 118)	130	129
Montreal	R	(309, 135)	3	4
Peppers	B	(249, 410)	131	132
Sailboat	G	(469, 406)	187	186

**Table 2.** Results of *NPCR* test

Test image name	No. of cipher rounds (proposed scheme)			No. of cipher rounds (conventional scheme)		
	1	2	3	1	2	3
F16	0.32416	<b>0.99605</b>	0.99605	0.32416	<b>0.99634</b>	0.99609
House	0.61242	<b>0.99600</b>	0.99608	0.61242	<b>0.99622</b>	0.99613
Montreal	0.62026	<b>0.99612</b>	0.99600	0.62026	<b>0.99620</b>	<b>0.99613</b>
Peppers	0.64716	<b>0.99610</b>	0.99616	0.64716	<b>0.99598</b>	<b>0.99606</b>
Sailboat	0.34846	<b>0.99603</b>	0.99609	0.34846	<b>0.99625</b>	<b>0.99607</b>

**Table 3.** Results of *UACI* test

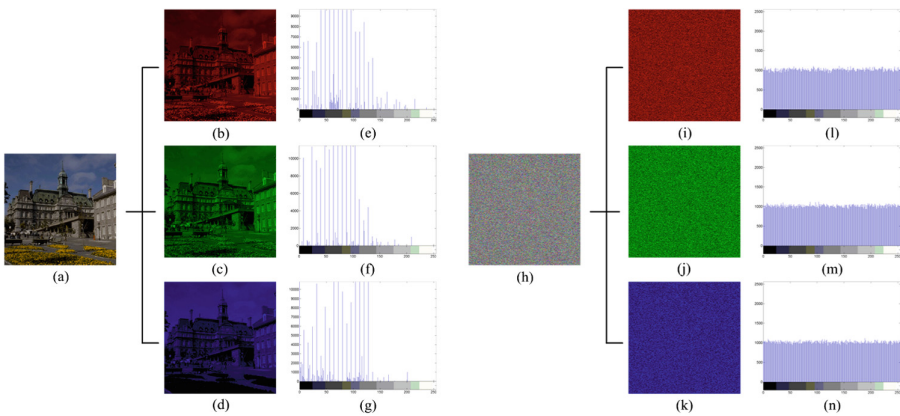
Test image name	No. of cipher rounds (proposed scheme)			No. of cipher rounds (conventional scheme)		
	1	2	3	1	2	3
F16	0.04068	<b>0.33441</b>	0.33477	0.00508	0.33612	<b>0.33461</b>
House	0.30731	<b>0.33446</b>	0.33493	0.00480	0.33718	<b>0.33447</b>
Montreal	0.07786	<b>0.33477</b>	0.33397	0.00973	0.33615	<b>0.33472</b>
Peppers	0.32461	<b>0.33440</b>	0.33444	0.00253	0.33270	<b>0.33420</b>
Sailboat	0.17489	<b>0.33456</b>	0.33492	0.08755	0.33404	<b>0.33467</b>

### 4.1 Brute-Force Analysis

In cryptography, a brute-force attack is a cryptanalytic attack that attempts to break a cipher by systematically checking all possible keys until the correct one is found. A key should therefore be long enough that this line of attack is impractical – i.e., would take too long to execute. As mentioned above, the initial states of the Hénon map,  $(x_0, y_0)$ , and the Lü system,  $(x_0, y_0, z_0)$ , are used as the permutation and substitution keys, respectively. As aforementioned, all the state variables in our algorithm are declared as 64-bit double-precision type, which gives 53 bits of precision. The two keys are independent of each other, and therefore the key length of the proposed scheme is  $53 \times 5 = 265$  bits. Generally, cryptographic algorithms use keys with a length greater than 100 bits are considered to be “computational security” as the number of operations required to try all possible  $2^{100}$  keys is widely considered out of reach for conventional digital computing techniques for the foreseeable future. Therefore, the proposed scheme is secure against brute-force attack.

### 4.2 Statistical Analysis

**Frequency distribution of pixel values.** A good image cryptosystem should flatten the frequency distribution of cipher-pixels values as such information has the potential to be exploited in a statistical attack. The frequency distribution of pixel values in an image can be easily explored by using histogram analysis. An image histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. The histograms of the three color channels of the “Montreal” test image and its output cipher-image produced by the proposed scheme are shown in Fig. 1. It’s clear from Figs. 1(l)–(n) that the pixel values of all the three color components of the output cipher-image are fairly evenly distributed over the whole intensity range, and therefore no information about the plain-image can be gathered through histogram analysis.



**Fig. 1.** Histogram analysis. (a) and (h) are the test image and its output cipher-image, respectively. (b)–(d) and (i)–(k) are the three color channels of (a) and (h), respectively. (e)–(g) and (l)–(n) are the histograms of (b)–(d) and (i)–(k), respectively.

The distribution of pixel values can be further quantitatively determined by calculating the information entropy of the image. Information entropy, introduced by Claude E. Shannon in his classic paper “A Mathematical Theory of Communication”, is a key measure of the randomness or unpredictability of information content. The information entropy is usually expressed by the average number of bits needed to store or communicate one symbol in a message, as described by

$$H(S) = - \sum_{i=1}^N P(s_i) \log_2 P(s_i), \quad (14)$$

where  $S$  is a random variable with  $N$  outcomes  $\{s_1, \dots, s_N\}$  and  $P(s_i)$  is the probability mass function of outcome  $s_i$ . It's obvious from Eq. (14) that the entropy for a random source emitting  $N$  symbols is  $\log_2 N$ . For instance, for a ciphered image with 256 color levels per channel, the entropy should ideally be 8, otherwise there exists certain degree of predictability which threatens its security.

The information entropies of above five test images and their output cipher-images are calculated, and the results are listed in Table 4. As can be seen from Table 4, the entropy of all the output cipher-images are very close to the theoretical value of 8. This means the proposed scheme produces outputs with perfect randomness and hence is robust against frequency analysis.

**Table 4.** Information entropies of the test images and their output cipher-images.

Test image name	Plain-image	Cipher-image
F16	6.663908	7.999784
House	7.485787	7.999748
Montreal	4.826442	7.999748
Peppers	7.669826	7.999757
Sailboat	7.762170	7.999741

**Correlation between neighboring pixels.** The simplest way to investigate the relationship between two neighboring pixels in an image is to use a scatter plot, which is carried out as follows. First, randomly select  $S_n$  pairs of neighboring pixels in each direction from the image. Then, the selected pairs is displayed as a collection of points, each having the value of one pixel determining the position on the horizontal axis and the value of the other pixel determining the position on the vertical axis.

Figures 2(a)–(c) and (d)–(f) show the scatter diagrams for horizontally, vertically and diagonally neighboring pixels in the red channel of the “Montreal” test image and its output cipher-image with  $S_n = 5000$ , respectively. Similar results can be obtained for the other two color channels. As can be seen from Fig. 2, most points in (a)–(c) are clustered around the main diagonal, whereas those in (d)–(f) are fairly evenly distributed. The results indicate that the proposed scheme can effectively eliminate the correlation between neighboring pixels in an input image.

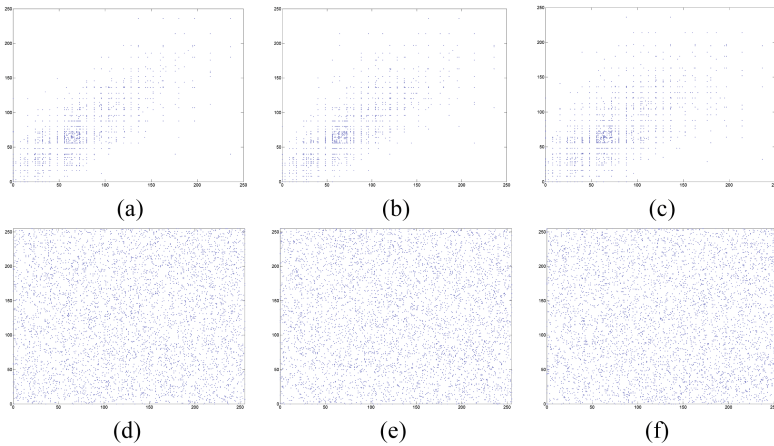
To further quantitatively measure the correlation between neighboring pixels in an image, the correlation coefficients  $r_{xy}$  for the sampled pairs are calculated according to the following three formulas:

$$r_{xy} = \frac{\frac{1}{S_n} \sum_{i=1}^{S_n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\frac{1}{S_n} \sum_{i=1}^{S_n} (x_i - \bar{x})^2\right) \left(\frac{1}{S_n} \sum_{i=1}^{S_n} (y_i - \bar{y})^2\right)}}, \tag{15}$$

$$\bar{x} = \frac{1}{S_n} \sum_{i=1}^{S_n} x_i, \tag{16}$$

$$\bar{y} = \frac{1}{S_n} \sum_{i=1}^{S_n} y_i, \tag{17}$$

where  $x_i$  and  $y_i$  form the  $i$ th pair of neighboring pixels.



**Fig. 2.** Graphical analysis for correlation of neighboring pixels. (a)–(c) and (d)–(f) are scatter diagrams for horizontally, vertically and diagonally neighboring pixels in the red channel of the “Montreal” test image and its output cipher-image, respectively. (Color figure online)

Table 5 lists the calculated correlation coefficients for neighboring pixels in the three color channels of the five test images and their output cipher-images. As can be seen from Table 5, the correlation coefficients for neighboring pixels in all the three color channels of the output cipher-images are very close to zero, and it further supports the conclusion drawn from Fig. 2.

**Table 5.** Correlation coefficients for neighboring pixels in the test images and their output cipher-images.

Test image name	Direction	Plain-image			Cipher-image		
		R	G	B	R	G	B
F16	Horizontal	0.9516	0.9703	0.9290	-0.0054	0.0141	0.0045
	Vertical	0.9735	0.9565	0.9636	-0.0031	-0.0156	-0.0273
	Diagonal	0.9271	0.9350	0.9109	0.0176	-0.0029	-0.0032
House	Horizontal	0.9575	0.9463	0.9691	0.0186	-0.0101	-0.0144
	Vertical	0.9513	0.9368	0.9735	-0.0216	0.0149	-0.0028
	Diagonal	0.9196	0.8933	0.9468	-0.0196	0.0045	0.0321
Montreal	Horizontal	0.8927	0.8862	0.9517	0.0263	0.0086	-0.0004
	Vertical	0.8870	0.8808	0.9418	-0.0204	-0.0016	0.0040
	Diagonal	0.7994	0.8024	0.9117	0.0130	0.0056	0.0045
Peppers	Horizontal	0.9691	0.9755	0.9674	-0.0064	-0.0027	-0.0247
	Vertical	0.9656	0.9824	0.9666	-0.0091	0.0072	-0.0116
	Diagonal	0.9583	0.9644	0.9465	-0.0116	-0.0333	-0.0074
Sailboat	Horizontal	0.9537	0.9677	0.9698	0.0034	0.0006	-0.0227
	Vertical	0.9558	0.9696	0.9704	0.0065	0.0056	-0.0125
	Diagonal	0.9422	0.9530	0.9515	0.0030	-0.0063	-0.0214

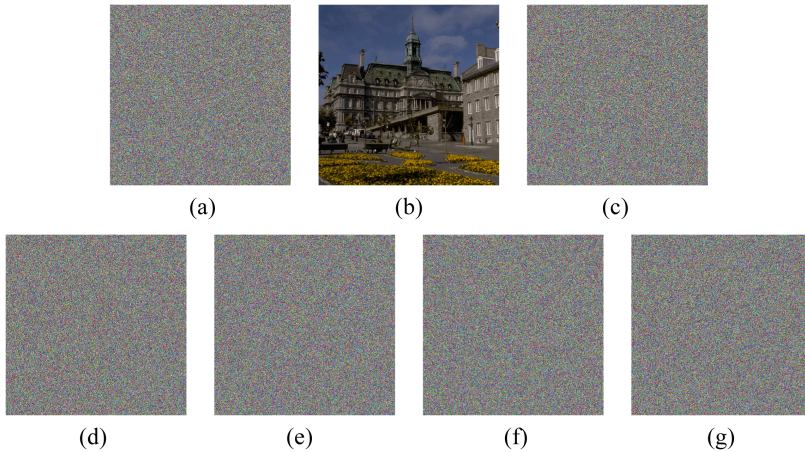
**Table 6.** Decryption keys used for key sensitivity test

Figure	Decryption key	
	Permutation part	Substitution part
3(b)	$x_0 = 0.825201589901473$ $y_0 = -0.206378406313441$	$x_0 = -5.11799279781735, y_0 = 7.69311459457213$ $z_0 = 25.4568382928529$
3(c)	$x_0 = \underline{0.825201589901472}$ $y_0 = -0.206378406313441$	$x_0 = -5.11799279781735, y_0 = 7.69311459457213$ $z_0 = 25.4568382928529$
3(d)	$x_0 = 0.825201589901473$ $y_0 = \underline{-0.206378406313442}$	$x_0 = -5.11799279781735, y_0 = 7.69311459457213$ $z_0 = 25.4568382928529$
3(e)	$x_0 = 0.825201589901473$ $y_0 = -0.206378406313441$	$x_0 = \underline{-5.11799279781736}, y_0 = 7.69311459457213$ $z_0 = 25.4568382928529$
3(f)	$x_0 = 0.825201589901473$ $y_0 = -0.206378406313441$	$x_0 = -5.11799279781735, y_0 = \underline{7.69311459457212}$ $z_0 = 25.4568382928529$
3(g)	$x_0 = 0.825201589901473$ $y_0 = -0.206378406313441$	$x_0 = -5.11799279781735, y_0 = 7.69311459457213$ $z_0 = \underline{25.4568382928528}$

### 4.3 Key Sensitivity Analysis

To evaluate the key sensitivity property of the proposed scheme, the ‘‘Montreal’’ test image is firstly encrypted using a randomly generated secret key: H6non map with initial conditions ( $x_0 = 0.825201589901473, y_0 = -0.206378406313441$ ) and L6 system with initial condition ( $x_0 = -5.11799279781735, y_0 = 7.69311459457213,$

$z_0 = 25.45683\ 82928529$ ), and the resulting cipher-image is shown in Fig. 3(a). Then the cipher image is tried to be decrypted using six decryption keys, one of which is exactly the same as the encryption key and the other five have only one-bit difference to it, as listed in Table 6. The resulting deciphered images are shown in Figs. 3(b–g), respectively, from which we can see that even an almost perfect guess of the key does not reveal any information about the original image. It can, therefore, be concluded that the proposed scheme fully satisfies the key sensitivity requirement.



**Fig. 3.** Results of key sensitivity test

## 5 Conclusions

This paper has proposed a new permutation-substitution type color image cipher based on chaotic Hénon Map and Lü System. To confuse the relationship between the ciphertext and the secret key, the positions of colored subpixels in the input image are scrambled using a pixel-swapping mechanism, which avoids the two main problems encountered when using the discretized version of area-preserving chaotic maps. To strengthen the robustness of the substitution procedure against chosen-plaintext attack, we introduced a new mechanism for associating the keystream with the plain-image by dynamically altering the values of the keystream elements under the control of plain-subpixels values during the subpixel values mixing process. Compared with other related mechanisms, the proposed mechanism allows the keystream sequence to be reused among different rounds of substitution operation as keystream generation process has no dependency on the plain-image. The proposed mechanism also helps increase the diffusion intensity and the experimental results indicate that the proposed scheme takes only two cipher rounds to achieve both desired *NPCR* and *UACI* values. We have carried out an extensive security analysis, which demonstrates the satisfactory security level of the new scheme. It can therefore be concluded that the proposed scheme provides a good candidate for online secure image communication applications.

**Acknowledgments.** This work was supported by the Fundamental Research Funds for the Central Universities (No. N150402004), and the Online Education Research Fund of MOE Research Center for Online Education (Qtone Education) (No. 2016YB123).

## References

1. Fridrich, J.: Symmetric ciphers based on two-dimensional chaotic maps. *Int. J. Bifurc. Chaos* **8**, 1259–1284 (1998)
2. Fu, C., Meng, W., Zhan, Y., et al.: An efficient and secure medical image protection scheme based on chaotic maps. *Comput. Biol. Med.* **43**, 1000–1010 (2013)
3. Chen, J., Zhu, Z., Fu, C., et al.: Reusing the permutation matrix dynamically for efficient image cryptographic algorithm. *Signal Process.* **111**, 294–307 (2015)
4. Wong, K.W., Kwok, B.S.H., Law, W.S.: A fast image encryption scheme based on chaotic standard map. *Phys. Lett. A* **372**, 2645–2652 (2008)
5. Fu, C., Lin, B., Miao, Y., et al.: A novel chaos-based bit-level permutation scheme for digital image encryption. *Opt. Commun.* **284**, 5415–5423 (2011)
6. Wu, Y., Zhou, Y., Aгаian, S., et al.: A symmetric image cipher using wave perturbations. *Signal Process.* **102**, 122–131 (2014)
7. Li, C., Li, S., Lo, K.T.: Breaking a modified substitution–diffusion image cipher based on chaotic standard and logistic maps. *Commun. Nonlinear Sci. Numer. Simul.* **16**, 837–843 (2011)
8. Li, C., Zhang, L.Y., Ou, R., et al.: Breaking a novel colour image encryption algorithm based on chaos. *Nonlinear Dyn.* **70**, 2383–2388 (2012)
9. Li, C., Xie, T., Liu, Q., et al.: Cryptanalyzing image encryption using chaotic logistic map. *Nonlinear Dyn.* **78**, 1545–1551 (2014)
10. Li, C., Liu, Y., Zhang, L.Y., et al.: Breaking a chaotic image encryption algorithm based on modulo addition and XOR operation. *Int. J. Bifurc. Chaos* **23**, 1350075 (2013)
11. Wang, Y., Wong, K.W., Liao, X., et al.: A chaos-based image encryption algorithm with variable control parameters. *Chaos, Solitons Fractals* **41**, 1773–1783 (2009)
12. Fu, C., Chen, J., Zou, H., et al.: A chaos-based digital image encryption scheme with an improved diffusion strategy. *Opt. Express* **20**, 2363–2378 (2012)
13. Chen, J., Zhu, Z., Fu, C., et al.: An improved permutation-diffusion type image cipher with a chaotic orbit perturbing mechanism. *Opt. Express* **21**, 27873–27890 (2013)
14. Cvitanović, P., Gunaratne, G.H., Procaccia, I.: Topological and metric properties of Hénon-type strange attractors. *Phys. Rev. A* **38**, 1503 (1988)
15. Lü, J., Chen, G.: A new chaotic attractor coined. *Int. J. Bifurc. Chaos* **12**, 659–661 (2002)



# A Distributed Authentication Protocol Using Identity-Based Encryption and Blockchain for LEO Network

Shuai Li<sup>1</sup>, Meilin Liu<sup>2</sup>, and Songjie Wei<sup>1</sup>(✉)

<sup>1</sup> School of Computer Science and Engineering,  
Nanjing University of Science and Technology, Nanjing 210094, China  
{lishuai, swei}@njust.edu.cn

<sup>2</sup> Shanghai Institute of Satellite Engineering,  
Shanghai Academy of Spaceflight Technology, Shanghai 200240, China  
meilinliu51@outlook.com

**Abstract.** LEO satellite-based mobile communication networks have gained enormous attention with the development of communication technology. However, the centralized authentication protocol in traditional satellite networks cannot accommodate to LEO satellite networks with frequent link switching. This paper proposes a fast and efficient authentication protocol in LEO satellite network by combining identity-based encryption and blockchain technology. We simulate the protocol in OPNET for validation with the node modules built in LEO satellite network. The new protocol demonstrates significant advantages in both performance and scalability with less communication overhead for different authentication scenarios in LEO satellite network.

**Keywords:** LEO satellite networks · Authentication · OPNET simulation  
Blockchain · Identity-based encryption

## 1 Introduction

The Low-Earth-Orbit (LEO) satellite network system represented by Iridium system and Globalstar system has become one of the most popular areas of research. Compared with traditional satellite network, because of its low orbit, LEO network has the advantages of short delay and low path-loss. In addition, a constellation of multiple satellites in LEO satellite network system brings true global coverage and the reuse of frequency is more effective. LEO satellite system plays an important role in mobile satellite communication system and is bound to be one of the most important parts within the future global satellite communication system.

Due to the openness of satellite network, communication in satellite network can be easily intercepted by non-authorized or malicious attackers. How to ensure the secure communication in satellite network is the key of achieving security in satellite network system. In communication systems, the use of encryption algorithm to maintain confidentiality is a common and effective method. There is big difference between the satellite network and the ground network in many aspects, such as computing

capability, storage space, high packet loss rate and dynamic topology, etc. Consequently, the traditional authentication represented by a series of protocol with certificates is no longer applicable in satellite scenarios. While considering the design of authentication protocol, we should not only ensure the security of communication, but also reduce the computation and storage overhead. Also the number of steps and nodes involved should be as few as possible.

Different from traditional satellite network, LEO satellite network has the characteristics of dynamic topology and frequent link switching. The authentication protocol running on it must be as light and efficient as possible in premise of ensuring security. Simplicity means that cryptographic computation used in authentication must not be too costly, and storage overhead should not be too high. High efficiency means that the response time of authentication protocol should be short. However, there are lots of drawbacks within the traditional centralized authentication protocols, such as complex computation, central bottleneck or high response time, so they are not efficient enough. We have analyzed the features of LEO satellite network, proposed and simulated a distributed area access authentication protocol based on the mechanism of Identity-based Encryption (IBE) and blockchain. IBE has the characteristics of fast key generation with specified identity string provided by users, which eliminates the cost of certificates used in traditional authentication protocols. Blockchain is decentralized, with reference to this characteristic, we can effectively avoid risks caused by the existence of bottleneck within authentication and implement efficient authentication.

## 2 Protocol Design

In the proposed protocol, the pair of public and private keys are generated quickly by using their identity and the private key of Key Generation Center (KGC). At the same time, based on blockchain, the trust chain consisted of KGC, satellites and users conduct the rapid handover authentication. The distributed storage technology of blockchain records users' registration, cancellation, login, logout, handover and other related log as plugin.

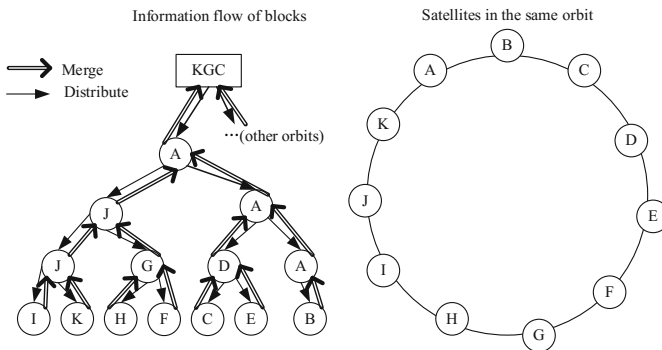


Fig. 1. Logical structure of this system

The protocol is designed in two parts: access authentication and handover authentication. During the process of access authentication, users and satellites can realize mutual authentication through their public and private keys, and the user’s authority is checked with his token. For handover authentication, the fast handover mainly depends on trust chain. Meanwhile, the relevant logs of authentication are recorded in the form of blocks, and these blocks will be merged and distributed between satellites and KGC.

We describe the logical structure of this system as shown in Fig. 1. A satellite in each orbit is selected as a logical root responsible for interaction of blocks with KGC. This logical structure is also the basis of blocks’ merging and distribution.

As background knowledge, we firstly introduce IBE and blockchain technology.

**2.1 Identity-Based Encryption**

In IBE [1], the user’s public key can be obtained directly by his unique identity string, such as phone number, email and so on. Thus it eliminated the overhead brought by certificates. By this way, we can create the mapping between identity and public key.

In addition, a trusted third-party KGC is required. KGC provides key generation services for different roles in this system. When registering, the user need to provide his identity to KGC, then KGC will use its private key and related system parameters to calculate the public and private key pairs of this user, and also securely transmit these to him/her. When sending confidential messages, the user no longer needs to rely on certificates, but only need to use receiver’s public key corresponding to his identity to encrypt messages and then send them. Due to this, IBE avoids the overhead of traditional certificate mechanism, which is just one of the advantages with IBE.

**Table 1.** Symbols and meanings

Symbol	Meaning	Symbol	Meaning
$ID_A$	User A’s ID	$U\_authority$	User’s authority
$ID_S$	Satellite S’s ID	$Start\_time$	Authority’s beginning time
$P_A$	User A’s public key	$Stop\_time$	Authority’s ending time
$P_S$	Satellite’s public key	$XX\_Sign$	XX’s signature
$P_{KGC}$	KGC’s public key	$Auth_{Token}$	User’s authorization token
$d_A$	User A’s secret key	$UserInfo$	User’s related info
$d_S$	Satellite S’s secret key	$Service$	Service that user applies for
$d_{KGC}$	KGC’s secret key	$result$	Result of authentication
$Encry_x()$	Function of encryption with $x$ as key	$Sign_x()$	Function of signing with $x$ as key
$Stime$	Time of handover	$Splace$	Place of handover

**2.2 Blockchain**

Blockchain [2] is the supportive technology of digital encrypted currency represented by Bitcoin. The core strengths are that it can build trust among distributed nodes and

ensure the integrity of data without being tampered or forged. Besides, blockchain supports customization by programming of smart contracts for diverse demands.

Data not being falsified maliciously and reliably decentralized trust are the two main superiorities. The former is guaranteed when each node in the network stores a complete copy of data. And the latter primarily depends on the effectiveness of consensus mechanism in blockchain with no need of Trusted Third Party (TTP) among nodes.

### 2.3 Principles and Processes

This protocol has two major parts: key management implemented with IBE, and authentication and record related logging based on both of blockchain and IBE. For protocol description, we make the symbolic conventions as shown in Table 1. When explaining the principles of each phase, all messages included in this protocol are timestamped by default, and the node receiving messages will always check the timestamp. The flow of this protocol is shown in Fig. 2.

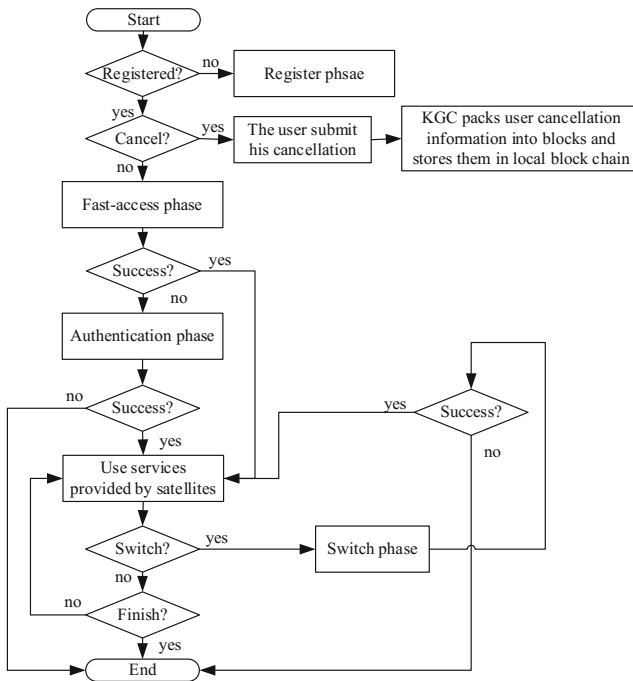


Fig. 2. Flowchart of authentication by the protocol

#### (1) Registration Phase

KGC is a trustworthy center which is responsible for calculating a user’s public key, private key, and user token for authority. If a user has already registered, he is allowed

for access the satellite system at any time during the token's valid period. Otherwise, a legal user identity is needed to be submitted to KGC to obtain a pair of public and private keys calculated by KGC, together with a token signed by KGC.

The calculation is as follows: user A provides his legal identity  $ID_A$  (such as *user: Alice@gmail.com* where "user" means it's the role of user), and then KGC uses hash function and  $P_{KGC}$  to calculate  $P_A$ . Next, KGC calculates  $d_A$  with  $d_{KGC}$ . The satellite will register in the same way before its launch.

Meanwhile, KGC constructs the token of user A and signs it with  $d_{KGC}$ . And  $ID_A||U\_authority||Start\_time||Stop\_time||KGC\_Sign$  is the format of  $Auth_{Token}$  where  $KGC\_Sign$  means signature of the first four fields in this token. After finishing these, KGC returns the pair of public and private keys, and also the token to user A safely. Afterwards, KGC packs the user's registration log into blocks which will be stored in local blockchain. At this point, user A has completed all the preparations before accessing to the satellite system for services.

## (2) Access Authentication Phase

The access authentication is shown in Fig. 3(a), and the four steps are as follows:

- (a) When user A wants to get access to satellite S, A firstly needs to check the identity of S, and then uses hash function to calculate  $P_S$  with  $P_{KGC}$ . Afterwards, A sends his identity to S.
- (b) While receiving this message, S checks the identity of A and searches for latest cancellations to check the validity of A. Then S calculates  $P_A$  accordingly, generates random number  $r$  together with session key  $k$ , and sends  $m1$  to A.  $m1$  is as follows:

$$m1 = Encry_{P_A}(r, k, timestamp, Sign_{d_S}(r, k, timestamp)) \quad (1)$$

- (c) After receiving this message, A decrypts it with  $d_A$ , verifies the signature of  $r$  and  $k$ , and then saves them. Thereafter, A sends  $m2$  to S,  $m2$  is as follows:

$$m2 = Encry_k(r, Auth_{Token}, Service, UserInfo) \quad (2)$$

The *UserInfo* contains the location, time and A's identity when authentication starts.

- (d) While receiving this message, S decrypts it with  $k$ , verifies the correctness of  $r$  and searches for the latest cancellations to verify the validity of current user. If A is valid, then S verifies the signature of A's  $Auth_{Token}$  with  $P_{KGC}$ . The session key  $k$  adopts symmetric encryption, such as Rijndael algorithm. Next, S checks whether  $ID_A$  in  $Auth_{Token}$  is consistent with the identity provided at the beginning or not. And then S provides service according to the authority and expiration time in  $Auth_{Token}$ . Then S sends back  $m3$ ,  $m3$  is as follows:

$$m3 = Encry_k(result, timestamp) \quad (3)$$

If all the steps mentioned above have no mistakes, S allocates relevant resources officially to establish a secure connection with A, moreover, S packs A's login log which contains *UserInfo* mainly into block, then stored it into S's local blockchain. Otherwise, S disconnects from A.

Once upon receiving the message from S, A decrypts it with  $k$ . If it's success response, then A obtain services through the secure channel between himself and S.

With comprehensive analysis, the essence of access authentication phase is to accomplish mutual authentication by IBE. User doesn't need to store the public key of each satellite in advance. Instead, only through the network identification of a satellite broadcasted, user can obtain the corresponding public key directly. And then the safety of authentication is ensured with IBE. During authentication, a session key is negotiated, and a secure channel is established after the success of authentication.

### (3) Fast-Access Authentication Phase

When a new user once get authenticated successfully, his related information will be stored in satellites. Therefore, due to data traceability of blockchain, when this user wants to get access to a satellite for service next time, he only needs to send  $m_4$  to the satellite,  $m_4$  is as follows:

$$ID_A, Service, ID_{S_2}, timestamp, Sign_{ID_A}(ID_A, Service, ID_{S_2}, timestamp) \quad (4)$$

$S_2$  stands for the satellite user A wants to access. Next, after receiving this,  $S_2$  calculates  $P_A$  according to  $ID_A$ , verifies the signature and then check out if  $ID_{S_2}$  in this message is same with its own. If there is no mistake, then it can search for data related to A in its local blockchain, return new session key encrypted with  $P_A$  and provide relevant service according to the data founded.

So, by this way, user can get access to satellites efficiently. The time complexity of searching in blockchain is  $\log_2(n)$ . However, if the satellite this user gets access to is not in the same orbit with the original satellite where the user once gets authenticated successfully, then the user can't take this fast-access way cause there are no related data in this current satellite. Therefore, if the user always wants to use the fast-access way, he needs to ensure that there is at least one satellite in each orbit he has ever got access to successfully by normal access authentication way.

### (4) Handover Authentication Phase

The handover authentication phase is shown in Fig. 3(b), and the four steps included are as follows:

- (a) Through the secure channel, user A informs the satellite (called as  $S_1$ ) of his leaving information including  $ID_A$  and  $ID_{S_2}$ .
- (b) While  $S_1$  receives such messages from A, it checks that whether the satellite A wants to switch to is its neighbor or not. If it is,  $S_1$  will pack A's handover log which is as  $(Stime, Splace, Service, ID_{S_1}, ID_{S_2}, ID_A)$  into block and store this in its local blockchain. Of course, handover log can also be extended according to business needs. At the same time,  $S_1$  returns  $m_5$  to A, which is generated as follows:

$$ID_{S_1}, ID_{S_2}, ID_A, timestamp, Service, Sign(ID_{S_1}, ID_{S_2}, ID_A, timestamp, Service) \tag{5}$$

- (c) After receiving m5, A disconnects from  $S_1$ , signs this message and sends it to  $S_2$ .
- (d) Subsequently,  $S_2$  checks timestamp of the message sent from A, and also checks out whether  $S_1$  is its neighbor. If not,  $S_2$  will refuse A's request. Otherwise,  $S_2$  calculates  $P_{S_1}$  and  $P_A$  to verify the signature in this message, and if it's right, then it needs to search for latest cancellations to check the validity of A. If A is valid, then  $S_2$  returns new session key signed with  $d_{S_2}$  and encrypted with  $P_A$  to A. Later,  $S_2$  officially allocates relevant resources and establishes secure connection with A by this new session key. Meanwhile,  $S_2$  packs A's handover log which depends on the receiving message mainly into block and stores this into its local blockchain.

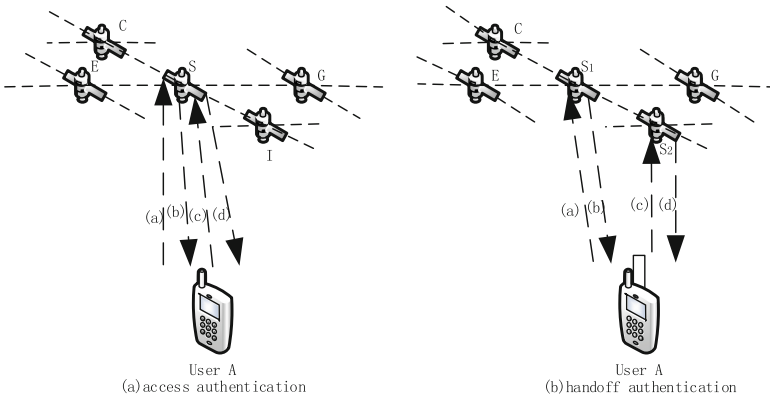


Fig. 3. Diagram of the protocol

Next, A decrypts the message receiving from  $S_2$  with  $d_A$ . If it's the success message, A will verify the new session key's signature and continue to obtain service through new secure channel between him and  $S_2$ . If any step goes wrong,  $S_2$  disconnects from A.

With comprehensive analysis of the principle in handover phase, it's to realize fast handover by the trust chain consisted of satellites and KGC. This is also the consensus among all satellites. In other words, when a user ever gets authenticated successfully which represents that this user has passed the check of one satellite in this system, other satellites should trust the result of authentication. Therefore, this way is more efficient than access authentication phase with low cost in calculation and communication.

As mentioned above, when user logs, logouts or switches among satellites, the satellite should record user's related logs (*UserInfo*), pack them into block and store these blocks into its local blockchain. When a user registers or cancels his identity,

KGC needs pack the related information into block. In system, the relevant information includes registration, cancellation, login, logout and handover these 5 types of records.

When it's time to update (depends on update period), each satellite sends its own latest blocks (namely, blocks that have not been sent out) to adjacent nodes according to the logical structure of system. KGC or satellites will merge these blocks receiving from other nodes with their own blockchain on the basis of timestamp. Finally, when the amount of data at satellite side reaches threshold, each satellite removes those blocks in accordance with rules (such as, for one user, only save the newest record), but all satellites must ensure that every block removed should be ever sent out.

When a user cancels his identity, KGC packs the user's cancellation record into block and stores the block in its local blockchain. And also KGC needs to push these blocks containing the newest cancellation records to logical root node of each orbit periodically or proactively. After receiving these, the logical root nodes distribute these blocks in their orbit successively according to the logical structure shown in Fig. 1.

Whether it is merging or distribution, once one node received blocks, it needs to verify the correctness of blocks' signature, and then integrates these blocks with its local blockchain. The structure of block in this protocol is consistent with Blockchain. Referring to the re-registration, a user reports the loss of his private key and registers with a new identity by the same way described in registration phase.

### (5) Security Analysis

For common attacks such as data tampering, eavesdropping, replay attack, man-in-the-middle attack etc., this protocol which has a good resistance mainly based on the security features of IBE and blockchain is robust. For the malicious attacker, he can't get plaintexts from the ciphertexts obtained by eavesdropping as long as it can't get the private key of any user or satellite; For replay-attacks, the protocol uses timestamp which can resist such attacks very well; For man-in-the-middle attack, the man in the middle can't register with the role of satellite and also the role of user that is already existed in the system. What's more, the attacker can't get the private key of KGC, satellites or registered users. Therefore, he can't disguise himself as any role of the system to conduct man-in-the-middle attack. Also, blockchain can ensure accuracy, completeness, consistency and traceability of data. However, KGC must be completely trustworthy required by IBE, which may have potential safety problems hidden in it. Nevertheless, the actual environment can meet this requirement as any certificate authority (CA) working in good condition.

## 2.4 Related Work

Referring to centralized authentication protocol used in traditional satellite network, Cruickshank has proposed an authentication protocol [3] that uses asymmetrical encryption algorithm. But the operations involved in his protocol are too complicated. Hwang proposed a different authentication protocol without a public-key cryptosystem [4]. Nevertheless, the shared secret key involved needs to be updated every time when any user wants to be authenticated. Chang proposed a mutual authentication protocol that requires only XOR and hash function [5], and during every authentication session, network control center (NCC) doesn't need to generate a private key and a temporary



identity for user. However, NCC is involved in each session during authentication, so its burden is heavy, and also the delay of authentication will be larger. In the papers of Hwang and Chang, NCC is a bottleneck, the performance of authentication protocol is limited to NCC, and once there is any problem with NCC, the authentication protocol will not work properly. Zheng et al. proposed an authentication protocol [6] which avoids these weakness by involving gateway in authentication. But the protocol proposed includes not only users and satellites but also the gateway and NCC, etc. And in this way, the number of interactive steps is increased which results in more response time of authentication. In Lin's paper [7], the advantages of symmetric encryption, asymmetric encryption and certificate system used in satellite network are analyzed.

In summary, PKI (Public Key Infrastructure) is still the typical and important base to implement key management. But with certificates, it was restricted in satellite scenarios which has resource constrain. And also, referring to decentralized authentication protocol used in satellite network, related research are relative lacking. In other resource constrain scenarios similar to satellite network, such as wireless sensor network, the authentication protocols intensively investigated focus on cluster mainly and centralized center is still necessary.

### 3 Simulation and Evaluation

Based on OpenSSL, PBC and GMP libraries, we implement IBE algorithm and compare it to RSA which is recommended by the ISO as the asymmetric encryption standard. For example, in Cruickshank's paper, he uses RSA to implement the function of signature and encryption. In order to analyze the performance of the proposed protocol, we simulate this protocol upon OPNET simulation platform.

#### 3.1 Comparison Between IBE and RSA

To test whether IBE can be used in practice, we compared its performance with RSA algorithm. During programming IBE algorithm, we use SHA1 algorithm that produces 160-bit digest as the hash function. As for RSA algorithm, we directly invoke it from OpenSSL.

The experiment environment used by the test program is Ubuntu 16.04LTS with 4 GB memory and corei5-4590@3.30 GHZ\*4 CPU. After running test program for twenty times, the computational overhead of two algorithms is shown in Fig. 4.

In this experiment, the bilinear pairing used by IBE is generated by the function whose prototype is `pbk_param_init_a_gen (pbk_param_t par, int rbits, int qbits)` in PBC, where `rbits` is 160 and `qbits` is 512 by default. The average time consumed for key generation, encryption and decryption in IBE are 7.251 ms, 1.468 ms and 1.369 ms respectively. For RSA, the time spent for key generation, encryption and decryption are 37.817 ms, 3.753 ms and 4.109 ms on average. It shows that IBE is superior to RSA, and this is mainly because IBE is based on bilinear pair while RSA is based on the difficulty of decomposing a large number. Hence, the performance of IBE can satisfy the need of practical application in satellite network, and some advanced LEO satellite systems such as iridium system already have their own processors in their

satellites which are superior in performance. Moreover, hash function, encryption, decryption and other calculations involved in this protocol can be designed and implemented within particular hardware, so as to further reduce the demand for computing capability of satellite. In terms of the development with IBE, Chinese Office of Security Commercial Code ministration issued the standard of SM9 algorithm which is one kind of identity-based-encryption, and SM9 has entered the phase of promotion. For the security of IBE algorithm, a paper [8] provides a rigorous demonstration.

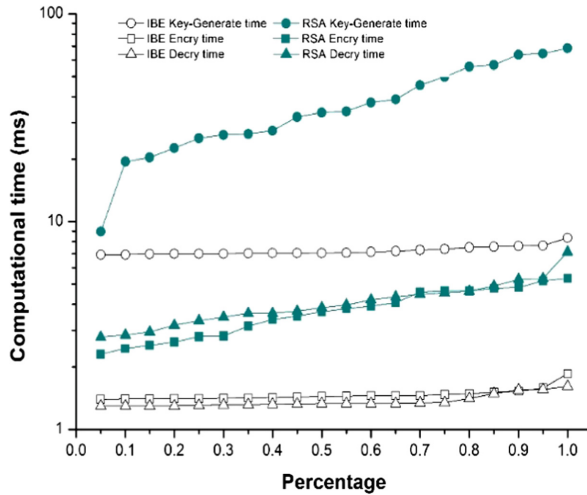


Fig. 4. Comparison between IBE and RSA

### 3.2 OPNET Modeling and Simulation

Due to low orbit, handover is frequent in LEO satellite network. So, in order to ensure the persistence of communication, the authentication protocol designed should be well adapted to this feature and wouldn't introduce too much overhead. In OPNET, we build LEO satellite network scenario [9] consisted of satellite nodes supporting applications attribute and analogous constellation of Iridium which has six orbits and eleven satellites per orbit for simulation without backup satellites. Orbital height is set at 780 km and orbital inclination is set to 86.4°. We use wlan\_workstation\_advanced node as user node. And the topology of simulated LEO satellite network is shown in Fig. 5.

Considering the relative motion between user and satellite, it is reasonable to set user node to be immobile during simulation, and satellites move in their own orbit. The process of this protocol is defined by tasks config. And there are mainly two phases, one is access authentication phase which is defined as challenge\_auth and also fast-access phase which is defined as fast\_access in task\_config object; The other one is handover phase which is defined as switchsat. The size and initialization time of

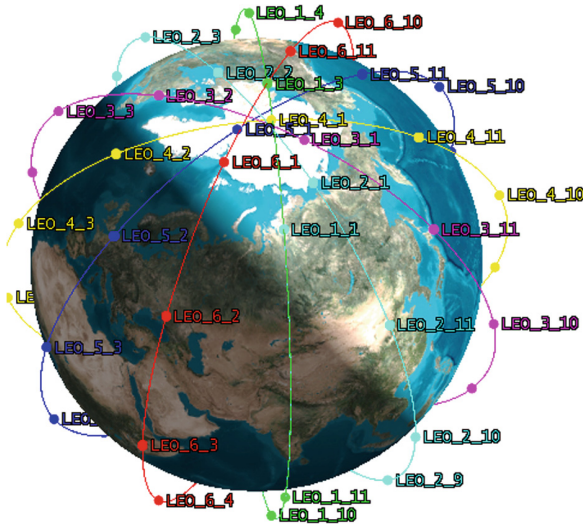


Fig. 5. Topology of simulated LEO satellite network

message used during simulation is based on the size of each field defined in each message and the performance of IBE together with the symmetric encryption (using the AES-192-ECB mode). For example, random number  $r$  used in the protocol is 4 bytes, identity string is no more than 30 bytes, timestamp is 15 bytes and separator between different fields is 2 bytes. Of course, it's just a basic setting for simulation which can be adjusted according to actual business needs. In addition, to build the entire LEO satellite network scenario, it's also necessary to set IP address, routing protocol, signal-to-noise ratio of user and satellite nodes, etc.

### 3.3 Interpretation of Result

Based upon the simulation scenario established, we simulate performance of the protocol in LEO satellite network by setting custom traffic between user and satellite nodes (based on Application config, Task config and Profile config object).

We firstly simulate a complete flow of the protocol, the whole simulation lasts for 500 s, the access authentication occurs at 150 s, the handover authentication occurs at 300 s and the fast-access authentication occurs at 400 s. The results of simulation are shown in Table 2.

Table 2. Response time and delay in each phase of the protocol.

Phase	Src	Dest	Response time (s)	Delay (s)
Access	User	$S_1$	0.17771	0.06737
Access	User	$S_1$	0.32246	0.07591
Fast-access	User	$S_2$	0.18039	0.05363
Handover	User	$S_1$	0.17816	0.05322
Handover	User	$S_2$	0.20689	0.05083

From Table 2, we can see that the response time of each phase in this protocol is less than 500 ms which is far superior to the cost of authentication in the paper [10] (10 s-level) and won't affect the quality of service (QoS) of satellites. At the same time, the packet delay is basically between 50 ms and 70 ms. Compared to this, the average time of encryption, decryption and other processing time can be ignored, this is also the feature that a practical authentication protocol should have. Besides, it's easy to find out the handover authentication phase saves about 100 ms to 150 ms comparing with access authentication phase, and this proves the advantages of fast handover. What's more, we can find out the response time of fast-access authentication phase is shorter than other phases which benefits from the traceability and correctness of data in blockchain.

Then, we adjust the simulation to make it last for two hours. During this simulation, the average interval time of handover is about 10 min which is consistent with iridium system. The results of simulation are shown in Figs. 6 and 7.

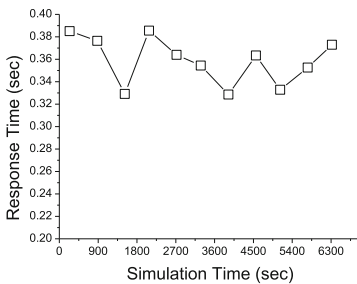


Fig. 6. Response time

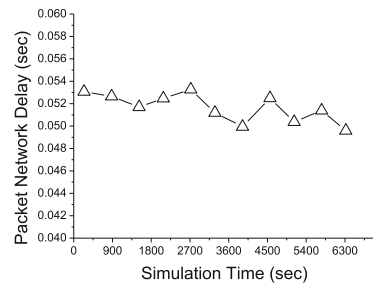


Fig. 7. Packet delay

From Figs. 6 and 7, we can see that the response time of handover authentication is about 360 ms and the delay is about 53 ms which is in accordance with the result of Table 2. And this also proves that our proposed protocol is stable and effective.

For storage overhead, the cost consists of two parts, the first is used to store the public key of KGC and related system parameters, the second is for session key. Taking the amount of users in iridium system which is 150,000 during its heyday as example, the storage used for storing session keys is about 24 MB when 150,000 users are all online at the same time. Therefore, the cost of key storage is much lower than this for each satellite, which is acceptable. Besides, the logging function in this protocol also brings cost of storage, and its size is mainly determined by the threshold for storing blocks. When the amount of blocks reaches threshold, the satellite will delete all related blocks according to the certain rule. In this respect, the threshold specified is the cost of storage for each satellite (for example, threshold can be set to 100 MB, but with the increasing number of users, it needs to be increased).

Assume that the arrival of user obeys poisson distribution and the service time obeys negative exponent distribution. The communication overhead of access, fast-access and handover authentication is  $R_a$ ,  $R_f$ ,  $R_h$ , the average number of users per

hour is  $\lambda$ , the average service time is  $1/\mu$  and the average interval time of handover is  $t$ . Then, for each satellite, the computational overhead brought by this protocol per hour is:

$$x_1 \times \frac{e^{-\lambda} \lambda^{x_1}}{x_1!} \times R_a + x_2 \times \frac{e^{-\lambda} \lambda^{x_2}}{x_2!} \times R_f + \frac{e^{-\lambda} \lambda^{(x_1+x_2)}}{(x_1+x_2)!} \times \left( e^{-\lambda n t} - e^{-\lambda(n+1)t} \right) \times n \times R_h \quad (6)$$

And  $x_1$  represents the number of users who get authenticated by access authentication while  $x_2$  is the number of users who get authenticated by fast-access authentication.

### 3.4 Advantages

According to the analysis and simulation results of this protocol, we can come to a conclusion that this protocol has the following advantages:

- (1) With IBE, this protocol eliminates the cost of certificate. Users and satellites can easily generate the corresponding pair of public and private keys with their own legal identity. Different roles can be distinguished from these identity and the other fields of messages used in this protocol can be extended due to diverse business needs, which provides a high scalability.
- (2) Based on IBE and blockchain, decentralized access authentication and fast handover among satellites are implemented. Additionally, the phases in this protocol are all completely off-line, which means the entire process of authentication only involves satellites and users, which avoids the bottleneck caused by centralized authentication.
- (3) The computational cost is very low. Take symmetric encryption-decryption as P, asymmetric encryption-decryption with IBE as E, signing as S and signature verification as V, the access authentication phase requires 2P, 1E, 1S and 2 V. The fast-access authentication phase includes only 1E, 1S and 1 V. The handover authentication phase only needs 1E, 3S and 3 V. Consequently, this protocol is very efficient.
- (4) Based on the consensus of trust chain, we store information about users and satellites with the technology of blockchain which ensures the accuracy, completeness, consistency and traceability of data within the block. The consensus of trust chain avoids the computational overhead brought by Proof of Work (PoW), and it's also able to make sure that only valid users can get access to valid satellites.
- (5) The logging function can be used as a pluggable module, and you can decide whether to use it according to the actual requirements. In practical cases, if there are only limited computing and storage capacity, we can also use specially designed hardware to do the function of hash, encryption and other types of computing. And also the technology of light chain node realized by SPV (Simplified payment verification) can be used for further reduction in the cost of

storage. However, light chain node will bring more overhead of communication, but it will have less influence on this protocol because of the low delay of LEO satellite network.

## 4 Conclusion and Future Work

Considering the dynamic topology and frequent link switching of LEO satellite network, this paper proposes a new distributed authentication protocol with IBE for key management together with encryption-decryption and blockchain for rapid handover authentication together with logging function. For validation, we simulate this protocol in OPNET. The results of theory analysis and simulation show that the protocol is secure, light-weighted and efficient in LEO satellite network.

Furthermore, in this protocol, although KGC is not involved in authentication, it's still the center for key management. Once it's breached, the security of whole system would be faced with great threat. In spite of some solutions that have solved this partially, there is no real all-around solutions. IBE establishes mapping relations between identity and public key through mathematical methods, now we are working on realizing this kind of mapping relations through smart contract on ethereum, and by applying this, there will be truly no center whether for authentication or key management.

**Acknowledgments.** This material is based upon work supported by the China NSF grant No. 61472189, the CASC Innovation Fund No. F2016020013, the State Key Laboratory of Air Traffic Management System and Technology No. SKLATM201703, and the Postgraduate Research & Practice Innovation Program of Jiangsu Province No. KYCX17\_0369.

## References

1. Wu, J., Long, Y., Huang, Q., et al.: Design and application of IBE email encryption based on Pseudo RSA certificate. In: International Conference on Computational Intelligence and Security, pp. 282–286. IEEE Computer Society (2016)
2. Patel, D., Bothra, J., Patel, V.: Blockchain exhumed. In: Asia Security and Privacy (ISEASP), 2017 ISEA. IEEE (2017)
3. Cruickshank, H.S.: A security system for satellite networks. In: International Conference on Satellite Systems for Mobile Communications and Navigation, pp. 187–190. IET (2002)
4. Hwang, M.S., Yang, C.C., Shiu, C.Y.: An authentication scheme for mobile satellite communication systems. ACM SIGOPS Oper. Syst. Rev. **37**(4), 42–47 (2003)
5. Chang, Y.F., Chang, C.C.: An efficient authentication protocol for mobile satellite communication systems. ACM SIGOPS Oper. Syst. Rev. **39**(1), 70–84 (2005)
6. Zheng, G., Ma, H.T., Cheng, C., et al.: Design and logical analysis on the access authentication scheme for satellite mobile communication networks. IET Inf. Secur. **6**(1), 6–13 (2012)
7. Qi, L., Zhi, L.: Authentication and access control in satellite network. In: Third International Symposium on Electronic Commerce and Security, pp. 17–20. IEEE Computer Society (2010)

8. Chen, L., Cheng, Z.: Security proof of Sakai-Kasahara's identity-based encryption scheme. In: Smart, Nigel P. (ed.) *Cryptography and Coding 2005*. LNCS, vol. 3796, pp. 442–459. Springer, Heidelberg (2005). [https://doi.org/10.1007/11586821\\_29](https://doi.org/10.1007/11586821_29)
9. Long, H.: *OPNET Modeler and Computer Network Simulation*. XiDian University Press, Xi'an (2006)
10. Zhibo, X., Ma, H.: Design and simulation of security authentication protocol for satellite network. *Comput. Eng. Appl.* **43**(17), 130–132 (2007)

# A Detection System for Distributed DoS Attacks Based on Automatic Extraction of Normal Mode and Its Performance Evaluation

Yaokai Feng<sup>1</sup>(✉), Yoshiaki Hori<sup>2</sup>, and Kouichi Sakurai<sup>1</sup>

<sup>1</sup> Faculty of Information Science and Electrical Engineering,  
Kyushu University, Fukuoka 819-0395, Japan

fengyk@ait.kyushu-u.ac.jp, sakurai@csce.kyushu-u.ac.jp

<sup>2</sup> Organization for General Education, Saga University, Saga 840-8502, Japan  
horiyo@cc.saga-u.ac.jp

**Abstract.** Distributed DoS (Denial-of-Service) attacks, or say DDoS attacks, have reportedly caused the most serious losses in recent years and such attacks are getting worse. How to efficiently detect DDoS attacks has naturally become one of the hottest topics in the cyber security community and many approaches have been proposed. The existing detection technologies, however, have their own weak points. For example, methods based on information theory must choose an information theoretic measures carefully which play an essential role on the detection performance and such methods are efficient only when there are a significantly large number of anomalies present in the data; signature-based methods can not deal with new kinds of attacks and new variants of existing attacks, and so on. The behavior-based ones have been thought to be promising. However, they often need some parameters to define the normal nodes and such parameters cannot be determined easily in advance in many actual situations. In our previous work, an algorithm without parameters was proposed for extracting normal nodes from the historic traffic data. In this paper, we will explain a practical off-line detection system for DDoS attacks that we developed based on that algorithm in a project called PRACTICE (Proactive Response Against Cyber-attacks Through International Collaborative Exchange). The general flow of our detection system and the main specific technologies are explained in details and its detection performance is also verified by several actual examples.

**Keywords:** Cyber security · DDoS attacks · Behavior-based detection  
Normal behavior mode · Frequency distribution

## 1 Introduction

The frequency and extent of damages caused by cyber attacks have been increasing greatly in recent years, despite many approaches for protecting and detecting attacks proposed by the network security community and many such practical



systems have been deployed in almost all the important organizations. The basic reason for this is that the technologies used by attackers are also becoming more sophisticated to obtain great economic and/or commercial profit or for national objectives. Of so many cyber attacks, DDoS attacks have reportedly caused the most serious losses in recent years [1]. Distributed attacks means those attacks conducted collaboratively by multiple, even a large number of hosts.

### 1.1 DDoS Attacks

Servers used in many organizations are often of great performance and multi-layer protection/detection measures have been taken including edge protection/detection systems, for example, firewall, DMZ (DeMilitarized Zone) [2]; host protection systems, for example, protection/detection systems for specific attacks or general anomalies; the data protection systems, for example, WAF (Web Application Firewall) [3]; and so on. Also, the network communication performance (e.g. the bandwidth, the communication device performance) has been very developed. Thus, it has become difficult for attackers to achieve their purposes by only one host, for example, denial of services, illegal login, and so on. As a result, actual DoS attacks are always conducted from a large group of hosts that have been compromised by the attackers, that is, so-called DDoS attacks. DDoS attacks have reportedly caused the most serious losses in recent years [1].

Since a DDoS attack exploits many compromised hosts (called bots) send accesses to the target simultaneously, the network bandwidth and/or the computer resources of the victim may be consumed heavily (even exhaustedly) and the victim is forced to deny providing services even to legitimate users. It was reported that DDoS attacks could expose 40% of businesses to losses of 100,000 US\$ or more in an hour at peak times [4].

Many actual distributed cyber attacks have been reported. According to BBC News [5] and an announcement from the Trend Micro Inc. [6], hundreds of thousands of computers were infected with the DNS Changer virus between 2007 and 2011. The botnet (formed from the compromised hosts) was built by a cyber-criminal gang based in Estonia who re-routed computers through fake servers to promote fake products. The number of the bots was up to more than 4,000,000.

Another example was reported in March, 2013 [7]. The non-profit anti-spam organization Spamhaus [8] was suffering from a large DDoS attack against their website at the beginning of 2013. The largest source of attack traffic against Spamhaus came from DNS reflection. In the pick of this attack, over 30,000 unique DNS resolvers were involved as reflectors. The open resolvers responded with DNS zone file, generating collectively attack traffic of approximately 75 Gbps. This translates to each open DNS resolver sending an average of 2.5 Mbps, which is small enough for most DNS resolvers. In this Spamhaus case, the largest pick of the attack traffic which occurred on Mar. 20 was up to over 300 Gbps [9] and finally Spamhaus server was knocked offline [7,9]. A larger

DDoS attack was reported by Cloudflare [10] on Feb. 13, 2014, in which NTP amplification attack [11] was used and the attack peaked at over 400 Gbps.

## 1.2 Existing Technologies for Detection of Cyber-Attacks

The existing detection technologies can be classified into several groups. For example, signature-based methods, volume-based methods, histogram-based methods, and so on. However, it is well-known that signature-based methods are not efficient for new variants and new kinds of attacks, because they can only detect the anomalies stored in a pre-prepared database. Volume-based methods need to determine the thresholds in advance, which is obviously not easy in most applications. In histogram-based methods, many statistic histograms are built as normal ones using clean historical data and the histograms are mapped into a high-dimensional feature space. Then the smallest bounding boxes of those normal histograms (points in feature space) is regarded as the normal region and are used to detect anomalies. This method, however, often has very high false negative rates and clean past traffic data need to be collected, which is obviously not easy, even is impossible, for most actual cases. Also, several methods based on information theory have also been proposed. However, such methods suffer from the following weak points: their performance is highly dependent on the choice of the information theoretic measure; they are efficient only when a significantly large number of anomalies exist in the data and it is difficult to associate an anomaly score with a test instance [12]. The work [12] gives a survey of traditional detection technologies. Some technologies based on change-point detection also have been proposed [13–15]. The session information based on the actions of communication protocols such as TCP and UDP was also used to detect scan attacks [16].

Although many approaches have been proposed, as a matter of fact, threshold-based approaches are still the mainstream methods in the existing IDSs (Intrusion Detection Systems). All the three important IDSs in the public domain, namely Snort [17], Bro [18] and NSM [19], use threshold-based approaches. For example, when Snort observes whether a scan occurred, Snort is configured by default to generate an alert only when 20 different ports are assessed within 60 s. Of course, the parameters can be adjusted manually. Note that, Snort does can not pick out distributed port scans [17, 20]. Another example is in the study [21], which defined distributed port scans by default as being “scans from five or more multiple sources aimed at a particular destination port in the same/24 subnet within one hour.” In the threshold-based methods, thresholds have to be determined in advance and play a key role in detection efficiency. In order to determine the thresholds, however, users have to consider the actual networks and rely on their own experiences since no easy way exists. However, the parameter-tuning is often difficulty even for experts.

Behavior-based detection has been attracting great attentions of many researchers and developers in the cyber security community and has been regarded as one of the most promising methods [22, 23]. Briefly speaking, it

is such methods that attacks are detected using normal behavior-modes (hereafter normal modes) extracted from historical traffics of the monitored networks. That is, anomalies are detected by comparing the real-time traffic patterns with normal nodes.

The biggest challenge being faced by behavior-based detection systems is that how to extract the thresholds that can reflect well the monitored networks from actual historical traffic data in which there are often noises/anomalies. In our previous work [22, 23], an algorithm for extracting normal modes from the historical traffic data has been proposed and discussed in detail. However, that algorithm needs two thresholds. After that, we also proposed one improved algorithm based on automatically thresholding [24, 25]. That is, the new algorithm need not any threshold for extracting the normal mode. In this study, based on our extraction algorithm we proposed in [24, 25], a practical off-line system is implemented for detecting Distributed cyber-attacks. The general flow of our detection system and the main specific technologies will be explained and its detection performance is also verified by several actual examples. This detection system has at least the following strong points.

- (1) The thresholds for distinguishing anomalies from the normal traffic is automatically extracted from the historical traffic. Thus, we no longer need to decide it manually in advance.
- (2) The extracted thresholds can reflect the features of the specific monitored networks. This is important because the network traffics in different organizations are perhaps very different from each other.
- (3) New kinds and new variants of attacks are also able to be picked out.
- (4) Detection is fast because the detection process is only a simple value-comparison.

This paper is organized as follows. Our detection engine is introduced in detail in Sect. 2 including the general idea, general flow, algorithm explanation and several specific technologies. Then, Sect. 3 is our detection cases. Finally, our conclusions and future work are drawn in Sect. 4.

## 2 Our Detection System

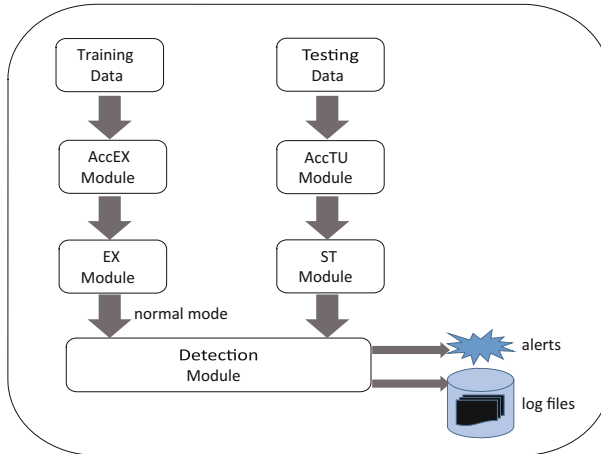
### 2.1 Problem Definition

In this detection system, all the destination ports are monitored separately and an alert will be given if the number of the unique source hosts accessed some port in a time unit increases greatly and suddenly.

In our detection system, the thresholds used to pick out suspicious distributed attacks for destination ports are learned from historical traffic data of the monitored network. We want to note that, this detection system can also be used to monitor specific hosts or to monitor applications. What we need to for those purposes is just to change the statistical objects.

## 2.2 General Flow

Figure 1 shows the general flow. From this figure, we can observe that, not only the attack alerts but also some important information related to the alerts are logged to several files for further analysis and investigations.



**Fig. 1.** General flow.

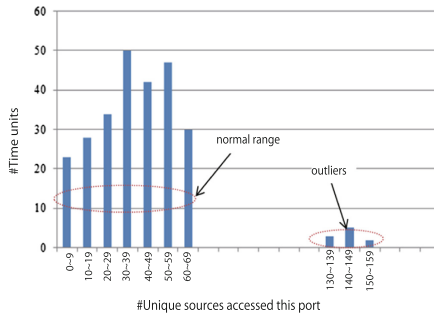
The AccTU Module (Accumulation for each time unit) is to accumulate the test traffic data for every time unit. At the end of each time unit, the ST Module (STatistic Module) is invoked to make statistics in the just-ended time unit. The AccEX Module (Accumulation for EXtracting the normal mode) accumulates the training traffic data in each time unit for extracting the normal mode. The time period for accumulating the data in this module should be long enough (e.g. one month) in order to obtain correct normal mode. The EX Module (EXtracting Module) is invoked to extract the normal mode. Once the current time unit ends in the test data, the Detection Module is invoked to decide whether a DDoS attack occurred or not in this time unit. This decision is made by comparing the statistical result from the test data with the normal mode extracted from the training data. While alerts are given if any, the important information are also stored to several log files (the details will be given later). Moreover, in the Detection Module, according to the relative comparison between the statistical result from the test data and the normal mode, the alert can be associated with a score (or say attack level), indicating the relative scale of the corresponding attack.

## 2.3 Normal Mode Extraction

The approach for automatically extracting the normal mode in the EX Module plays crucial role in this system. Whether correct or roughly correct thresholds

can be extracted or not directly affects the detection result. By contrast, the ST Module is just for making statistics and the Detection Module is just a value comparison. In this subsection, the algorithm for automatically extracting the normal mode that was proposed in our previous work [24, 25] is explained. The biggest advantage of this algorithm is that no parameters are needn't in it (unlike the other algorithms). Thus, the tough parameter-tuning is removed. Simply, this process consists of the following steps.

**Data Preparation:** (AccTU Module in Fig. 1) Collecting traffic data for extracting the normal mode. This is done separately for each of the destination ports because each of the destination ports need a threshold of its own.



**Fig. 2.** Frequency distribution and behavior mode.

**Step 1.** Building frequency distributions. For each of the monitored destination ports, a frequency distribution is built on the number of the unique source hosts that accessed this port in a time unit. Figure 2 shows an example. In the horizontal axis, the number of unique source hosts is divided into bins, while the vertical axis denotes the number of time units. That is, in how many time units the number of unique source hosts drops to the range indicated by the corresponding bin.

**Step 2.** Extracting the thresholds from the frequency distributions. In Fig. 2, the small bins or bin-group located on far-right from the largest group are regarded as outliers, i.e. the anomalies in the learning data. After all the outliers have been removed, the x-axis range of the remaining bins is regarded as the threshold for picking out distributed attacks for the corresponding destination port, which indicates the range of the number of the unique source hosts accessing this port in one time unit.

An algorithm for extracting normal mode from historical traffic was proposed in our previous study [22]. In this algorithm, two parameters indicating area-threshold and distance-threshold need to be predefined. However, determining the two parameters is not an easy issue for real applications. After that, we proposed an extraction algorithm utilizing an evaluation function [24], in which no parameters are needed.

For one frequency distribution an example is shown in Fig. 2, the end-point of every bin is a candidate of the extracting result. Each of the candidates is given a score by a evaluation function and the candidate having the lowest score is reported as the final extracting result. In this way, the extracting result can be decided according to the frequency distribution itself, or say, according to the dataset distribution. Thus, parameters are no longer needed.

The frequency distribution can be denoted by  $(f_1, f_2, \dots, f_n)$ , where  $f_i$  ( $1 \leq i \leq n$ ) is the y-value of the  $i^{th}$  bin;  $n$  is the index of the right-most non-zero bin. In addition, the candidates of the extracting result are denoted by  $\{r_1, r_2, \dots, r_n\}$ , where  $r_i$  ( $1 \leq i \leq n$ ) is the x-value of the end position of the  $i^{th}$  bin. For example,  $r_4$  is 39 in Fig. 2.

This proposal is based on the basic consideration that, when a bin is examined whether or not it should be discarded as an outlier, its y-value and its distance from the origin should be taken into account. The less its y-value is and the farther it is from the origin, the more the possibility that it is regarded as an outlier should be. If a bin should be regarded as an outlier then, all the bins located its right-hand are also thought as outliers at the same time. Thus, when one bin is checked to decide if it is regarded as an outlier, all the bins at its right-hand should also be taken into consideration. Based on these considerations, the evaluation function for the candidate  $r_i$  ( $1 \leq i \leq n$ ) is designed as below.

$$Score(r_i) = \frac{\sum_{q=i}^n sb(f_q)}{n-i} = \frac{\sum_{q=i}^n \frac{f_q}{q}}{n-i}.$$

In this equation,  $f_q/q$  is the local score of the  $q^{th}$  bin, which considers the current bin only. The final score is the average of all the local scores of the bins located on the right of the current bin. At last, The candidate having the lowest score will be the final extracting result of the normal mode. Ties are resolved by choosing the larger one. That is,

$$\begin{aligned} Result &= \arg \min_{r_i} Score(r_i) \\ &= \arg \min_{r_i} \frac{\sum_{q=i}^n \frac{f_q}{q}}{n-i}. \end{aligned}$$

In this way, a normal mode (a threshold) can be obtained for every destination port. However, according to our investigations, in actual situations, there are many in-active ports. Thus, there are many destination ports to which no packets were sent to in the training data accumulated in the AccEX Module. Thus, a minimum threshold must be prepared. In the current version of this system, the minimum threshold is set to five by default.

## 2.4 Data and Global Alerts

The early stage of this study was a part of a project supported by Japan Government, called PRACTICE (Proactive Response Against Cyber-attacks Through

**Table 1.** A real example of Alert-List file.

Alert ID	Date	Time	Des port	Protocol	Alert score
001	03/31/2015	18:58:59	23	TCP	90
002	03/31/2015	18:58:59	4040	TCP	60
003	03/31/2015	18:58:59	8080	TCP	80
004	03/31/2015	18:58:59	9002	TCP	60
005	03/31/2015	22:58:59	9944	TCP	60
006	03/31/2015	22:58:59	10001	TCP	60
...	...	...	...	...	...

International Collaborative Exchange) until FY2015. We use the darknet traffic data sent to our detection system from 10 sensors of PRACTICE deployed in different countries. The countries include Singapore, Philippines, Thailand, USA, France, Japan, Malaysia, Maldives, Brazil and Indonesia. Thus, our detection system can report the detection result for each sensor and also can investigate the timing relationship of the alerts among different sensors.

In this study, the alerts occurring in more than three sensors simultaneously are called *global alerts*. Global alerts should be paid special attention since the same alerts occurred in different regions or countries at the same time.

## 2.5 Output Files

Besides the alert list is given, the details of each alert and the details of the global alerts are also logged as separate files for further investigation and analysis.

- **Alert-List file** for each sensor. Table 1 shows an example, where each line is an alert including alert ID, date, time, port-type (TCP or UDP), port-number, alert score, normal mode and the actual number of the unique source hosts. The alert score is defined according to how many times the actual number of the unique source IPs is more than the normal mode.
- **Alert-Details file** for each sensor. This file has the details of every alert in the mentioned-above AlertList file, including SensorID, Detection date/time, Targeted port-number/port-type, Unique source IP list of this alert. The last number in the parenthesis is the number of the unique source IP addresses of this alert. Table 2 shows an example. Note that the highest octet of the real IP addresses in this table are omitted for the privacy.
- **Global-Alerts file**. Table 3 shows an example of global alert list, where one block is one global attack including date-time, port-number/port-type, the corresponding sensors and the common source IPs of this alert.

**Table 2.** A real example of Alert-Details file.

Alert ID=001 Total number of the unique source IPs = 26 The source IP list: *.20.70.114, *.164.208.51, *.90.11.181, *.6.167.142, *.4.227.197, .....
Alert ID=002 Total number of the unique source IPs = 58 The source IP list: *.20.72.14, *.164.20.31, *.90.11.18, *.4.227.197, *.240.68.153, .....
Alert ID=003 .....

**Table 3.** A real example of Global-Alerts file.

Global alert ID=G001 Destination port=23/TCP Involved sensors = sensor053, sensor056, sensor057, sensor058, sensor059, sensor062 Total number of the common unique source IPs = 106 The common source IP list: *.22.60.24, *.16.38.51, *.91.134.11, *.74.27.27, *.24.8.62, .....
Global alert ID=G002 Destination port=8080/TCP Involved sensors = sensor053, sensor056, sensor062, sensor068 Total number of the unique source IPs = 68 The common source IP list: .....

### 3 Detection Cases

#### 3.1 Detection Examples for Individual Sensors

The traffic data in the previous month of the test data were used as the training data and the time unit was set to four hours. Detection results (alert lists) are well consistent with our manually investigation results. Examples are shown in Figs. 3, 4 and 5.

#### 3.2 Detection Examples of Global Alerts

Using the data from the project PRACTICE, the global alerts from our system and the observation of SANS Internet Storm Center [26] (called SANS hereafter)



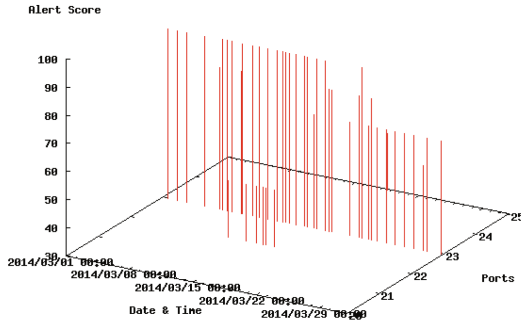


Fig. 3. Detection result for Sensor050 (TCP ports, Mar. 2014).

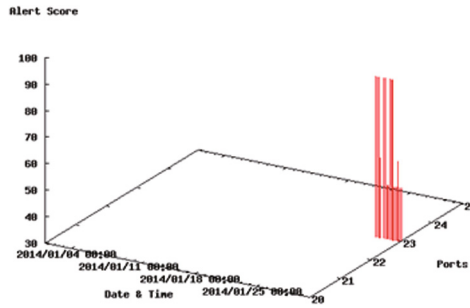


Fig. 4. Detection result for Sensor050 (TCP ports, Jan. 2014).

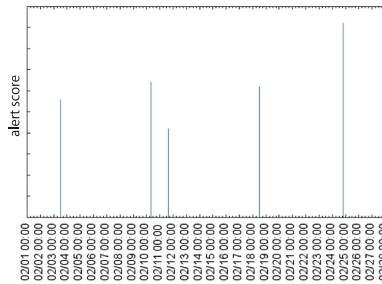
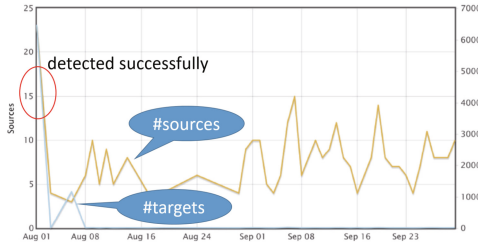


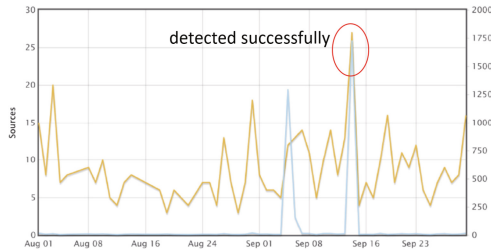
Fig. 5. Detection result for Sensor053 (UDP ports, Feb. 2014).

were compared in order to observe (to some extent, not exactly) the accuracy of the detection performance of our system. Here are some examples.

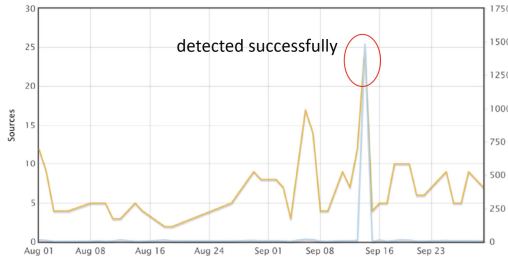
- Global alert 1:
  1. Alert's date and time: 2015/08/01-04:00:00
  2. Destination port: 5924/TCP
  3. Six sensors got involved: sensor053, sensor056, sensor057, sensor058, sensor059, sensor062



**Fig. 6.** The observation of SANS Internet Storm Center and our detection result for 5924/TCP. (Color figure online)



**Fig. 7.** The observation of SANS Internet Storm Center and our detection result for 28015/TCP. (Color figure online)



**Fig. 8.** The observation of SANS Internet Storm Center and our detection result for 8529/TCP. (Color figure online)

Figure 6 shows the observation of this port from SANS in the same time period. The pick marked a red circle seems to correspond to our detection result.

- Global alert 2:
  1. Alert’s date and time: 2015/09/14 16:00:00
  2. Destination port: 28015/TCP
  3. Five sensors got involved: sensor053, sensor056, sensor057, sensor058, sensor062

Figure 7 shows the observation from SANS in the same time period. The pick marked a red circle seems to correspond to our detection result.

- Global alert 3:
  1. Alert’s date and time: 2015/09/14 20:00:00

2. Destination port: 8529/TCP
3. Six sensors got involved: sensor053, sensor056, sensor057, sensor058, sensor059, sensor062

Figure 8 shows the observation from SANS in the same time period. The pick marked a red circle seems to correspond to our detection result.

### 3.3 Observations

From the experiments in this section, we can observe that our detection system is easy to implement and is efficient to detect DDoS attacks. Some actual attacks reported by the SANS Internet Storm Center were captured successfully by our system.

## 4 Conclusion and Future Work

Based on automatically extracting of normal modes using our algorithm, a detection system was implemented for DDoS attacks, one kind of popular and serious attacks. In our system, the thresholds for picking out DDoS attacks are extracted from past traffic data of the monitored network, instead of being given from the user in advance, which is not easy for most situations.

In the future, we will try to find the attack groups of the detected attacks, which is seemingly possible if the source IP addresses of the attacks are analyzed carefully. Moreover, we will try to find the relations among the attacks to the separated destination ports.

**Acknowledgments.** This work was partially supported by Proactive Response Against Cyber-attacks Through International Collaborative Exchange (PRACTICE), Ministry of Internal Affairs and Communications, Japan and partially supported by Strategic International Research Cooperative Program, Japan Science and Technology Agency (JST).

Also, this work was partially supported by JSPS KAKENHI Grant Numbers JP17K00187 and JP16K00132.

This research was also partially supported by Management Expenses Grants of Cybersecurity Center, Kyushu University.

## References

1. Xu, S.: Collaborative attack vs. collaborative defense. In: Bertino, E., Joshi, J.B.D. (eds.) CollaborateCom 2008. LNICSSITE, vol. 10, pp. 217–228. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-03354-4\\_17](https://doi.org/10.1007/978-3-642-03354-4_17)
2. Wiki: DMZ (computing). [https://en.wikipedia.org/wiki/DMZ\\_\(computing\)](https://en.wikipedia.org/wiki/DMZ_(computing)). Accessed 6 Mar 2017
3. Wiki: WAF. <https://en.wikipedia.org/wiki/WAF>. Accessed 6 Mar 2017
4. ComputerWeekly News: <http://www.computerweekly.com/news/4500243431/DDoS-losses-potentially-100k-an-hour-survey-shows>. Accessed 11 June 2016

5. BBC News: Internet lost for thousands using temporary FBI servers. <http://www.bbc.com/news/technology-18769088>. Accessed 6 Mar 2017
6. Trendmicro: Operation Ghost Click. <http://www.trendmicro.co.uk/security-intelligence/research/operation-ghost-click/>. Accessed 6 Mar 2017
7. Cloudflare: The DDoS That Knocked Spamhaus Offline. <https://blog.cloudflare.com/the-ddos-that-knocked-spamhaus-offline-and-ho/>. Accessed 6 Mar 2017
8. Spamhaus Project. <https://www.spamhaus.org/>. Accessed 6 Mar 2017
9. Internet Watch News, 26 April 2013 (in Japanese). <http://internet.watch.impress.co.jp/docs/interview/597628.html>. Accessed 6 Mar 2017
10. Technical Details Behind a 400 Gbps NTP Amplification DDoS Attack, 13 February 2014. <https://blog.cloudflare.com/technical-details-behind-a-400gbps-ntp-amplification-ddos-attack/>. Accessed 6 Mar 2017
11. Wiki: NTP server misuse and abuse. [https://en.wikipedia.org/wiki/NTP\\_server\\_misuse\\_and\\_abuse](https://en.wikipedia.org/wiki/NTP_server_misuse_and_abuse). Accessed 6 Mar 2017
12. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: a survey. *ACM Comput. Surv.* **41**(3), 1–72 (2009)
13. Kim, M.S., Kang, H.J., Hong, S.C.: A flow-based method for abnormal network traffic detection. In: Proceedings of the IEEE/IPIP Network Operations and Management Symposium, pp. 599–612 (2004)
14. Kensuke, K., Hideitsu, H., Murata, N.: Change-point detection in a sequence of bags-of-data. *IEEE Trans. Knowl. Data Eng.* **27**(10), 2632–2644 (2015)
15. Kensuke, K., Hideitsu, H., Murata, N.: Change-point detection in a sequence of bags-of-data. In: The IEEE International Conference on Data Engineering (ICDE), pp. 1560–1561 (2016)
16. Treurniet, J.: A network activity classification schema and its application to scan detection. *IEEE/ACM Trans. Netw.* **19**(5), 1396–1404 (2011)
17. Snort User’s Manual. <http://www.snort.org/docs>. Accessed 11 June 2016
18. The Bro Internet Security Monitor. <https://www.bro.org/>. Accessed 11 June 2016
19. Network and Security Manager (NSM). [https://www.juniper.net/documentation/en\\_US/release-independent/nsm/information-products/pathway-pages/nsm/product/index.html](https://www.juniper.net/documentation/en_US/release-independent/nsm/information-products/pathway-pages/nsm/product/index.html). Accessed 11 June 2016
20. Gates, C.: The modeling and detection of distributed port scans: a thesis proposal, Technical report CS-2003-01, Dalhousie University (2003)
21. Yegneswaran, V., Barford, P., Ullrich, J.: Internet intrusions: global characteristics and prevalence. In: Proceedings of the 2003 ACM Joint International Conference on Measurement and Modeling of Computer Systems, pp. 138–147 (2003)
22. Feng, Y., Hori, Y., Sakurai, K., Takeuchi, J.: A behavior-based method for detecting distributed scan attacks in darknets. *J. Inf. Process. (JIP)* **21**(3), 527–538 (2013)
23. Feng, Y., Hori, Y., Sakurai, K., Takeuchi, J.: A behavior-based method for detecting outbreaks of low-rate attacks. In: Proceedings of the 3rd Workshop on Network Technologies for Security, Administration and Protection (NETSAP), pp. 267–272 (SAINT 2012) (2012)
24. Feng, Y., Hori, Y., Sakurai K.: A proposal for detecting distributed cyber-attacks using automatic thresholding. In: Proceedings of the 10th Asia Conference on Information Security (AsiaJCIS 2015), pp. 152–159 (2015)
25. Feng, Y., Hori, Y., Sakurai, K.: A behavior-based online engine for detecting distributed cyber-attacks. In: Choi, D., Guilley, S. (eds.) WISA 2016. LNCS, vol. 10144, pp. 79–89. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-56549-1\\_7](https://doi.org/10.1007/978-3-319-56549-1_7)
26. SANS Internet Storm Center. <https://isc.sans.edu/>

# A Unified Model for Detecting Privacy Leakage on Android

Xueqi Ren<sup>1</sup>, Xinming Wang<sup>1</sup>, Hua Tang<sup>1</sup>, Zhaohui Ma<sup>1,2</sup>, Jiechao Wu<sup>1</sup>,  
and Gansen Zhao<sup>1</sup> (✉)

<sup>1</sup> School of Computer Science, South China Normal University,  
Guangzhou 510631, China  
gzhao@m.scnu.edu.cn

<sup>2</sup> School of Information Science and Technology, Guangdong University of Foreign  
Studies, Guangzhou 510420, China

**Abstract.** Since Android application may leak user's private information, the issue of Android privacy leakage has rased significant concerns. Various approaches are proposed to detect privacy leakage with different indicators to determine privacy leaks. In this paper, we propose a unified security model to determine privacy leakage. Our unified model includes three typical indicators for detecting privacy leaks, which are sensitive transmission, user intention and application behavior. The proposed model formalizes privacy leakage behavior based on information flows and state transitions. We identify three typical frameworks of privacy leakage. By analyzing the security model, it is feasible to use our model to implement three typical frameworks.

**Keywords:** Privacy · Privacy leakage detection · Privacy model  
Sensitive data · Sensitive data transmission

## 1 Introduction

Nowadays, Mobile Internet has been growing rapidly. Mobile devices are becoming popular and essential for communication. China Statistical Report on Internet Development, issued by CNNIC, shows that the number of people using Internet on mobile phones in China reaches 695 million by the end of December, 2016. This indicates that mobile phone dominates the area of Internet-enabled devices. Moreover, the report of Statista indicates that Android OS captures a huge market share of mobile platform with 86.2%. That means Android OS has the largest number of user groups.

Due to the high mobility, portability and convenience in mobile device, people are increasingly relying on it and willing to interact with each other using their mobile phones in daily life. Mobile phone has a great volumn of privacy data which relate to communication, social network, financial assets and life record of users, e.g. phone id, contacts, call logs, communication records, GPS track and so on. Therefore, as Android OS has a huge market share, the security of Android mobile phones has attracted considerable attention.

However, due to the lack of control in Android application distribution (i.e. user can download apps from third-party application store), it is possible for users to download and install a malicious app that leaks privacy information of users. In addition, as the number of mobile application functions increases, the potential attack surface also increases. Privacy information leakage may result in damage to people's usage of mobile phone and the property of users.

In order to address this issue, researchers detect privacy leakage on Android using static/dynamic analysis, taint analysis and machine learning [1]. Most of the researches consider a privacy leakage to occur when sensitive data are transferred to outside the device. That means, there is a flow of data from an sensitive information source to a sink. But it is not suitable for all privacy leakage problem exactly. For example, users send their location to receive newest local detailed information about weather when they use weather forecast app. In this case, users are aware of this kind of transmission of private data. Therefore, this kind of transmission should not be classified as privacy leakage.

State-of-the-art privacy leakage detection on Android focus on three aspects:

- Transmission between the sensitive data source and outside the boundary. In order to simplify the research problem, many previous research works use this indicator for privacy leakage, and deeply optimize the relevant detection methods on this indicator.
- Consistency of the sensitive data transmission and the user intention. That is, the indicator for privacy leakage detection is whether the transmission of sensitive data match user intention or not. If the transmission of sensitive data is expected by the user, it will be considered as necessary and not a leakage. For example, it determines a privacy leakage to occur if sensitive data transfer outside the device without user permission.
- Consistency of the sensitive data transmission and the application functional behavior. User intention is subjective and diverse. Depending on various user requirement, different users may have different choices when they are in the same situation. In addition, some findings [2] indicate that users demonstrate very low rates of comprehension to permission system. Therefore, in some situation, user intention may lead to privacy leakage in the case of incapable of comprehending system context by some users. It requires an objective indicator to help us determine privacy leakage.

This paper investigates and summarizes related work on privacy leakage under Android platform, and identifies typical frameworks of privacy leakage. This paper proposes a unified security model that is able to formalize the behavior of privacy leakage on Android with all three indicators for detecting privacy leakage. We analyze the security model and use our unified model to implement three typical frameworks.

The main contribution of this paper is as follows:

- We conduct a survey on related works, and identify typical privacy leakage detection frameworks.

- We formalize privacy leakage behaviors based on information flows and state transitions.
- We propose a privacy leakage model which considers and combines three typical indicators for detecting privacy leakage.
- We implement several typical privacy leakage detection frameworks using the proposed model.

The rest of this paper is organized as follows. Section 2 presents a survey on existing works. Section 3 presents an introduction on three typical privacy leakage detection frameworks. Section 4 proposes a unified privacy leakage detection model. Section 5 implements the three privacy leakage detection frameworks using the proposed model. Section 6 concludes the paper.

## 2 Related Work

### 2.1 Transfer of Sensitive Data

The transfer of sensitive data is the main indicator for most existing privacy leakage detection researches.

McClurg and colleagues [6] indicate that normal users are unable to manually keep track of sensitive data used by third-party apps because of time consumption and background knowledge. Therefore, they propose a system to keep track of how the sensitive data is being used inside the app and to detect privacy leakage using dynamic taint analysis. However, due to the high cost of time consumption of dynamic analysis, some researches [7, 8] use static analysis to detect privacy leaks.

AndroidLeaks [3] uses static taint analysis to detect privacy leakage based on transferring data to network. Static taint analysis detects the security issue by labeling the sensitive data and then tracking the propagation of labeled data in program. But there are many other channels which can transfer data to outside the phone such as SMS, NFC. Mann and colleagues [4] summarize a comprehensive category of source and sink including location data, file output, content resolver and so on. Similarly, SCANDAL [5] also determines if there is a data flow from a sensitive source to a sink.

At the view of the privacy leakage among components and applications, FlowDroid [9] uses static analysis to detect intra-component privacy leaks in Android apps. Epicc [10], Amandroid [11] and IccTA [12] perform an ICC analysis to detect inter-component privacy leaks in Android apps. DIALDroid [13] considers the problem of inter-app data leaks.

### 2.2 User Intention

Actually, many benign apps also send sensitive data to outside the device. Therefore, analyzing transmission of sensitive data by considering user intention is more reasonable.

AppInspector [15] determines a privacy leakage as a disclosure of user privacy without user permission or prompt for user. TaintDroid [14] uses dynamic taint analysis to monitor how the private data is being used by app. Capper [16] models the user's decisions with a context-aware policy enforcement mechanism. When a leak is detected, Capper inquires user's decision and is able to remember users preference in the same context.

However, it is possibly infeasible to inquire user whenever a sensitive leak is detected. Pegasus [17] is able to automatically detect sensitive operations being performed without user permission. User intentions are encoded as temporal property of permission event graphs. Then, Pegasus is able to check if the application satisfies the property. AppIntent [18] is able to provide a sequence of GUI manipulations corresponding to the sequence of events that lead to the data transmission, thus helping an analyst to determine if the data transmission is user intended or not.

### 2.3 Application Functional Behavior

Due to the diversity of user intention, user behavior is subjective. Some researches [19,20] indicate that users are not completely understand the Android permission system. There are a few researches focus on the consistency of the sensitive data transmission and the application functional behavior.

DroidJust [21] formulates the problem of sensitive information leakage as a justification problem. It justifies if a sensitive transmission serves any functions of app. AppProfiler [22] describes application behavior profiles. In order to allow users to make informed decision, it creates a knowledge base of mapping between API calls and application behavior types. AppProfiler focuses on detecting how the Android API is used instead of determining what is an unacceptable privacy violation. Based on the characteristics of Android system, VetDroid [23] utilizes the permission usage instead of system call to reconstruct and analyze sensitive behavior in app.

## 3 Typical Framework

Privacy leakage detection is mainly implemented by three approaches: transfer of sensitive data analysis, user intention analysis and application functional behavior analysis.

### 3.1 Transfer of Sensitive Data

AndroidLeaks is a typical framework of privacy leakage based on a transmission between the sensitive data source and outside the boundary. AndroidLeaks defines related sensitive data and operations according to Android permission system. The process of detecting privacy leaks in AndroidLeaks as follows:

Figure 1 shows the analysis process for every Android application. First, AndroidLeaks decompiles apk file, and analyzes AndroidManifest file to obtain



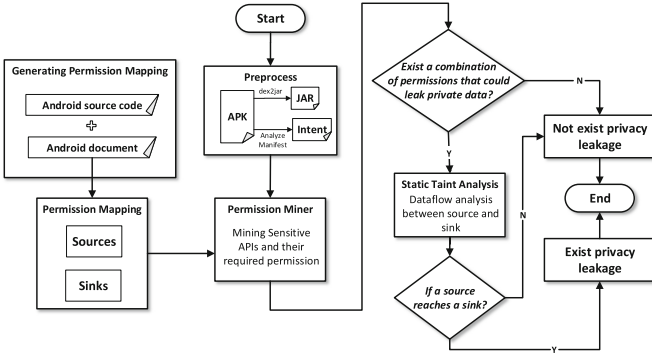


Fig. 1. The process of AndroidLeaks.

permissions and components required by application. It is important to consider the API call that is able to access sensitive permission because accessing sensitive permission means accessing sensitive data. AndroidLeaks automatically constructs the mapping between API calls and sensitive permissions by analyzing the Android Framework source code. AndroidLeaks extracts subset of APIs about sensitive permission as source and sink.

Then, according to call graph and permission mapping, permission miner analyzes sensitive APIs and required permissions of target app, and detects where to call source method or sink method. If the application contains source and sink, then it uses static taint analysis to determine if there is sensitive data flow from a source to a sink. A flow between source and sink indicates a privacy leakage to occur.

### 3.2 User Intention

AppInspector is a typical framework of privacy leakage based on consistency of the sensitive data transmission and the user intention. AppInspector determines a privacy leakage if a disclosure of sensitive data occurs without user permission or notification for users. The process of detecting privacy leaks in AppInspector as follows:

Figure 2 shows the analysis process. First, for every application, AppInspector generates arbitrary sequences of UI and sensor input events. After delivering these inputs to app, information flow tracking module tracks explicit flows and implicit flows. For explicit flow, AppInspector uses dynamic taint analysis to label sensitive data and detects whether the labeled data will leave device through network or not. For implicit flow, AppInspector tracks control dependencies. During tracking, AppInspector logs some information related to sensitive actions in application. In addition, execution exploration module uses concolic execution to explore diverse execution path.

Privacy analysis module analyzes the tracking information flows and the data records to detect malicious behavior. AppInspector determines if a notification

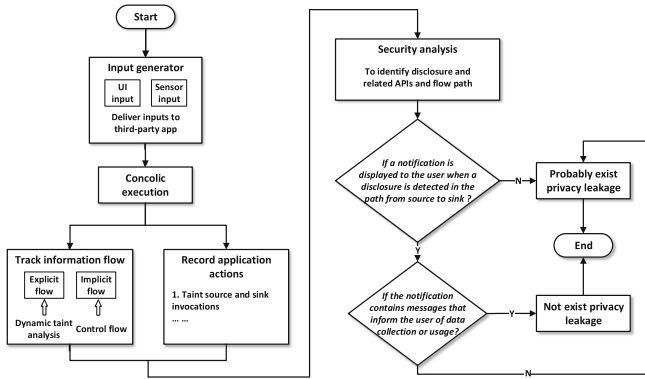


Fig. 2. The process of AppInspector.

that contains messages informing the user of data collection or requesting permission to transmit the data is displayed to the user when a disclosure is detected. A disclosure without such notification indicates a privacy leakage to occur.

### 3.3 Application Functional Behavior

DroidJust is a typical framework of privacy leakage based on consistency of the sensitive data transmission and the application functional behavior. DroidJust tries to relate sensitive transmission with application function and determine if the transmission is justifiable. The process of detecting privacy leaks in DroidJust as follows:

Figure 3 shows the analysis process. First, DroidJust transforms the Dalvik bytecode into the Jimple. The research indicates that app functions are experienced by users during interaction between the user and the app. During the interactions, users are prompted by the changes of sensible phone states (SPS), so app functions are expressed by changing SPS. Then, DroidJust performs a sensitive information transmission analysis, which searches sensitive information flow by parsing permission specifications and uses static taint analysis to identify the sensitive information transmissions. If a response flow is from server and used to change SPS, this flow is called sensible information reception (SIR). Next, DroidJust performs sensible information reception analysis to identify SIR with inbound flows as source and the APIs that can change SPS as sinks.

Finally, DroidJust correlates the identified sensitive information flows and response flows to determine if a sensitive information flow is justifiable. If a sensitive transmission can be linked to a SIR, it is justifiable, otherwise, it is unjustifiable.

## 4 Security Model

Privacy is the sensitive information of entity. Data is a carrier of information, i.e. we express the information by data. Privacy leakage on Android is a sensitive

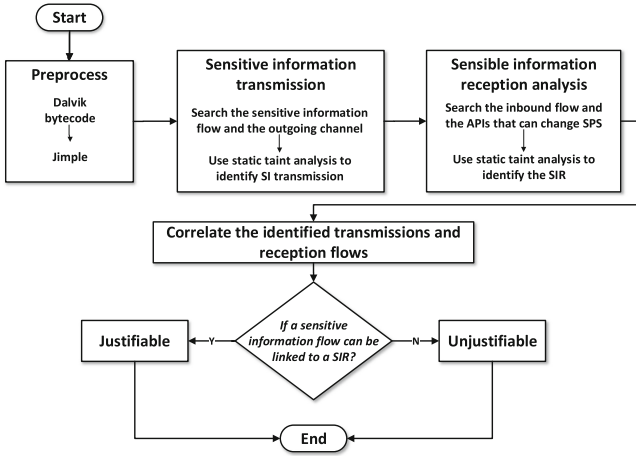


Fig. 3. The process of DroidJust.

transmission flow of data and is generated during the interaction between users and devices. Such sensitive transmission flow transfers sensitive data to malicious area intentionally or unintentionally.

### 4.1 Privacy Leakage Behavior

We construct a symbol system for privacy leakage behavior under Android platform and model the abstract world of privacy leakage. The process of changing data is described as the state transfer in the system. Prior typical frameworks mainly analyze privacy leakage behavior according to taint analysis technique. Taint analysis is able to analyze the data flow of program, mark and trace specific data, then determine whether one specific source reaches another specific sink or not. Similarly, in our security model, we determine a privacy leakage to occur by tracing which type of target area the specific data is transferred to.

Target area is the destination that data will arrive at in the flow. Target area consists of two areas as follow:

- *Security Area*: Security Area is a legal area. If all sensitive data of system reach security area in the flow finally, there is no existing privacy leakage behavior in the system.
- *Irregular Area*: Irregular Area is an illegal area. If there is a sensitive data of system reaches irregular area in the flow finally, privacy leakage behavior exists in the system.

Let  $R = (R_{sec}, R_{nsec})$  be the set of target areas, where  $R_{sec}$  is the security area, and  $R_{nsec}$  is the irregular area. For  $R_{sec}$  and  $R_{nsec}$ , there are:

$$R_{sec} \cup R_{nsec} = R \tag{1}$$

$$R_{sec} \cap R_{nsec} = \emptyset \tag{2}$$

For example, let  $R = (r_1, r_2)$ , which represents  $R$  divides into two areas,  $r_1$  and  $r_2$ ,  $r_1 \in R_{sec}$ ,  $r_2 \in R_{nsec}$ . But the partition of area is dynamic. Hence, there is a process  $f$  such that

$$f : (r_1, r_2) \rightarrow (r_3, r_4)$$

In this case,  $R = (r_3, r_4)$ , where  $r_3 \in R_{sec}$  and  $r_4 \in R_{nsec}$ . Process  $f$  denotes the dynamic partition of target area.

For the process of data transmission, the related modelling elements of system are as follow:

- $V$ : The set of states. State  $v \in V$ , and  $v \in P(D \times R)$ .  $D$  denotes the data set of current state.  $D$  has two categories: tainted data and normal data.  $R$  is the area that data will arrive at.  $R$  also has two categories: security area and irregular area.  $P(D \times R)$  is the power set of  $D \times R$ , that means  $D \times R$  is the set of all possible subsets.  $P(D \times R)$  represents the data in  $D$  reaches the area in  $R$ .

$$(d, r) \in v$$

denotes data  $d$  arrives at area  $r$  in state  $v$ . In our model, we focus on tainted data.

- $C$ : The set of operations or commands, which can trigger the state transfer.
- $W$ : The set of state transfer.

$$W \subseteq V \times V \times C$$

The formula represents an entity makes an operation in  $C$ , moving system from one state in  $V$  to another state in  $V$ .

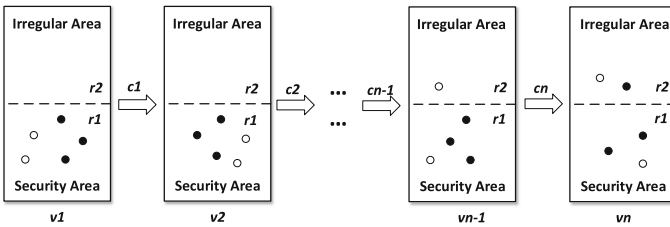


Fig. 4. The unified security model.

Figure 4 shows the unified security model. State  $v_1$  transfers to state  $v_2$  by operation  $c_1$ . The target area is divided into  $r_1$  and  $r_2$ , where  $r_1 \in R_{sec}$  and  $r_2 \in R_{nsec}$ . The solid circle represents a tainted data. The hollow circle represents a normal data. In Fig. 4, state  $v_1$  and  $v_2$  are secure, state  $v_{n-1}$  is also secure, but state  $v_n$  is not secure. Tainted data is sensitive data and we focus on sensitive data instead of normal data. Hence, state  $v_{n-1}$  is secure because there is only a

normal data arrives at irregular area, and state  $v_n$  is not secure because there is a tainted data falls in irregular area.

Assuming that a privacy leakage behavior occurs in a system. We model this system and let  $(1, 2, \dots, i, \dots, n) \in T$  denote a sequence of times.

- $X = V^T$ :  $x \in X$  denotes a sequence of states of system.
- $Y = C^T$ :  $y \in Y$  denotes a sequence of operations of system.

Therefore, the  $i$ th state of system is represented as  $x_i, i \in T$ . The system contains a sequence of states  $\{x_1, x_2, \dots, x_n\}$ . Similarly, system also contains a sequence of operations  $\{y_1, y_2, \dots, y_n\}$ . Every operation is able to trigger a state transfer. If there is a sequence of states  $x_0, x_1, \dots, x_n$  and a sequence of operations  $y_1, \dots, y_n$  such that

$$x_0 \xrightarrow{y_1} x_1 \xrightarrow{y_2} \dots \xrightarrow{y_n} x_n$$

then it indicates that the transfer from  $x_0$  to  $x_1$  is triggered by  $y_1$ , the transfer from  $x_1$  to  $x_2$  is triggered by  $y_2$ , etc. In other words, there is a data flow between  $x_0$  and  $x_n$ .

$\Sigma(W \times C \times x_0) \subseteq X \times Y$  represents the system, and  $x_0$  is the initial state of the system. For  $i \in T$ , the system is in state  $x_{i-1} \in V$ , the system make an operation  $y_{i-1} \in C$ , then the system transfers to state  $x_i$ .  $(x, y) \in \Sigma(W \times C \times x_0)$  if and only if for all  $i \in T, (x_i, x_{i-1}, y_i) \in W$ .

### 4.2 Determination of Privacy Leakage

The privacy leakage behavior is expressed by sequences of actions. In our security model, we use state transfer to describe behavior.

**Definition 1.**  $(v, v', c) \in V \times V \times C$  is a transfer of system  $\Sigma(W \times C \times x_0)$  if and only if  $\{\exists(x, y) \in \Sigma(W \times C \times x_0)\} \wedge \{\exists i \in T, (v, v', c) = (x_i, x_{i-1}, y_i)\}$ .

It is comprehensive to consider user intention when modelling privacy leakage behavior. User permission is the concern in the process of state transfer in our model.

**Definition 2.** Let  $(x, y) \in \Sigma(W \times C \times x_0)$  and  $x_0 \xrightarrow{y_1} x_1 \xrightarrow{y_2} \dots \xrightarrow{y_n} x_n$ . The behavior of system satisfies the user intention if and only if the system satisfies:

$$\exists y_i \in \{\text{Set of User Permission Operations}\}$$

It is comprehensive to consider application functional behavior when modelling privacy leakage behavior. The specific sequences of operations which satisfies the application functional behavior is the concern in the process of state transfer in our model.

**Definition 3.** Let  $(x, y) \in \Sigma(W \times C \times x_0)$  and  $x_0 \xrightarrow{y_1} x_1 \xrightarrow{y_2} \dots \xrightarrow{y_n} x_n$ . Assuming there is a sequence of specific operations that satisfies application functional behavior  $\{z_1, z_2, \dots, z_n\} \subseteq C$ . The system satisfies the application functional behavior if and only if the system satisfies:

$$\{y_1, y_2, \dots, y_n\} = \{z_1, z_2, \dots, z_n\}$$

**Definition 4.** Let  $(x, y) \in \Sigma(W \times C \times x_0)$ . The condition of the process of dynamic partition of target area  $f$  is:

$$\exists y_i \in \{\text{Set of User Permission Operations}\} \vee \exists \{y_1, y_2, \dots, y_n\} = \{z_1, z_2, \dots, z_n\}$$

**Constraint Condition.**  $\forall (x_i, x_{i-1}, y_i) \in W$ , state  $x_{i-1}$  transfers to state  $x_i$  if and only if, for any  $(d, r) \in x_i$  and  $(d, r) \in x_{i-1}$ , there does not exist  $(d_j, R_{sec}) \in x_{i-1} \wedge (d_k, R_{nsec}) \in x_i$ ,  $d_j, d_k \in D$  are tainted data.

**Determination Rule.** For system  $\Sigma(W \times C \times x_0)$ ,  $x_0$  is the initial state. If  $\forall (x_i, x_{i-1}, y_i) \in W$  satisfies Constraint Condition, then system  $\Sigma(W \times C \times x_0)$  is secure. If  $\exists (d_j, R_{sec}) \in x_{i-1} \wedge \exists (d_k, R_{nsec}) \in x_i$  and  $d_j, d_k \in D$  is tainted data, then system  $\Sigma(W \times C \times x_0)$  has privacy leakage behavior.

### 5 Model Analysis

By Definitions 2 and 3, operations  $\{y_1, y_2, \dots, y_n\} \in C$  correlate to user intention and application functional behavior. By Definition 4,  $f$  is enforced when there is an operation that belongs to the set of user permission operations or a sequence of operations that equals to a sequence of specific application functional behavior operations. If process  $f$  is enforced, then security area and irregular area will have new boundaries. Therefore, when operations about user intention or application functional behavior are detected, target area will be redivided. Actually, the repartition area has a significant effect on which area the data will arrives at. For example, in Fig. 5, there is a process  $f : (r_1, r_2) \rightarrow (r_3, r_4)$ , when  $r_1, r_3 \in R_{sec}$ ,  $r_2, r_4 \in R_{nsec}$ . Before enforcing process  $f$ , data  $d$  is in  $r_1$  which is security area. After enforcing process  $f$ , data  $d$  may in the same position, but the boundary of security area and irregular area is redefined. Hence, data  $d$  is in  $r_4$  which is irregular area in this example.

For the typical framework of privacy leakage based on the sensitive transmission, it uses the transfer of sensitive data as indicator to determine privacy

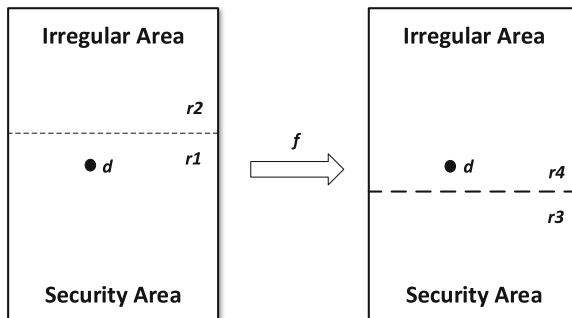


Fig. 5. Example of the repartition target area.

leakage. If sensitive data of system transfers to outside the system, then privacy leakage behavior is detected. That means, the area of outside the system is likely to outside the security boundary. Actually, it is able to abstractly describe this behavior as a sensitive data arrives at a irregular area. Therefore, our unified security model can apply to this typical framework to detect privacy leakage. Take AndroidLeaks as an example, assume system  $\Sigma(W \times C \times x_0)$  has sensitive data  $d$  which are tainted data. For  $i \in T$ , there is a state transfer  $(x_i, x_{i-1}, y_i) \in W$ ,  $\exists(d_j, R_{sec}) \in x_{i-1} \wedge \exists(d_k, R_{nsec}) \in x_i$ , then AndroidLeaks determines whether the system has privacy leaks.

For the typical framework of privacy leakage based on user intention, it uses the consistency of the sensitive data transmission and the user intention as indicator to determine privacy leakage. A sensitive data transmission with user intention may not be true privacy leaks. Take AppInspector as an example. Assume the target system  $\Sigma(W \times C \times x_0)$  has a set of user permission operations.  $\Sigma(W \times C \times x_0)$  performs  $x_0 \xrightarrow{y_1} x_1 \xrightarrow{y_2} \dots \xrightarrow{y_n} x_n$ . In some cases, a sensitive transmission means that there is a data  $d$  arrives at  $R_{nsec}$ . However, a user permission operation is also detected by AppInspector, then the original  $d$  may arrive at  $R_{sec}$  after enforcing the process  $f$ . Finally, AppInspector is able to detect privacy leaks by *Determination Rule*.

For the typical framework of privacy leakage based on application functional behavior, it uses the consistency of the sensitive data transmission and the application functional behavior as indicator to determine privacy leakage. A system transfers sensitive data may serve any functions of application. Take DroidJust as an example. Assume the position of changing SPS is a security area. The system  $\Sigma(W \times C \times x_0)$  performs  $x_0 \xrightarrow{y_1} x_1 \xrightarrow{y_2} \dots \xrightarrow{y_n} x_n$ . DroidJust detects a sensitive transmission as well as a SIR that can be linked. In a sense, DroidJust also uses a flow analysis. Therefore, we are able to define the security area and irregular area to detect privacy leaks by *Determination Rule*. In other cases, a framework [22] focuses on mapping between API calls and application behavior types and detecting how the APIs are used. Therefore, it generates a sequence of application functional behavior operations  $\{z_1, z_2, \dots, z_n\} \subseteq C$ . Then, the following determination is also able to utilize *Determination Rule*.

## 6 Conclusion

In this paper, we investigate the issues of detecting privacy leakage on Android, summarize three typical privacy leakage detection framework and analyze their detection process. We propose a unified security model to determine privacy leakage on Android. Our unified model focuses on three indicators, including the transmission between the sensitive data source and the boundary outside, the consistency of the sensitive data transmission and the user intention as well as the consistency of the sensitive data transmission and the application functional behavior. The proposed model defines a determine rule to detect privacy leaks and suitably apply to the typical frameworks. In future work we plan to implement an experimental verification for perfecting the work by using open source tools.

**Acknowledgments.** This research is supported in part by China MOE Doctoral Research Fund (No. 20134407120017), Guangdong Nature Science Fund (No. S2012030006242).

## References

1. Liang, H., Wu, D., Xu, J., Ma, H.: Survey on privacy protection of android devices. In: 2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing (CSCloud), pp. 241–246. IEEE (2015)
2. Felt, A.P., Chin, E., Hanna, S., Song, D., Wagner, D.: Android permissions demystified. In: Proceedings of the 18th ACM Conference on Computer and Communications Security, pp. 627–638. ACM (2011)
3. Gibler, C., Crussell, J., Erickson, J., Chen, H.: AndroidLeaks: automatically detecting potential privacy leaks in android applications on a large scale. In: Katzenbeisser, S., Weippl, E., Camp, L.J., Volkamer, M., Reiter, M., Zhang, X. (eds.) TRUST 2012. LNCS, vol. 7344, pp. 291–307. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-30921-2\\_17](https://doi.org/10.1007/978-3-642-30921-2_17)
4. Mann, C., Starostin, A.: A framework for static detection of privacy leaks in android applications. In: Proceedings of the 27th Annual ACM Symposium on Applied Computing, pp. 1457–1462. ACM (2012)
5. Kim, J., Yoon, Y., Yi, K., Shin, J., Center, S.: Scandal: static analyzer for detecting privacy leaks in android applications. In: MoST, vol. 12 (2012)
6. McClurg, J., Friedman, J., Ng, W.: Android privacy leak detection via dynamic taint analysis. *EECS* **450**, 2013 (2013)
7. Yang, Z., Yang, M.: Leakminer: detect information leakage on android with static taint analysis. In: 2012 Third World Congress on Software Engineering (WCSE), pp. 101–104. IEEE (2012)
8. Matsumoto, S., Sakurai, K.: A proposal for the privacy leakage verification tool for android application developers. In: Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication, p. 54. ACM (2013)
9. Arzt, S., Rasthofer, S., Fritz, C., Bodden, E., Bartel, A., Klein, J., Le Traon, Y., Outeau, D., McDaniel, P.: Flowdroid: precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for android apps. *ACM Sigplan Not.* **49**(6), 259–269 (2014)
10. Outeau, D., McDaniel, P., Jha, S., Bartel, A., Bodden, E., Klein, J., Le Traon, Y.: Effective inter-component communication mapping in android with epicc: an essential step towards holistic security analysis. In: Proceedings of the 22nd USENIX Security Symposium, pp. 543–558 (2013)
11. Wei, F., Roy, S., Ou, X., et al.: Amandroid: a precise and general inter-component data flow analysis framework for security vetting of android apps. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 1329–1341. ACM (2014)
12. Li, L., Bartel, A., Bissyandé, T.F., Klein, J., Le Traon, Y., Arzt, S., Rasthofer, S., Bodden, E., Outeau, D., McDaniel, P.: IccTA: detecting inter-component privacy leaks in android apps. In: Proceedings of the 37th International Conference on Software Engineering, vol. 1, pp. 280–291. IEEE Press (2015)
13. Bosu, A., Liu, F., Yao, D.D., Wang, G.: Collusive data leak and more: large-scale threat analysis of inter-app communications. In: Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security, pp. 71–85. ACM (2017)



14. Enck, W., Gilbert, P., Han, S., Tendulkar, V., Chun, B.-G., Cox, L.P., Jung, J., McDaniel, P., Sheth, A.N.: Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Trans. Comput. Syst. (TOCS)* **32**(2), 5 (2014)
15. Gilbert, P., Chun, B.-G., Cox, L., Jung, J.: Automating privacy testing of smartphone applications. Duke University (2011)
16. Zhang, M., Yin, H.: Efficient, context-aware privacy leakage confinement for android applications without firmware modding. In: *Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security*, pp. 259–270. ACM (2014)
17. Chen, K.Z., Johnson, N.M., D’Silva, V., Dai, S., MacNamara, K., Magrino, T.R., Wu, E.X., Rinard, M., Song, D.X.: Contextual policy enforcement in android applications with permission event graphs. In: *NDSS* (2013)
18. Yang, Z., Yang, M., Zhang, Y., Gu, G., Ning, P., Wang, X.S.: Appintnet: analyzing sensitive data transmission in android for privacy leakage detection. In: *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, pp. 1043–1054. ACM (2013)
19. Felt, A.P., Ha, E., Egelman, S., Haney, A., Chin, E., Wagner, D.: Android permissions: user attention, comprehension, and behavior. In: *Proceedings of the Eighth Symposium on Usable Privacy and Security*, p. 3. ACM (2012)
20. Kelley, P.G., Consolvo, S., Cranor, L.F., Jung, J., Sadeh, N., Wetherall, D.: A conundrum of permissions: installing applications on an android smartphone. In: Blyth, J., Dietrich, S., Camp, L.J. (eds.) *FC 2012*. LNCS, vol. 7398, pp. 68–79. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-34638-5\\_6](https://doi.org/10.1007/978-3-642-34638-5_6)
21. Chen, X., Zhu, S.: DroidJust: automated functionality-aware privacy leakage analysis for android applications. In: *Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, p. 5. ACM (2015)
22. Rosen, S., Qian, Z., Mao, Z.M.: AppProfiler: a flexible method of exposing privacy-related behavior in android applications to end users. In: *Proceedings of the Third ACM Conference on Data and Application Security and Privacy*, pp. 221–232. ACM (2013)
23. Zhang, Y., Yang, M., Xu, B., Yang, Z., Gu, G., Ning, P., Wang, X.S., Zang, B.: Vetting undesirable behaviors in android apps with permission use analysis. In: *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, pp. 611–622. ACM (2013)

# Multi-party Security Computation with Differential Privacy over Outsourced Data

Ping Li<sup>1</sup>, Heng Ye<sup>2</sup>, and Jin Li<sup>1</sup>(✉)

<sup>1</sup> School of Computer Science and Educational Software, Guangzhou University,  
Guangzhou 510006, China

liping26@mail2.sysu.edu.cn, jinli71@gmail.com

<sup>2</sup> Beijing Key Laboratory of Security and Privacy in Intelligent Transportation,  
Beijing Jiaotong University, 3 Shangyuancun, Beijing 100044, China

heng.ye@bjtu.edu.cn

**Abstract.** Differential privacy has received considerable attention for privacy-preserving machine learning applications. In particular, in the cloud computing environment, data are outsourced from different users. Processing outsourced computations on the joint distribution of multi-party's data under multiple public keys with differential privacy is a significant and difficult problem. In this paper, we propose a scheme named 1, multi-party security computation with differential privacy over outsourced data (MSCD) by using a combination of public-key encryption with a double decryption algorithm (DD-PKE) and  $\epsilon$ -differential privacy to solve this problem. In our work, the cloud server adds the corresponding different statistical noises according to different queries of the data analyst, which differs from previous works in which noise is added by the data provider. In the random oracle model, our scheme is proven to achieve the goal of outsourced computation on the data sets of multiple parties without privacy leakage.

**Keywords:** Differential privacy · Homomorphic encryption  
Outsourcing computation

## 1 Introduction

Machine learning is the process of programming computers to optimize a performance criterion using example data or prior experience. Because of its powerful ability to process large amounts of data, machine learning has been applied in various fields in recent years, including pattern recognition, artificial intelligence, signal processing, networks and data mining.

With increases in the sizes of data and computations, machine learning has been performed on the cloud computing platform due to its limitless computation and storage abilities. However, the cloud server is untrusted; if the outsourced data contain personal information, then such data are vulnerable to attack when processing and storing data. Once these data are collected by companies and

organizations, they can be permanently retained. At this time, cloud users are unable to control how the cloud uses these data, and they are unable to determine what the cloud learns from these data.

Based on this situation, cloud users have increasing concerns regarding the use of their sensitive data. To ensure the fundamental right of users' privacy, appropriate data privacy-preserving techniques to control the risk of leakage must be performed before making data available for computation and analysis. In the existing data security literature, there are two main privacy-preserving methods for controlling the risk of leakage: *data encryption* and *data distortion*.

Data encryption is the process of using an encryption algorithm on the original data in the distributed computing environment such that the original data are hidden and the privacy is protected. In general, techniques based on data encryption have a particular property, such as homomorphic encryption, full homomorphic encryption, order-preserving encryption and so on. To protect the privacy of cloud users' sensitive data, users only outsource their encrypted data to the cloud for storing and processing data. Data encryption has been applied in many fields [26, 28–30]. Data distortion [1, 8] perturbs the original data, such as adding statistical noise or swapping data, but it retains the statistical property over the processed data. Examples of data distortion approaches include  $k$ -anonymity [32, 33],  $l$ -diversity [2],  $t$ -close [34] and  $m$ -invariance [48].

Differential privacy protection is a data-distortion-based method, but the added statistical noise is unrelated to the size of the data set. Therefore, for large data sets, a high level of privacy protection can only be achieved by adding a tiny amount of noise. Dwork [12] provided the original definition of  $\epsilon$ -differential privacy protection. Later several variations on the formal definition of differential privacy, such as [9, 13, 23, 49], were proposed. However, differential privacy protection differs from data distortion in that it defines a very strict attack model and in that it provides a rigorous, quantitative representation and proof about the privacy disclosure risk. Consequently, differential privacy as a practical tool to protect confidential sensitive data from unauthorized access has been widely used in real life.

**Motivation.** In the cloud computing environment, because the cloud is untrusted, multiple users will choose some privacy-preserving technologies to pretreat their original data before uploading the data to the cloud for storing and processing. How to construct a privacy-preserving machine learning model to learn the joint distribution of multi-party settings with multiple public keys remains a challenge. To address this task and to improve the efficiency and accuracy of calculations, we propose a scheme named **Multiparty Security Computation with Differential privacy (MSCD)** over encrypted data to solve this problem.

**Our Contributions.** In our MSCD scheme, we assume that the cloud server and data analyst are not colluding and that they are semi-honest but untrusted. We show that our MSCD scheme is secure in the semi-honest model, the main contributions of this work are summarized as follows:

- In this work, the cloud server has the authority to add different statistical noises to the outsourced data set according to different queries of the data

analyst rather than the data providers adding statistical noise by themselves with only one application.

- We use a DD-PKE cryptosystem to preserve the privacy of the data providers' data sets.
- In our MSCD scheme, the machine learning task is performed on a randomized data set with  $\epsilon$ -differential privacy rather than on the encrypted data set.

## 2 Related Work

Our work is based on several well-known lines of research. In the existing literature, some works are related to ours, and to better understand the relevant research areas, we describe the related work in the following.

***Differential Privacy in Secure Multi-party Computation.*** In the distributed multi-party setting, given  $n$  local data sets  $D_1, D_2, \dots, D_n$  and a function  $f$ , one might aim to compute  $f(\cup_{i=1}^n D_i)$  while satisfying  $\epsilon$ -differential privacy on each local data set  $D_i$  ( $i \in [1, n]$ ). Based on the modulo addition-based encryption scheme, Ács and Castelluccia [21] constructed a differential privacy smart metering system. Toward performing statistical queries over multiple users' private data securely, Chen et al. [37] presented a distributed differential privacy system based on the Goldwasser-Micali (GM) cryptosystem [41]. Based on [21, 22], Goryczka and Xiong [43] proposed an enhanced and fault-tolerant scheme, such that the multiple private data are aggregated with differential privacy.

***Differential Privacy in Multi-party Data Publishing.*** Fung et al. [6] provided some methods and tools for publishing data while guaranteeing data privacy. Mohammed et al. [35] presented a new anonymization algorithm for the non-interactive scenario. In [42], Goryczka et al. considered the problem of how to perform collaborative data publishing for anonymizing horizontally partitioned data of multiple parties with  $m$ -privacy. Chen et al. [38] considered the emerging data publishing setting. Later, Chen et al. [39] considered the high-dimensional data publishing setting. Recently, Hua et al. [24] proposed a privacy-preserving utility verification mechanism for set-valued data in collaborative data publishing. Wang et al. [36] designed an online aggregate monitoring scheme named RescueDP for real-time spatio-temporal crowd-sourced data publishing with differential privacy.

***Differential Privacy in Machine Learning.*** Machine learning research has recently turned to privacy-preserving machine learning with differential privacy protection, including boosting [15], Bayesian inference [7, 16, 20, 27, 44], empirical risk minimization (ERM) [10, 25, 40], stochastic gradient descent (SGD) [5, 45, 47], and linear and logistic regressions [50].

Dwork et al. [14] first considered the topic of privacy-preserving learning. Friedman and Schuster [4] considered machine learning within the framework of differential privacy. Abadi et al. [31] developed new algorithm techniques for learning and a refined analysis of privacy costs under differential privacy, which can be proven to train deep neural networks with differential privacy.

### 3 Preliminaries

In this section, we present some notations and cryptographic primitives that will be used throughout this paper.

#### 3.1 Notations

We denote by  $\mathbb{R}^n$  the  $n$ -dimensional real space and by  $\mathbb{R}^+(\mathbb{Z}^+)$  the space of all positive real (integer) numbers. Let  $[1, n]$  be a set from 1 to a natural number  $n$ . Let  $p, q$  be two primes, and let  $N = qp^2$ . We write  $\mathbb{Z}_p$  as a set of  $\{0, 1, \dots, p-1\}$ , and  $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$ . Let  $\mathcal{X} = \{x \in \mathbb{Z}_{p^2} \mid x \equiv 1 \pmod p\}$  be the  $p$ -Sylow subgroup of  $\mathbb{Z}_{p^2}$ . We use  $\#$  to denote the order of a set or an element, such as  $\mathbb{Z}_{p^2}$  is a cyclic group with order  $p(p-1)$ , i.e.,  $\#\mathbb{Z}_{p^2} = p(p-1)$ , and the order of  $\mathcal{X}$  is  $\#\mathcal{X} = p$ . Based on this fact, we define a function  $\mathcal{L}$  over  $\mathcal{X}$  as follows:

$$\begin{aligned} \mathcal{L} : \mathcal{X} &\rightarrow \mathbb{Z}_p \\ \mathcal{L}(x) &:= \frac{x-1}{p}. \end{aligned} \tag{1}$$

**Definition 1** (*Negligible Functions*). We say that a function  $neg : \mathbb{N} \rightarrow \mathbb{R}$  is negligible if for every positive polynomial  $poly(\cdot)$  and for all sufficiently large  $n$ ,

$$neg(n) < \frac{1}{poly(n)}.$$

We use  $|\cdot|$  to denote the size of data set  $D$  or the bit length of  $x$ , and we use  $\oplus$  to denote the addition mod 2 of the binary vectors.

#### 3.2 Diffie-Hellman Problem over $\mathbb{Z}_N$

The Diffie-Hellman problem [17] as a cryptographic primitive has been widely used in many cryptographic schemes.

**Definition 2** (*p-Diffie-Hellman, p-DH*). Let  $\mathcal{P}(\kappa)$  be a set of all prime numbers with length  $\kappa$ . For any  $p, q \in \mathcal{P}(\kappa)$ , define  $N = qp^2$  and let  $\mathbb{G}_p = \{x \in \mathbb{Z}_N \mid \#(x^{p-1} \pmod{p^2}) = p\}$ . The Diffie-Hellman problem is defined as follows: Given three elements  $a, b \leftarrow \mathbb{Z}_p$ ,  $g \leftarrow \mathbb{G}_p$ , and  $(g^a \pmod N, g^b \pmod N)$ , find  $g^{ab} \pmod N$ .

**Definition 3** (*t-Diffie-Hellman, t-DH*). The  $t$ -DH problem is defined as follows: Given three elements  $a, b \in [1, 2^t - 1] (t > \kappa)$ ,  $g \leftarrow \mathbb{G}_p$ , and  $(g^a \pmod N, g^b \pmod N)$ , find  $g^{ab} \pmod N$ .

#### 3.3 Homomorphic Encryption

Assume there is a public-key encryption scheme  $\mathcal{E} = (\text{KeyGen}, \text{Enc}, \text{Dec})$ , let  $[x]_{\text{pk}}$  be the encrypted data of  $x$  with public key  $\text{pk}$ , i.e.,  $[x]_{\text{pk}} = \text{Enc}(\text{pk}, x)$ .

If the encryption scheme  $\mathcal{E}$  is an additively homomorphic cryptosystem, then an encryption  $[a + b]_{\text{pk}}$  can be computed by  $[a + b]_{\text{pk}} = [a]_{\text{pk}} \otimes [b]_{\text{pk}}$ , where all computations are performed in the encrypted domain. If a message  $a \in \mathbb{P}$  is multiplied by a constant  $b$ , then its encryption of  $ab$  can be computed by  $[ab]_{\text{pk}} = [a]_{\text{pk}}^b$ .

## 4 Problem Formulation

### 4.1 Problem Statement

In this paper, we consider the following problem: *Due to privacy concerns, multiple data providers encrypt their local data sets before uploading to the cloud for storing and processing the data. Based on these aggregated encryptions under different public keys, the cloud sever generates a synthetic data set, in which different statistical noises are added according to different applications. We aim to release this synthetic data with  $\epsilon$ -DP and to perform a privacy-preserving machine learning model on these synthetic data.*

### 4.2 Differential Privacy

Mechanism  $\mathcal{M}$  is a randomized function mapping a data set  $D$  into an output in a range space, i.e.,  $\mathcal{M} : D \rightarrow \text{Range}(\mathcal{M})$ , which is used to publish information. The definition of differential privacy as follows:

**Definition 4** ( $\epsilon$ -Differential Privacy,  $\epsilon$ -DP [12]). *A random mechanism  $\mathcal{M}$  is said to be  $\epsilon$ -differential privacy if for any pair of neighboring data sets  $D$  and  $D'$  and for any possible anonymized data set  $O$  in output range space  $\text{Range}(\mathcal{M})$ ,*

$$\Pr[\mathcal{M}(D) = O] \leq e^\epsilon \times \Pr[\mathcal{M}(D') = O] \quad (2)$$

where the probability  $\Pr[\cdot]$  is taken over the randomness of mechanism  $\mathcal{M}$  and also shows the risk of privacy disclosure.

In this definition,  $\epsilon$  is a predefined privacy parameter for controlling the privacy budget, the smaller  $\epsilon$  is, the stronger is the privacy protection. The formal definition of sensitivity is given below.

**Definition 5** (Sensitivity). *Assume that  $f$  is a numeric query function that maps a data set  $D$  into a  $d$ -dimensional real space  $\mathbb{R}^d$ , i.e.,  $f : D \rightarrow \mathbb{R}^d$ . For any pair of neighboring data sets  $D$  and  $D'$ , the sensitivity  $f$  is defined as*

$$\Delta f = \max_{D, D'} \|f(D) - f(D')\|_{L_1} \quad (3)$$

where  $\|\cdot\|_{L_1}$  denotes the  $L_1$  norm.

**Theorem 1** (*Laplace Mechanism*). Let  $\sigma \in \mathbb{R}^+$ , and  $f$  is a numeric query function that maps a domain  $D$  into a  $d$ -dimension real space  $\mathbb{R}^d$ , i.e.,  $f : D \rightarrow \mathbb{R}^d$ . The computation  $\mathcal{M}$

$$\mathcal{M}(\mathbf{x}) = f(\mathbf{x}) + (\text{Lap}_1(\sigma), \text{Lap}_2(\sigma), \dots, \text{Lap}_d(\sigma)) \tag{4}$$

provides  $\epsilon$ -differential privacy, where the noise  $\text{Lap}_i(\sigma)$  ( $i \in [1, d]$ ) is drawn from the Laplace distribution with scaling parameter  $\sigma$ , whose density function is

$$p(\sigma) = \frac{1}{2\sigma} \exp(-|x|/\sigma). \tag{5}$$

Here, the parameter  $\sigma = \Delta f/\epsilon$  is controlled by the privacy budget  $\epsilon$  and the function's sensitivity  $\Delta f$ .

### 4.3 Public-Key Encryption with a Double Decryption Algorithm

Here, we give a definition of public-key encryption scheme that have a double decryption algorithm, denoted as DD-PKE as follows.

**Definition 6** (*DD-PKE*). A public-key encryption scheme with a double decryption  $\Pi = (\text{Setup}, \text{KeyGen}, \text{Enc}, \text{uDec}, \text{mDec})$  consists of the following PPT algorithms:

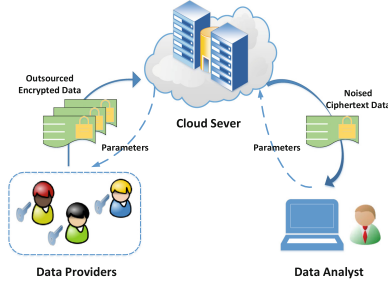
1.  $(\text{pp}, \text{msk}) \leftarrow \text{Setup}(1^\kappa)$ : It takes the system security parameter  $\kappa$  as input and outputs a tuple  $(\text{pp}, \text{msk})$ , where  $\text{pp}$  is a public system parameter and  $\text{msk}$  is the master secret key, which is only known to the master entity.
2.  $(\text{pk}, \text{sk}) \leftarrow \text{KeyGen}(\text{pp})$ : The user takes the  $\text{pp}$  as input and outputs public key  $\text{pk}$  and secret key  $\text{sk}$ .
3.  $c \leftarrow \text{Enc}(\text{pp}, \text{pk}, m)$ : It takes  $\text{pp}$ , a user's  $\text{pk}$  and a message  $m \in \mathbb{P}$  as input and outputs a ciphertext  $c \in \mathbb{C}$ .
4.  $m \leftarrow \text{uDec}(\text{pp}, \text{sk}, c)$ : User entity takes  $\text{pp}$ , a user's  $\text{sk}$  and a ciphertext  $c \in \mathbb{C}$  as input and returns a message  $m \in \mathbb{P}$  or a special symbol  $\perp$ .
5.  $m \leftarrow \text{mDec}(\text{pp}, \text{pk}, \text{msk}, c)$ : Master entity takes the  $\text{pp}$ ,  $\text{pk}$ ,  $\text{msk}$  and a ciphertext  $c \in \mathbb{C}$  as input and returns a message  $m \in \mathbb{P}$  or a special symbol  $\perp$  that indicates that the ciphertext was rejected.

## 5 System and Adversary Models

In this section, we present the definitions of our system model and the adversary model.

### 5.1 System Model

Our system consists of a cloud server  $\mathcal{C}$ , a data analyst  $\mathcal{DA}$  and a data provider set  $\mathcal{DP}$  (see Fig. 1).



**Fig. 1.** System model under consideration

Assume that the data provider set  $\mathcal{DP}$  contains  $n$  data providers, denoted as  $P_1, P_2, \dots, P_n$ , and that each data provider  $P_i \in \mathcal{DP}$  is a cloud user who keeps data set  $D_i = \{(\mathbf{x}_j^i, \mathbf{y}_j^i) \in \mathbf{X} \times \mathbf{Y} : j \in [1, p_i], i \in [1, n]\}$ . Each  $D_i$  ( $i \in [1, n]$ ) is of size  $p_i$  with data vector  $\mathbf{x}_j^i \in \mathbb{R}^d$ , and the corresponding binary label  $y_j^i \in \mathbf{Y} := \{0, 1\}$ . To protect the privacy of data set, each data provider  $P_i \in \mathcal{DP}$  encrypts its sensitive data set  $D_i$  ( $i \in [1, n]$ ) using the DD-PKE scheme with its own public key  $\text{pk}_i$  before outsourcing to  $\mathcal{C}$ .

$\mathcal{C}$  is semi-honest and holds a data center, it can aggregate the combined data sets from the various cloud users and publish the data sets according to the task of the data analyst  $\mathcal{DA}$ . Note that  $\mathcal{C}$  owns the aggregated data sets encrypted with *different public keys*. Due to the privacy concerns,  $\mathcal{C}$  adds a Laplace noise  $\boldsymbol{\eta}_i$  to the ciphertext for each data provider  $P_i \in \mathcal{DP}$ . Thus, the noise-added encrypted data set can be computed by  $\text{Enc}(\text{pk}_i, \hat{D}_i) = \text{Enc}(\text{pk}_i, D_i) \otimes \text{Enc}(\text{pk}_i, \boldsymbol{\eta}_i)$ . Then,  $\mathcal{C}$  publishes these noise-added data sets encrypted with *different public keys*.

$\mathcal{DA}$  downloads the published data from  $\mathcal{C}$  and decrypts them because  $\mathcal{DA}$  holds the master secret key  $\text{msk}$ . Later,  $\mathcal{DA}$  obtains the aggregated data set  $\hat{D} = \bigcup_{i=1}^n \hat{D}_i$ , where  $\hat{D}_i$  has added noise and is randomized. Based on  $\hat{D}$ ,  $\mathcal{DA}$  can train a machine learning model with  $\epsilon$ -DP.

## 5.2 Adversary Model

In this paper, we assume that data provider  $P_i \in \mathcal{DP}$  ( $i \in [1, n]$ ),  $\mathcal{C}$  and  $\mathcal{DA}$  are semi-honest but untrusted. Additionally, we assume that there is no collusion between  $\mathcal{C}$  and  $\mathcal{DA}$ , between any two data providers or between any data provider and  $\mathcal{DA}$ . Based on this security assumption, we present an active adversary  $\mathcal{A}$  in our scheme. The aim of  $\mathcal{A}$  is to obtain the plaintext of  $\mathcal{DP}$ 's data with the following abilities:

1.  $\mathcal{A}$  is able to collude with  $\mathcal{DA}$  to obtain plaintexts of all ciphertext data downloaded from  $\mathcal{C}$  by running an interactive protocol.
2.  $\mathcal{A}$  may corrupt  $\mathcal{C}$  to guess the plaintexts of all ciphertext data outsourced from  $P_i \in \mathcal{DP}$  ( $i \in [1, n]$ ) and all data sent from  $\mathcal{DA}$  by performing an interactive protocol.
3.  $\mathcal{A}$  may corrupt some data providers of  $\mathcal{DP}$  to generate plaintext information of other data providers' ciphertexts.



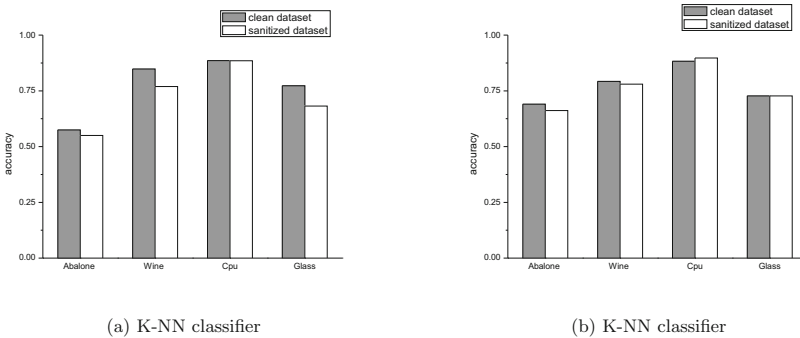
## 6 Our Solution

In this section, we describe our solution for the problem formulated in Sect. 4. We first present the main steps of our solution. We then describe in detail the construction of our solution based on the DD-PKE cryptosystem  $\mathcal{H}$  in the remaining subsections.

1. **Initialization.** In this step,  $\mathcal{DA}$  runs a **Setup** algorithm to set up the DD-PKE system and distributes the system’s public parameter  $\mathbf{pp}$  to the cloud server.
2. **Data Uploading.** After obtaining the system’s public parameter  $\mathbf{pp}$  sent from  $\mathcal{C}$ , data providers generate their own public/secret keys using the algorithm **KeyGen** and upload the encrypted data set under their own public key to  $\mathcal{C}$ .
3. **Noise Adding.** In this phase, according to the differential application or queries of  $\mathcal{DA}$ , cloud server  $\mathcal{C}$  adds differential Laplace noise to these outsourced ciphertexts. Here, the Laplace noises are encrypted under the corresponding outsourced ciphertexts. Later,  $\mathcal{C}$  publishes these noise-added ciphertexts to  $\mathcal{DA}$ .
4. **Machine Learning-Based  $\epsilon$ -DP.** After downloading the noise-added ciphertexts from  $\mathcal{C}$ ,  $\mathcal{DA}$  can decrypt these ciphertexts using the **mDec** algorithm since he has the master private key  $\mathbf{msk}$ . Then,  $\mathcal{DA}$  keeps a synthetic data set with added noise. Based on this new data set,  $\mathcal{DA}$  can learn a machine learning model with  $\epsilon$ -DP.

## 7 Simulation Results

In this section, we show how we can use our scheme to preserve data privacy according to the DD-PKE cryptosystem  $\mathcal{H}$  and  $\epsilon$ -differential privacy for machine



**Fig. 2.** Simulation results: (a) K-NN classifier with the parameter  $k = 1$ . (b) K-NN classifier with the parameter  $k = 5$ .

learning by adding noise to the input. All simulations are conducted on a PC with an AMD A4-3300M APU with Radeon (TM) HD Graphics 1.90 GHz and 6 GB of RAM. In this paper, we use the data sets **Abalone**, **Wine**, **Cpu** and **Glass** as our test data sets, which can be downloaded from the UCI Machine Learning Repository. To simulate the K-nearest neighbor (K-NN) classifier’s performance, we withdraw  $\frac{1}{10}$  of the data sets’ records to compose the test data set. Additionally, we choose the privacy level  $\epsilon = 0.1$  to apply the Laplace mechanism to our four data sets (see Fig. 2).

## 8 Security Analysis

In this section, we first present the security analysis of the basic cryptographic encryption primitive and  $\epsilon$ -DP before analyzing the security of our MSCD scheme.

### 8.1 Analysis of Encryption Primitive

We consider the security of the DD-PKE cryptosystem  $\Pi$  against *adaptive chosen ciphertext attacks* (CCA2) as follows:

**Definition 7** (*IND-DD-CCA2*). Let  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  be an adversary. We say that a DD-PKE cryptosystem  $\Pi = (\text{Setup}, \text{KeyGen}, \text{Enc}, \text{uDec}, \text{mDec})$  has the *indistinguishability property against adaptive chosen ciphertext attacks* (IND-DD-CCA2 secure) if, for any system’s security parameter  $\kappa \in \mathbb{N}$  and any PPT IND-DD-CCA2 adversary  $\mathcal{A}$  against the cryptosystem  $\Pi$ , the function

$$\begin{aligned} \text{Adv}_{\Pi, \mathcal{A}}^{\text{ind-cca2}}(\kappa) &= 2Pr[(\text{pk}, \text{sk}) \leftarrow \text{KeyGen}(1^\kappa); \\ &\quad (m_0, m_1, \text{state}) \leftarrow \mathcal{A}_1^{\mathcal{O}^u}(\text{pk}); \\ &\quad b \leftarrow \{0, 1\}; y \leftarrow \text{Enc}(\text{pk}, m_b) : \\ &\quad \mathcal{A}_2^{\mathcal{O}^m}(\text{pk}, m_0, m_1, \text{state}, y) = b] - 1. \end{aligned}$$

is negligible, where  $\mathcal{O}^u(\cdot) = \text{uDec}(\cdot)$  and  $\mathcal{O}^m(\cdot) = \text{mDec}(\cdot)$ . ‘state’ is secret information, possibly including  $\text{pk}$  and messages  $m_0$  and  $m_1$  with  $|m_0| = |m_1|$ .

**Definition 8** (*OTE*). Let  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  be a PPT adversary. We say that a symmetric encryption  $SE = (\text{Enc}, \text{Dec})$  is *OTE secure* if the function

$$\begin{aligned} \text{Adv}_{SE, \mathcal{A}}^{\text{OTE}} &= 2Pr[\kappa \leftarrow \{0, 1\}^l; \\ &\quad (m_0, m_1, \text{state}) \leftarrow \mathcal{A}_1(\cdot); \\ &\quad b \leftarrow \{0, 1\}; c^* \leftarrow \text{Enc}(\kappa, m_b) : \\ &\quad b \leftarrow \mathcal{A}_2(m_0, m_1, \text{state}, y) = b] - 1 \end{aligned}$$

is negligible.

## 8.2 Analyzing the Security of $\epsilon$ -Differential Privacy

In this subsection, we describe the following theorem:

**Theorem 2** (*Parallel Composition*). *Let  $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_n$  be  $n$  mechanisms, where each mechanism  $\mathcal{M}_i$  ( $i \in [1, n]$ ) provides  $\epsilon_i$ -DP. Let  $D_1, D_2, \dots, D_n$  be  $n$  arbitrary disjoint data sets of the input domain  $D$ . For a new mechanism  $\mathcal{M}$ , the sequence of  $\mathcal{M}(\mathcal{M}_1(D_1), \mathcal{M}_2(D_2), \dots, \mathcal{M}_n(D_n))$  provides  $(\max_{1 \leq i \leq n} \epsilon_i)$ -DP.*

## 8.3 Security of MSCD Scheme

On the one hand, our MSCD scheme is based on DD-PKE cryptosystem  $\Pi$ , the security of which can be guaranteed by the following theorem:

**Theorem 3.** *In the random oracle model, our solution based on DD-PKE cryptosystem  $\Pi$  is IND-DD-CCA2 secure if the computational  $t$ -DH problem is intractable and the  $SE$  is OTE secure for a message  $m \in \mathbb{P}$  and randomness  $r \in \mathbb{Z}_{2^{\kappa}-1}$ .*

*Proof.* Assume that the  $t$ -DH problem is not intractable and that the symmetric encryption scheme  $SE = (Enc, Dec)$  is not OTE secure. Here,  $Enc(\lambda, m) = \lambda \oplus (m || r)$ . Then, a PT adversary  $\mathcal{A}$  exists that can solve the  $t$ -DH problem, or it must break the  $SE$  scheme with non-negligible probability. We can construct an adversary  $\mathcal{A}^*$  with the help of  $\mathcal{A}$ , which either solve the  $t$ -DH problem or break the  $SE$  scheme.

Suppose that  $\mathcal{B}^*$  breaks the  $SE$  scheme, which runs  $\mathcal{A}$  as a subroutine. Since  $\mathcal{A}$  can solve the  $t$ -DH problem and break the  $SE$  scheme with non-negligible probability, then  $\mathcal{A}^*$  and  $\mathcal{B}^*$  have no advantage against the IND-DD-PKE and OTE according to the above algorithms, respectively. Consequently,  $\text{Adv}_{\Pi, \mathcal{A}^*, \mathcal{B}^*}^{\text{ind-cca2}}(\cdot)$  is with non-negligible probability.

On the other hand, our MSCD scheme can protect against the system attacker described in Sect. 5.2. First, if  $\mathcal{A}$  has corrupted  $\mathcal{DA}$  or  $\mathcal{C}$  to obtain the outsourced data,  $\mathcal{A}$  cannot obtain the corresponding plaintext by using the master key  $msk$  because of the IND-DD-CCA2 security of our MSCD scheme. Second,  $\mathcal{A}$  has corrupted some data providers of  $\mathcal{DP}$  and obtains the public/private keys of these corrupted data providers. Due to the non-interactive and independent key generation algorithm of the data providers, these multiple private/public keys are uncorrelated. Therefore,  $\mathcal{A}$  still cannot decrypt the ciphertext.

## 9 Conclusion

In this paper, we proposed MSCD, a scheme for a differential privacy outsourcing computation system over the joint distribution of a multi-party setting under multiple public keys, which allows multiple data providers to outsource encrypted data sets to a cloud server for storing and processing data. In our work, the cloud

server can add different statistical noises to the outsourced data sets according to the different queries of the data analyst, which is different from existing works (i.e., data providers add statistical noise by themselves). Our work is mainly based on DD-PKE and  $\epsilon$ -DP technology, which can be proven to achieve the goal of outsourced computation on multi-party's data sets without privacy leakage in the random oracle model.

**Acknowledgments.** This work was supported by Natural Science Foundation of Guangdong Province for Distinguished Young Scholars (2014A030306020), Guangzhou scholars project for universities of Guangzhou (No. 1201561613), Science and Technology Planning Project of Guangdong Province, China (2015B010129015), National Natural Science Foundation of China (Nos. 61472091, 61702126) and National Natural Science Foundation for Outstanding Youth Foundation (No. 61722203).

## References

1. Polak, A.C., Goeckel, D.L.: Identification of wireless devices of users who actively fake their RF fingerprints with artificial data distortion. *IEEE Trans. Wirel. Commun.* **14**(11), 5889–5899 (2015)
2. Machanavajjhala, A., Kifer, D., Gehrke, J., Venkatasubramanian, M.: L-diversity: privacy beyond  $k$ -anonymity. *ACM Trans. Knowl. Discov. Data (TKDD)* **1**(1), 3 (2007)
3. Yao, A.C.: How to generate and exchange secrets. In: *Proceedings of the 27th IEEE Symposium on Foundations of Computer Science (FOCS)*, Toronto, Canada, pp. 162–167 (1986)
4. Friedman, A., Schuster, A.: Data mining with differential privacy. In: *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 493–502. ACM (2010)
5. Rajkumar, A., Agarwal, S.: A differentially private stochastic gradient descent algorithm for multiparty classification. In: *International Conference on Artificial Intelligence and Statistics*, pp. 933–941 (2012)
6. Fung, B.C.M., Wang, K., Chen, R., Yu, P.S.: Privacy-preserving data publishing: a survey of recent developments. *ACM Comput. Surv.* **42**(4), 1–53 (2010)
7. Yang, B., Sato, I., Nakagawa, H.: Bayesian differential privacy on correlated data. In: *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, pp. 747–762. ACM (2015)
8. Liew, C.K., Choi, U.J., Liew, C.J.: A data distortion by probability distribution. *ACM Trans. Database Syst. (TODS)* **10**(3), 395–411 (1985)
9. Chatzikokolakis, K., Andrés, M.E., Bordenabe, N.E., Palamidessi, C.: Broadening the scope of differential privacy using metrics. In: De Cristofaro, E., Wright, M. (eds.) *PETS 2013*. LNCS, vol. 7981, pp. 82–102. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-39077-7\\_5](https://doi.org/10.1007/978-3-642-39077-7_5)
10. Talwar, K., Thakurta, A., Zhang, L.: Private empirical risk minimization beyond the worst case: the effect of the constraint set geometry. *arXiv preprint arXiv:1411.5417* (2014)
11. Okamoto, T., Uchiyama, S.: A new public-key cryptosystem as secure as factoring. In: Nyberg, K. (ed.) *EUROCRYPT 1998*. LNCS, vol. 1403, pp. 308–318. Springer, Heidelberg (1998). <https://doi.org/10.1007/BFb0054135>

12. Dwork, C., McSherry, F., Nissim, K., Smith, A.: Calibrating noise to sensitivity in private data analysis. In: Halevi, S., Rabin, T. (eds.) TCC 2006. LNCS, vol. 3876, pp. 265–284. Springer, Heidelberg (2006). [https://doi.org/10.1007/11681878\\_14](https://doi.org/10.1007/11681878_14)
13. Dwork, C., Rothblum, G.N.: Concentrated differential privacy. arXiv preprint [arXiv:1603.01887](https://arxiv.org/abs/1603.01887) (2016)
14. Dwork, C., Naor, M., Pitassi, T., Rothblum, G.N.: Differential privacy under continual observation. In: Proceedings of the Forty-second ACM Symposium on Theory of Computing, pp. 715–724. ACM (2010)
15. Dwork, C., Rothblum, G., Vadhan, S.: Boosting and differential privacy. In: FOCS (2010)
16. Dimitrakakis, C., Nelson, B., Zhang, Z., Mitrokotsa, A., Rubinstein, B.: Differential privacy for bayesian inference through posterior sampling. *J. Mach. Learn. Res.* **18**(11), 1–39 (2017)
17. Diffie, W., Hellman, M.E.: New directions in cryptography. *IEEE Trans. Inf. Theory* **22**(6), 644–654 (1976)
18. Kiltz, E., Malone-Lee, J.: A general construction of IND-CCA2 secure public key encryption. In: Paterson, K.G. (ed.) Cryptography and Coding 2003. LNCS, vol. 2898, pp. 152–166. Springer, Heidelberg (2003). [https://doi.org/10.1007/978-3-540-40974-8\\_13](https://doi.org/10.1007/978-3-540-40974-8_13)
19. Barthe, G., Köpf, B., Olmedo, F., Zanella-Béguelin, S.: Probabilistic relational reasoning for differential privacy. In: POPL (2012)
20. Barthe, G., Farina, G.P., Gaboardi, M., Arias, E.J.G., Gordon, A., Hsu, J., Strub, P.Y.: Differentially private Bayesian programming. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 68–79. ACM (2016)
21. Ács, G., Castelluccia, C.: I have a DREAM! (DiffeRentially privatE smArt Metering). In: Filler, T., Pevný, T., Craver, S., Ker, A. (eds.) IH 2011. LNCS, vol. 6958, pp. 118–132. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-24178-9\\_9](https://doi.org/10.1007/978-3-642-24178-9_9)
22. Shi, E., Chan, H.T.H., Rieffel, E., Chow, R., Song, D.: Privacy-preserving aggregation of time-series data. In: Annual Network and Distributed System Security Symposium (NDSS). Internet Society (2011)
23. Mironov, I., Pandey, O., Reingold, O., Vadhan, S.: Computational differential privacy. In: Halevi, S. (ed.) CRYPTO 2009. LNCS, vol. 5677, pp. 126–142. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-03356-8\\_8](https://doi.org/10.1007/978-3-642-03356-8_8)
24. Hua, J., Tang, A., Fang, Y., Shen, Z., Zhong, S.: Privacy-preserving utility verification of the data published by non-interactive differentially private mechanisms. *IEEE Trans. Inf. Forensics Secur.* **11**(10), 2298–2311 (2016)
25. Hamm, J., Cao, P., Belkin, M.: Learning privately from multiparty data. arXiv preprint [arXiv:1602.03552](https://arxiv.org/abs/1602.03552) (2016)
26. Li, J., Yan, H.Y., Liu, Z.L., Chen, X.F., Huang, X.Y., Wong, D.C.S.: Location-sharing systems with enhanced privacy in mobile online social networks. *IEEE Syst. J.* <https://doi.org/10.1109/JSYST.2015.2415835>
27. Foulds, J., Geumlek, J., Welling, M., Chaudhuri, K.: On the theory and practice of privacy-preserving Bayesian data analysis. arXiv preprint [arXiv:1603.07294](https://arxiv.org/abs/1603.07294) (2016)
28. Li, J., Huang, X.Y., Li, J.W., Chen, X.F., Xiang, Y.: Securely outsourcing attribute-based encryption with checkability. *IEEE Trans. Parallel Distrib. Syst.* **25**(8), 2201–2210 (2014)
29. Li, P., Li, J., Huang, Z.G., Li, T., Gao, C.Z., Yiu, S.M., Chen, K.: Multi-key privacy-preserving deep learning in cloud computing. *Future Gener. Comput. Syst.* (2017). <https://doi.org/10.1016/j.future.2017.02.006>

30. Li, P., Li, J., Huang, Z.G., Gao, C.Z., Chen, W.B., Chen, K.: Privacy-preserving outsourced classification in cloud computing. *Cluster Comput.*, 1–10 (2017). <https://doi.org/10.1007/s10586-017-0849-9>
31. Abadi, M., Chu, A., Goodfellow, I., McMahan, H.B., Mironov, I., Talwar, K., Zhang, L.: Deep learning with differential privacy. In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 308–318. ACM (2016)
32. Sweeney, L.:  $k$ -anonymity: a model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **10**(5), 557–570 (2002)
33. Sweeney, L.: Achieving  $k$ -anonymity privacy protection using generalization and suppression. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **10**(5), 571–588 (2002)
34. Li, N., Li, T., Venkatasubramanian, S.:  $t$ -closeness: privacy beyond  $k$ -anonymity and  $l$ -diversity. In: *IEEE 23rd International Conference on Data Engineering, ICDE 2007*, pp. 106–115. IEEE (2007)
35. Mohammed, N., Chen, R., Fung, B.C.M., Yu, P.S.: Differentially private data release for data mining. In: *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 493–501. ACM (2011)
36. Wang, Q., Zhang, Y., Lu, X., Wang, Z., Qin, Z., Ren, K.: RescueDP: real-time spatio-temporal crowd-sourced data publishing with differential privacy. In: *The 35th Annual IEEE International Conference on Computer Communications, IEEE INFOCOM 2016*, pp. 1–9. IEEE (2016)
37. Chen, R., Reznichenko, A., Francis, P., Gehrke, J.: Towards statistical queries over distributed private user data. In: *NSDI*, vol. 12, p. 13 (2012)
38. Chen, R., Fung, B.C.M., Mohammed, N., Desai, B.C., Wang, K.: Privacy-preserving trajectory data publishing by local suppression. *Inf. Sci.* **231**, 83–97 (2013)
39. Chen, R., Xiao, Q., Zhang, Y., Xu, J.: Differentially private high-dimensional data publication via sampling-based inference. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 129–138. ACM (2015)
40. Bassily, R., Smith, A., Thakurta, A.: Differentially private empirical risk minimization: efficient algorithms and tight error bounds. arXiv preprint [arXiv:1405.7085](https://arxiv.org/abs/1405.7085) (2014)
41. Goldwasser, S., Micali, S.: Probabilistic encryption & how to play mental poker keeping secret all partial information. In: *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing*, pp. 365–377. ACM (1982)
42. Goryczka, S., Xiong, L., Fung, B.C.M.:  $m$ -privacy for collaborative data publishing. *IEEE Trans. Knowl. Data Eng.* **26**(10), 2520–2533 (2014)
43. Goryczka, S., Xiong, L.: A comprehensive comparison of multi-party secure additions with differential privacy. *IEEE Trans. Dependable Secur. Comput.* **14**, 463–477 (2015)
44. Su, S., Tang, P., Cheng, X., Chen, R., Wu, Z.: Differentially private multi-party high-dimensional data publishing. In: *2016 IEEE 32nd International Conference on Data Engineering (ICDE)*, pp. 205–216. IEEE (2016)
45. Song, S., Chaudhuri, K., Sarwate, A.D.: Stochastic gradient descent with differentially private updates. In: *2013 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 245–248. IEEE (2013)
46. Youn, T.-Y., Park, Y.-H., Kim, C.H., Lim, J.: An efficient public key cryptosystem with a privacy enhanced double decryption mechanism. In: Preneel, B., Tavares, S. (eds.) *SAC 2005*. LNCS, vol. 3897, pp. 144–158. Springer, Heidelberg (2006). <https://doi.org/10.1007/11693383.10>

47. Wu, X., Kumar, A., Chaudhuri, K., Jha, S., Naughton, J.F.: Differentially private stochastic gradient descent for in-RDBMS analytics. arXiv preprint [arXiv:1606.04722](https://arxiv.org/abs/1606.04722) (2016)
48. Xiao, X., Tao, Y.:  $M$ -invariance: towards privacy preserving re-publication of dynamic datasets. In: Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, pp. 689–700. ACM (2007)
49. Huang, Z., Mitra, S., Dullerud, G.: Differentially private iterative synchronous consensus. In: Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society, pp. 81–90. ACM (2012)
50. Zhang, J., Zhang, Z., Xiao, X., Yang, Y., Winslett, M.: Functional mechanism: regression analysis under differential privacy. VLDB 5(11), 1364–1375 (2012)

# REW-SMT: A New Approach for Rewriting XACML Request with Dynamic Big Data Security Policies

Ha Xuan Son<sup>1</sup>(✉), Tran Khanh Dang<sup>1</sup>, and Fabio Massacci<sup>2</sup>

<sup>1</sup> Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam  
hxson@ctuet.edu.vn, khanh@hcmut.edu.vn

<sup>2</sup> University of Trento, Trento, Italy  
fabio.massacci@unitn.it

**Abstract.** Application of dynamic policy has brought benefits to distributed systems, cloud systems, and social network. However, there are no previous studies focused on solving authorization problems in the dynamic policy. In this paper, we focus on analyzing the way of policy change and providing solutions in the dynamic policy environment. The contribution of this paper is two-fold: including the solution for changing policy even when the access request has been granted by the policy and we provide an XACML-based implementation that incorporates the rewriting request model. Experiential results with real-world policies have established the practical and theoretical value of our newly introduced approach.

**Keywords:** XACML · Big data · SMT · Dynamic policy  
Rewriting request · Policy evaluation

## 1 Introduction

With the growing popularity of cloud computing and social networks, the risks of sensitive data breaches have increased due to the fact that more and more enterprises and individuals tend to store their private information on the Internet. The access control models have established for such systems and, in particular, to ensure that sensitive data can only be accessed by authorized users. Until now, several access control models have been proposed for the specification and enforcement of access control mechanisms such as Mandatory Access Controls (MAC), Discretionary Access Control (DAC), Role-based Access Control (RBAC) and most recently Attribute-based Access Control (ABAC) [4]. Among these models, the eXtensible Access Control Markup Language (XACML) is an approved Organization for the Advancement of Structured Information Standards (OASIS) [13]. The authorization in XACML based on combining policies possibly specified by independent authorities.

In conventional static policy environments, it is very difficult for applications to interwork with new services [3]. Moreover, the static policy has significant



limitations in the cloud system, distributed system, as commented on in the literature [5], respectively. Dunlop et al. [3] define a dynamic policy environment that would involve the following:

- Policy include a variable which containing reference to run-time or location.
- Policy is able to be altered during the request time.
- Ability to create, update or delete policy during the request time.
- Ability to dismiss assigned policy during the request time.

One of our key motivation in this paper is to support behavioral flexibility by changing policy without recoding or stopping the system [7, 14]. This implies that the policy should be possible to update the rules dynamically or conditions interpreted by distributed entities to modify their behavior. For examples, Mazurek et al. [10] defined the system, in which users can update their policy dynamically in response to access requests. Thus, how can we guarantee the authorization problem in the dynamic environment. However, we identify several fundament which have not addressed yet in the two approaches that are unsupported on dynamic update policies. Firstly, the model is unsupported both update and delete conditions in rules and policy(set). In the relationship between policies and rules, it is the lack of both combining algorithms and obligation correctly. Finally, they are used only for a special structure of the policy language in XACML due to the lack of supporting complex `Rule` and `Policy (Set)` structure. This paper presents a method using the definitions based on rewriting requests in XACML to satisfy the requirement of both security and without recoding or stopping the system [14]. We analyze two problems including policy change continuously during the request time even when the access request has been granted and re-evaluated when the policy which has granted to the request was deleted.

In this paper, we employ SMT solving for the policy evaluation processing of access requests in XACML. Our proposal follows a similar Turkmen et al. [18, 19] approach that happened in analyzing access requests evaluated to a certain decision. However, the contribution of this paper is two-fold: First, it presents dynamic adaptability of behavior by changing policy without recoding or stopping the system. Second, it provides an XACML-based implementation that incorporates the rewriting request model.

The rest of the paper is organized as follows. In the next section, we provide relevant background material and related work. Section 3 is devoted to the approach based on the encoding of XACML policies as SMT formulas. In the following sections, we demonstrate Rew-SMT model and our implementation. In Sect. 5, we evaluate our implementation. Finally, we conclude the paper with a summary of our results and discuss opportunities for further research.

## 2 Background and Related Works

### 2.1 Satisfiability Modulo Theories (SMT)

SMT [2] generalizes boolean satisfiability (SAT) by using theories of various data structures that includes equality reasoning, arithmetic, fixed-size bit-vectors,

arrays, quantifiers and the other useful first-order theories. SMT solver establishes the satisfiability of a formula in not only decidable fragments first-order logic (e.g. the BSR fragment) but also decidable theories (e.g., Linear Arithmetics). To specify SMT formulas, we follow an extended version of the SMT-LIB (v2) and standard based on first order logic<sup>1</sup>. On the other hand, SMT solvers enable applications in several areas, such as extended static checking, predicate abstraction, test case generation, and bounded model checking over infinite domains. In the following, we show how SMT solvers can be used to support the policy evaluation in XACML v3.0 standard; see [19] for formal definitions.

## 2.2 Policy Evaluation

XACML enjoys large adoption in industry due to its simplicity and its powerful extension mechanism. Most of the studies has been focused on modelling, analysis and testing of XACML policies [8, 16, 17]. Recent work solves the policy evaluation problem directly in different ways: Decision Diagrams (DDs) [9], Binary Decision Diagrams (BDDs) [12], Multi-data-types Interval Decision Diagrams (MIDD) [11]. Liu et al. [9] proposed XEngine, a scheme for efficient XACML policy evaluation. This approach aims at improving the performance of PDP by using decision diagrams in XACML policy evaluation. The limitation is only solves the part of XACML policy and the numericalization does not support comparison functions. Pina et al. used BDDs theories to build two trees: *Matching Tree* based on the binary search algorithm with an aim at the fast searching of applicable rules, and *Combining Tree* support evaluation of applicable rules in [12]. Ngo et al. [11] approach is similar to [12] in the spirit. However, the approach based on *Decision Diagrams* in the general support only simple comparison functions (eight *less than*, *greater than* or *equal*).

## 2.3 Dynamic Policy

Many previous studies have begun looking at dynamic policies/rules evaluation as opposed to static policies [1, 6, 15]. Ammar et al. [1] approach implicitly updates policy rules based on dynamically inferring a query classification. They do not dynamically update the original policy definitions, but implicitly incorporate context into rule the evaluation. Kabbani et al. [6] presented an approach for specifying and enforcing dynamic authorization policies based on situations. The solution of authors proposes to make a policy dynamic by modifying authorization rules when the conditions and circumstances are changed. However, the drawback is the rule management complexity when considering many situations and many rules. Son et al. introduced Rew-XAC model in [15], based on XACMLv3. The Rew-XAC model carry out the rewriting request by computing a fuzzy function in policy based on an access request. However, they are use only for a special class of the policy structures, which means the policy's structure has only one policy in a policy set and one rule in a policy. Moreover, Rew-XAC theory does not support combining algorithms and multiple decisions.

<sup>1</sup> <http://www.smtlib.org>.

### 3 Approach

#### 3.1 Scenario

In this section, we present the sample encoding of XACML policies in [15] into SMT formulas and we will use as a running example through the paper.

**Policy p:** *When a nurse is in the patient’s room within the time interval 8:00 Am - 8:00 PM, she is able to access the personal information of those patients whose age is greater than 16, address is in the south of Vietnam (i.e. Can Tho, Ho Chi Minh, Ca Mau), treatment department is the Heart Center and disease is the hypertension. One way to model this policy is to represent (the negation of) these constraints as **Deny** rules and then to combine the resulting rules using **deny overrides** (dov):*

$$P : \langle p[\text{com\_al\_ID}], r_1, r_2, \dots, r_n \rangle \quad (1)$$

where  $p[\text{com\_al\_ID}]$  is the ID of combining algorithms and  $r_1, r_2, \dots, r_n$  denoted from **Rule 1** to **Rule n**. Our approach is to analyze policies (3) to gain high-efficiency structure while still preserving XACML correctness features.

$p[\text{dov}]$  : *subject* : “nurse”  $\wedge$  *resource* : “patients’ record”  $\wedge$  *action* : “read”  
 $r_1[\text{Deny}]$  : *current – time* < 8 : 00AM  $\vee$  *current – time* > 8 : 00PM  
 $r_2[\text{Deny}]$  : *current – place*  $\neq$  Patient room  
 $r_3[\text{Deny}]$  : *address*  $\notin$  {Can Tho, Ho Chi Minh, Ca Mau}  
 $r_4[\text{Deny}]$  : *age* < 16  
 $r_5[\text{Deny}]$  : *disease*  $\neq$  Hypertension  
 $r_6[\text{Deny}]$  : *department*  $\neq$  Heart Center  
 $r_7[\text{Permit}]$  : **true**

#### 3.2 Policy Formalization

We use (3) to create applicable constraints, which can construct and aggregate logical expressions. The applicability constraints defined as below:

$ac_0$ : “nurse”  $\in$  **subject**  
 $ac_1$ : “patients’ record”  $\in$  **resource**  
 $ac_2$ : “read”  $\in$  **action**  
 $ac_3$ :  $\forall v \in$  **current – time**  $v > 8 : 00PM$   
 $ac_4$ :  $\forall v \in$  **current – time**  $v < 8 : 00AM$   
 $ac_5$ :  $\forall v \in$  **current – place**  $v \neq$  Patient room  
 $ac_6$ :  $\forall v \in$  {Can Tho, Ho Chi Minh, Ca Mau}  $v \neq$  **address**  
 $ac_7$ : **age**  $\leq 16$   
 $ac_8$ : **disease**  $\neq$  Hypertension  
 $ac_9$ : **department**  $\neq$  Heart Center  
 $ac_{10\dots 18}$ :  $\forall v \notin$  att

$att \in \{subject, resource, action, current - time, current - place, address, age, disease, department\}$ . Constraints  $ac_{10} \dots ac_{18}$  address Indeterminate cases if the access request is lack of any attributes in  $att$ .

Each policy has a target element that specifies *applicability constraint*. The target element is used to define their scope, it means that the attribute values are matched with the incoming request attribute value. Applicable constraints are used to divide the policy space into three *applicable space (AS)* i.e.  $\langle AS_{AS}, AS_{IN}, AS_{NA} \rangle$  that denote for *Applicable Space, Indeterminate Space* and *Not Applicable space* respectively. We represent the applicability space of a policy element is  $\langle AS_{AS}, AS_{IN} \rangle$  which means an access request  $req \in AS_{NA}$  iff  $req \notin AS_{AS} \cup AS_{IN}$ . We symbolize (3) as the target of rule  $\langle AS_{AS}, AS_{IN} \rangle$  and use above applicable constraints to construct a policy formalization:

$$\begin{array}{ll}
 r_1: \langle (ac_3 \cup ac_4), ac_{13} \rangle & r_5: \langle ac_8, ac_{17} \rangle \\
 r_2: \langle ac_5, ac_{14} \rangle & r_6: \langle ac_9, ac_{18} \rangle \\
 r_3: \langle ac_6, ac_{15} \rangle & r_7: \langle \mathbf{true}, \emptyset \rangle \\
 r_4: \langle ac_7, ac_{16} \rangle &
 \end{array}$$

We present in Table 1 a succinct syntax that is the verbose syntaxes used in the policy evaluation theorem which can recursively compute the decision space  $DS$  of rule or policy (set) as follows; see [19] for formal definitions:

**Table 1.** The syntax abstract of policy evaluation theorem

Key word	Meaning of key work
$r$	<i>Rule</i>
$p$	<i>Policy(Set)</i>
$T$	<i>Target</i>
$C$	<i>Condition</i>
$e$	<i>Effect</i>
$x$	$x \in \{\text{Permit, Deny}\}$
$ca$	<i>Combining Algorithm</i>
$DS_{ca}$	<b>applyCa</b> ( $ca, DS_{q_1}, \dots, DS_{q_n}$ )

We describe the decision space of a policy as a tuple  $\langle DS_P, DS_D, DS_{IN}, DS_{NA} \rangle$ . Additionally, if the access request does not fall in  $DS_P \cup DS_D \cup DS_{IN}$  then that access request falls in  $DS_{NA}$ . We can reduce and transform expression in (3) to a decision space of the rules as the tuple  $\langle DS_P^{r_i}, DS_D^{r_i}, DS_{IN}^{r_i}, DS_{NA}^{r_i} \rangle$  by using the recursive computation of the policy evaluation which is shown in Fig. 1:

$$\begin{array}{l}
 DS_x^r = \begin{cases} AS_{AS}^T \cap AS_{AS}^C & \text{if } e = x \\ \emptyset & \text{Otherwise} \end{cases} \\
 DS_{IN(e)}^r = \begin{cases} AS_{IN}^T \cap AS_{IN}^C & \text{if } e = x \\ \emptyset & \text{Otherwise} \end{cases} \\
 DS_{IN(PD)}^r = \emptyset \\
 DS_{NA}^r = AS_{NA}^T \cup (AS_{AS}^T \cap AS_{NA}^C) \\
 DS_x^p = (AS_{AS}^T \cap DS_x^{ca}) \\
 DS_{IN(e)}^p = [(AS_{AS}^T \cup AS_{IN}^T) \cap DS_{IN(e)}^{ca}] \cup (AS_{IN}^T \cap DS_x^{ca}) \\
 DS_{IN(PD)}^p = (AS_{AS}^T \cup AS_{IN}^T) \cap DS_{IN(PD)}^{ca} \\
 DS_{NA}^p = AS_{NA}^T \cap DS_{NA}^{ca}
 \end{array}$$

**Fig. 1.** The theorem of policy evaluation

- $r_1: \langle \emptyset, (ac_3 \cup ac_4) \cap \overline{ac_{13}}, ac_{13}, \overline{(ac_3 \cup ac_4)} \cap \overline{ac_{13}} \rangle$
- $r_2: \langle \emptyset, ac_5 \cap \overline{ac_{14}}, ac_{14}, \overline{ac_5} \cap \overline{ac_{14}} \rangle$
- $r_3: \langle \emptyset, ac_6 \cap \overline{ac_{15}}, ac_{15}, \overline{ac_6} \cap \overline{ac_{15}} \rangle$
- $r_4: \langle \emptyset, ac_7 \cap \overline{ac_{16}}, ac_{16}, \overline{ac_7} \cap \overline{ac_{16}} \rangle$
- $r_5: \langle \emptyset, ac_8 \cap \overline{ac_{17}}, ac_{17}, \overline{ac_8} \cap \overline{ac_{17}} \rangle$
- $r_6: \langle \emptyset, ac_9 \cap \overline{ac_{18}}, ac_{18}, \overline{ac_9} \cap \overline{ac_{18}} \rangle$
- $r_7: \langle \mathbf{true}, \emptyset, \emptyset, \emptyset \rangle$

Turkmen et al. [19] demonstrated the encoding of *Decision Space (DS)* for the combining algorithm of the policy  $p$  which combine two policies  $p_1$  and  $p_2$ . The **deny-override combining algorithm** can be defined as follows:

$$DS_D^p = DS_D^{p_1} \cup DS_D^{p_2} \tag{2}$$

$$DS_P^p = (DS_P^{p_1} \cup DS_P^{p_2}) \setminus (DS_D^p \cup DS_{INPD}^p \cup DS_{IND}^p \cup DS_P^p) \tag{3}$$

$$DS_{NA}^p = DS_{NA}^{p_1} \cap DS_{NA}^{p_2} \tag{4}$$

$$DS_{INP}^p = (DS_P^{p_1} \cup DS_P^{p_2}) \setminus (DS_D^p \cup DS_{INPD}^p \cup DS_{IND}^p \cup DS_P^p) \tag{5}$$

$$DS_{IND}^p = (DS_D^{p_1} \cup DS_D^{p_2}) \setminus (DS_D^p \cup DS_{INPD}^p) \tag{6}$$

$$DS_{INPD}^p = \{(DS_{INPD}^{p_1} \cup DS_{INPD}^{p_2}) \cup \alpha \cup \beta\} \setminus DS_D^p \tag{7}$$

where:

$$\alpha = [DS_{IND}^{p_1} \cup (DS_{INP}^{p_2} \cup DS_P^{p_2})] \quad \beta = [DS_{IND}^{p_2} \cup (DS_{INP}^{p_1} \cup DS_P^{p_1})]$$

The overall *Indeterminate* space can be obtained by union of three sub-space, namely  $IN_{PD}$ ,  $IN_P$ , and  $IN_D$ .

$$DS_{IN}^p = DS_{INPD}^p \cup DS_{INP}^p \cup DS_{IND}^p$$

The next step describes the mechanisms to create a decision space of the overall policy as a tuple  $\langle DS_P, DS_D, DS_{IN}, DS_{NA} \rangle$  from the XACML logical expression by the rules in (3).

$$\begin{aligned}
 DS_D^p &= \mathbf{A} \cap \mathbf{B} & DS_{IN}^p &= \mathbf{A} \cap \mathbf{C} \cap \bar{\mathbf{B}} \\
 DS_P^p &= \mathbf{A} \cap \bar{\mathbf{B}} & DS_{NA}^p &= \bar{\mathbf{A}}
 \end{aligned}$$

where:

$$\begin{aligned}
 \mathbf{A} &= (ac_0 \cap ac_1 \cap ac_2) \\
 \mathbf{B} &= (ac_3 \cup ac_4 \cup ac_5 \cup ac_6 \cup ac_7 \cup ac_8 \cup ac_9) \\
 \mathbf{C} &= (ac_{10} \cup ac_{11} \cup ac_{12} \cup ac_{13} \cup ac_{14} \cup ac_{15} \cup ac_{16} \cup ac_{17} \cup ac_{18})
 \end{aligned}$$

After that, we transform this decision space to the policy as SMT formulas containing  $\{F_P, F_D, F_{IN}, F_{NA}\}$  using a background theory  $\mathbf{T}$ . Where  $F_P, F_D, F_{IN}, F_{NA}$  “are much sorted first-order formulas encoding Permit, Deny, Indeterminate decision spaces of  $p$  respectively with some of their terms interpreted in  $\mathbf{T}$ ” [19]. In the detail, the decision space of the policy builds up from the constraints, moreover, the definition of constraint based on arithmetic and numerical comparison functions requires the theory of linear arithmetic. The conditions of rule/policy(set) translates to applicability constraints ( $ac_i$ ) which defined by using the theory of string (e.g.  $\mathbf{disease} \neq \mathbf{Hypertension}$ ), bag (e.g.  $\forall v \in \{\mathbf{CanTho}, \mathbf{HoChiMinh}, \mathbf{CaMau}\}$ ) and set (e.g.  $\mathbf{age} \leq 16$ ). Generally, the background theory is the way of encoding the applicability constraints built up from combining of different theories. For instance, in this paper we used three theories (i.e. string, array and set) to construct the background theory.

### 3.3 Encoding XACML Request

This section symbolizes an access request as a set of general attribute and use the above policy formalization as to construct a tuple request policy  $\tau$  from its value. The tuple  $\tau$  of request  $Q$  is defined as follows:

$$\text{Tuple } \tau = \langle Q, \text{List}\{p_1, p_2, \dots, p_n\} \rangle \quad (8)$$

where:

$$\begin{aligned}
 Q &: \langle att_1 = v_1, att_2 = v_2, \dots, att_n = v_n \rangle \\
 \text{List}\{p_1, p_2, \dots, p_n\} &= \begin{cases} p_1 = \langle F_D^{r_1}, F_P^{r_1}, F_{IN}^{r_1}, F_{NA}^{r_1} \rangle \\ p_2 = \langle F_D^{r_2}, F_P^{r_2}, F_{IN}^{r_2}, F_{NA}^{r_2} \rangle \\ \dots \\ p_n = \langle F_D^{r_n}, F_P^{r_n}, F_{IN}^{r_n}, F_{NA}^{r_n} \rangle \end{cases}
 \end{aligned}$$

Policy evaluation is the process to find an applicable policy and grant permission to an access request based on the comparison between the logical expression in rules and the value of attributes ( $att$ ) in the access request. The request  $Q$  is followed by a policy only if all  $att$  of  $Q$  matched with logical expressions  $X_i$  of that policy. The request below is best described as an example of rewriting request when the policy had updated.

**Request Q:** *At 8:00 AM, a nurse who is in one patient’s room requests to read the personal information of the patient whose age is greater than 50, the address is in Ho Chi Minh, and disease is hypertension [15].*

**Q:**  $\langle$ **subject** = “nurse”, **current-time** = 8:00 AM, **current-place** = patients’ room, **action** = read, **resource** = patients’ record, **age** > 50, **address** = “Ho Chi Minh”, **disease** = “Hypertension” $\rangle$ .

**Table 2.** Evaluation process

	Policy	Request	Response
Subject	Nurse	Nurse	True
Resource	Patients’ record	Patients’ record	True
Action	Read	Read	True
Current - time	$v < 8:00 \text{ AM}$ $v > 8:00 \text{ PM}$	8:00 AM	False
Current - place	$v \neq \text{Patients’ room}$	Patients’ room	False
Address	$v \neq \{\text{Can Tho, Ho Chi Minh, Ca Mau}\}$	Ho Chi Minh	False
Age	$v \leq 16$	$v > 50$	False
Disease	$v \neq \text{Hypertension}$	Hypertension	False
Department	$v \neq \text{Heart Center}$	<b>IN</b>	<b>IN</b>

Table 2 summarizes the evaluation process of the request and the policy in scenario section. The information from the second row to the fourth one presents the response of the comparison in *Subject*, *Resource*, *Action*. The values can be returned including **Permit**, **Deny** or **Indeterminate**. On the one hand, the value of the **Department** is returned **IN** (**IN** denoted **Indeterminate**) and the **Effect** element of this policy is “*Deny*”, the result is **Indeterminate Deny**). After that, these responses send to PEP in which an **obligation** will be generated. The **obligation** requires user or system to provide the value of **Department**. According to the input value, the request will be returned to “**Permit**” if this value is *Heart center* or “**Deny**” otherwise. Assign **M** as the input value after fulfilling the requirement of **Obligation service**, the new request as follow:

**Q’:**  $\langle$ **subject** = “nurse”, **current-time** = 8:00 AM, **current-place** = patients’ room, **action** = read, **resource** = patients’ record, **age** > 50, **address** = “Ho Chi Minh”, **disease** = “Hypertension”, **department** = “**M**” $\rangle$ .

### 3.4 Change Policy Definition

As introduced before, our purpose is to support dynamic adaptability by changing the policy while still preserving security. We analyze in two aspects, i.e. the policy change continuously and the deleted policy. In this section, we define the policy change and classify the changeability of policy thanks to the level of the

interaction between the policy and their protected. In the former case, we find the other once in **Policy Administrator Point (PAP)** to replace the original policy. In the latter case, based on the property of changing, we classify the change policy into 3 cases, namely *delete*, *insert* and *update* in Table 3:

**Table 3.** The cases of inside policy changing

	Delete	Insert	Update
Explain	Delete applicable constraints	Insert applicable constraints	Update applicable-constraints, effect, combining algorithm
Re-evaluated	No	Yes	Yes
Rewrite	No	Yes	No

- **Delete:** If the systems delete any applicable constraints, the database protected by these rules/policy(set) are increased, this means we do not need to re-evaluated or rewrite the original request and go to next step.
- **Insert:** If the systems insert one or more applicable constraints, the database protected by these rules/policy(set) are decreased, this means we should rewrite the original request and re-evaluate before going to next step.
- **Update:** If the systems change any conditions, effect, or combining algorithm; the database protected by these rules/policy(set) are changed, this means we should be re-evaluated before going to the next step. In this cases, we do not need to rewrite the original request.

## 4 Implementation

### 4.1 Rew-SMT Model

In this section, we describe the Rew-SMT model and their work-flow, Fig. 2 illustrates our design in more detail. Although, our model is similar to the XACMLv3.0 to achieve better performance we replace the original evaluate process by the approach presented in [19] SMT technique. In particular, we focus on solving the authorization in dynamic policy environment by two-stage analyzing that is before and after updating policy. The architecture of Rew-SMT model is shown in Fig. 2 which implements authorization in eight steps:

1. PAP (Policy Administrator Point) writes policies and policy sets. After that, PAP denotes the decision space of the policy as a tuple  $\langle DS_P^{p_i}, DS_D^{p_i}, DS_{IN}^{p_i}, DS_{NA}^{p_i} \rangle$  and makes them available to PDP (Policy Decision Point).
2. Clients send an access request to PEP (Policy Enforcement Point). The access request will be transferred an attribute assignment by providing the values for those attributes  $\langle att_1^{Q_i} = v_1, att_2^{Q_i} = v_2, \dots, att_n^{Q_i} = v_n \rangle$ .



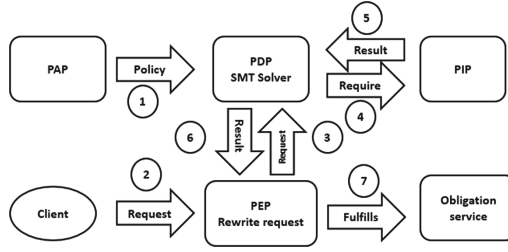


Fig. 2. Rew-SMT model

3. PEP sends the request for accessing to PDP in requests format.
4. PDP sends the request attributes to PIP (Policy Information Point).
5. PIP obtains the request attributes. PIP gets a value from the database then returns the requested attributes to the PDP component.
6. PDP evaluates the request and policies based on the SMT solver. Before PDP returns a response to PEP, it translates those responses into  $\langle DS_P^{p_i} = \{True/False\}, DS_D^{p_i} = \{True/False\}, DS_{IN}^{p_i} = \{True/False\}, DS_{NA}^{p_i} = \{True/False\} \rangle$ . According to those responses, the system could (not) grant a policy to the access request.
7. Depending on the update policies or not, we have two cases in this step.
  - PEP rewrites a request if the case is “insert condition” or PDP re-evaluates the requests when “delete policy” and “update condition” cases.
  - Otherwise, PEP fulfills Obligations.
8. If the value of the response is **Permit**, then the PEP accesses to the resource, otherwise, it returns **Deny** to the access requester.

## 4.2 Query Rewriting Algorithm

In this section, we describe some functions used that help us to better understanding the process of building Rew-SMT mechanism.

The Algorithm 1 finds the policies that returned “*Indeterminate*” or “*Permit/Deny*” from a *List\_policy* in the evaluation process. If the method returns a “*Permit/Deny*” policy, the access requests could (not) be granted by this policy, respectively. However, if the method returns the list of “*Indeterminate*” policy, the obligation will be generated and sends to users/systems. After that, based on the input data, the system could (not) grant a policy to the access request.

**translatorPolicy**( $p$ ): which parses the policy and inserts them into a tuple  $\langle DS_P^p, DS_D^p, DS_{IN}^p, DS_{NA}^p \rangle$  in line 1.

**evaluate**( $Q, \{p_1, p_2, \dots, p_n\}$ ): call evaluation process in line 3.

**isPermit/Deny/IN**( $X_{p_i}$ ): the input of these functions is the result of the evaluation functions. The value of these functions is **true** if the corresponding element of tuple  $X_{p_i}$ , *Permit*, *Deny*, *Indeterminate* is **true**. After that, we classify the evaluation scope to 3 sub-scope. Firstly, if the value of **isPermit**() or **isDeny**() in line 5 is **true**, the method stop the loop and return the policy in line 7. Otherwise, in line 9 and 10. The system similarly analyzes next policy in line 12.

**Algorithm 1.** findReplacePolicy() function

---

```

0: Input:  $List\_policy \{p_1, p_2, \dots, p_n\}$   $Q: \langle att_1 = v_1, att_2 = v_2, \dots, att_n = v_n \rangle$ 
   Output: Policy  $\{p_x \dots, p_{x+m}\}$ 
1:  $\langle DS_P^{p_i}, DS_D^{p_i}, DS_{IN}^{p_i}, DS_{NA}^{p_i} \rangle \leftarrow \text{translatorPolicy}(p_i)$ 
2: policyID = {}
3:  $X = \text{evaluate}(Q, List\_policy)$ 
4: for  $p_1, p_2, \dots, p_n$  do
5:   if  $p_i.isPermit(X_{p_i}) == \text{true}$  or  $p_i.isDeny(X_{p_i}) == \text{true}$  then
6:     policyID.add(i)
7:     return getPolicyByID(policyID)
8:   end if
9:   if  $p_i.isIN(\text{evaluate}(X_{p_i})) == \text{true}$  then
10:    policyID.add(i)
11:   end if
12:   i ++
13: end for
14: return getPolicyByID(policyID)

```

---

**getPolicyByID(ID):** get policy from the ID of policy. if the value of *policyID* is 0, return as **null** in line 14.

**Algorithm 2.** policyChange() function

---

```

0: Input:  $p: \langle DS_P^p, DS_D^p, DS_{IN}^p, DS_{NA}^p \rangle$   $Q: \langle att_1 = v_1, att_2 = v_2, \dots, att_n = v_n \rangle$ 
   Output:  $Q$ 
1: temp = getTypeChangePolicy(p), request  $Q^*$ 
2: while temp do
3:   case 1: Delete
4:     return  $Q$ 
5:   case 2: Update
6:     evaluate( $Q, p$ )
7:     return  $Q$ 
8:   case 3: Insert
9:     tempAtt = getLackAtt( $Q^*, p$ );
10:     $v = \text{Obligation}(\text{tempAtt})$ ;
11:     $Q^* = \text{rewriteRequest}(Q, v)$ 
12:    evaluate( $Q^*, p$ )
13:   return  $Q^*$ 
14: end while

```

---

The Algorithm 2 classifies a policy change from their property (*Delete*, *Update* or *Insert*) The input of this method is a policy and a request. In line 1, *temp* is a type of change. If the value of *temp* is “Delete”, the method returns the request  $Q$  in line 3 and 4. If the type of change is “Update” we re-evaluate before returning the request  $Q$  between line 5 and 7. In another cases,  $\text{getLackAtt}(Q^*, p)$  returns the attributes if it does not exist in the request. Otherwise, function returns null.

The input values are stored in  $v$  after fulfilling the requirement of *Obligation* method in line 10. We rewrite a new request by combining the value stored in  $v$  with the original request  $Q$ . After that we could evaluate the new request with the policy  $p$  before return  $Q^*$  from line 11 to 13.

**getTypeChangePolicy( $p$ )**: the output of this function returns type of change policy (i.e. Delete, Update, Insert).

**getLackAtt( $Q^*$ ,  $p$ )**: the output of this method returns the attribute when an access request matches to the policy.

**Obligation( $att$ )**: get data from user or system.

**rewriteRequest( $Q, v$ )**: create the new request from new attribute and their value with the original request.

## 5 Experiments

**Environment:** We run experiment on JRE 1.8, a Windows 7 Ultimate system with Intel i7 core 2.2 GHz and 6 GB RAM. The experiment datasets are XACML 3.0 policies. We use three samples real-world policy i.e. Continue-a<sup>2</sup>, KMarket<sup>3</sup> and GEYSER<sup>4</sup> shown in Table 4.

**Table 4.** Sample real-world policies

Datasets	Policy levels	# Policy set	# Policies	# Rules	# Attributes	Operators
Continue-a	6	111	266	298	14	=
KMarket	2	1	3	12	3	= (58.8%)- higher than (41.2%)
GEYSER	3	6	7	33	3	=

**Performance analysis:** In our experiments, we compare our implementation with normal algorithm. In the normal algorithm, in which re-evaluate the request if PAP have any update policy. This section has two scenario include policy delete and inside policy changing.

In our first testbed, we compare rewriting algorithm with normal one. We randomly generate 1000 requests for each dataset and report average run time. Figure 3 shows the performance of a thousand requests for rewriting algorithm and normal one. In continue-a, the performance time of rewriting algorithm higher than that of normal algorithm due to Rew-SMT spends time not only for policies conversion, request conversion and evaluation but also for rewriting the access requests when Rew-SMT could not find the appropriate policy. On the

<sup>2</sup> Continue-a: the policy is taken and converted by [11].

<sup>3</sup> KMarket: a real-world policy taken from Balana in 2013.

<sup>4</sup> GEYSER: <http://www.geysers.eu/>.

other hand, if each policy of the policy set has simple structure and few attributes then using the rewriting algorithm will give better performance than the normal algorithm which is demonstrated in GEYSERS and KMarket experiment.

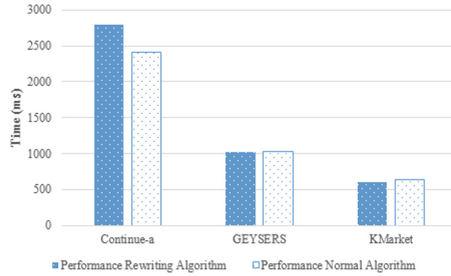


Fig. 3. Comparison between rewriting algorithm and normal algorithm

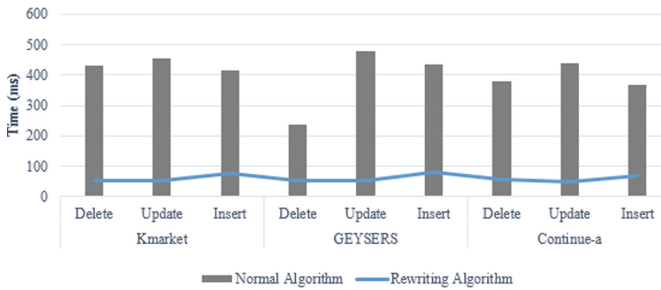


Fig. 4. Comparing normal and rewriting algorithm

In our second testbed, we randomly generate 100 couple of request-policy for each case (delete, insert, and update) and report the performance time of two algorithm in Fig. 4. The performance of rewriting algorithm is better than that of the normal algorithm due to the fact that Rew-SMT does not spend time for policies conversion and it just evaluates the access request on one policy which can be update.

## 6 Conclusion

In this paper, we introduced a solution to deal with the problem of security dynamic policy in big data repository. The proposed solution has four elements: a translator translated from policy to logic expressions are defined as applicable constraints, re-evaluate request (delete policy and update applicable constraints), rewriting request (insert applicable constraints), and obligation. We

define the Rew-SMT model which can be used both in XACML and in dynamic policy system. Our approach solves the problem in two aspects including delete policy and change inside policies. Our solution implemented in the Rew-SMT protocol that has both the efficient evaluation of structural complexity and results in three real-life policies. In the experiment we compare our approach with the normal scenario, it shows that our approach is higher performance than normal scenario.

When PEP sends the request to fulfill the obligations, it is likely, as we have generated the obligation, that we will send the requirement to user or system and ignore the redundancy and conflict between them. Thus an interesting area for future work is to prevent the conflict and redundancy of obligation. On the other hand, a future implementation may extend Rew-SMT and the policy analysis to cover complexity structure and conflict avoidance. These are all open issues that we hope to address in the near future.

**Acknowledgements.** This work was partially funded by the project of Ho Chi Minh City University of Technology under the contract number TSH-KHMT-2016-24 and was supported by AC-Lab (HCMUT) DISI-Labs (UNITN). Sincerely thank to Ngo Chan Nam, Lam Tuan Anh and Tran Luong Khiem who provided feedback on early revisions.

## References

1. Ammar, N., et al.: XACML policy evaluation with dynamic context handling. *IEEE Trans. Knowl. Data Eng.* **27**, 2575–2588 (2015)
2. Barrett, C.W., Sebastiani, R., Seshia, S.A., Tinelli, C.: Satisfiability modulo theories. *Handb. Satisf.* **185**, 825–885 (2009)
3. Dunlop, N., et al.: Dynamic policy model for large evolving enterprises. In: *Enterprise Distributed Object Computing Conference*, pp. 193–197. IEEE (2001)
4. Hu, V.C., et al.: Guide to attribute based access control (ABAC) definition and considerations (draft). NIST Special Publication 800-162 (2013)
5. Jaiswal, C., Nath, M., Kumar, V.: Location-based security framework for cloud perimeters. *IEEE Cloud Comput.* **1**(3), 56–64 (2014)
6. Kabbani, B., et al.: Specification and enforcement of dynamic authorization policies oriented by situations. In: *New Technologies, Mobility and Security*, pp. 1–6 (2014)
7. Laborde, R., et al.: An adaptive XACMLv3 policy enforcement point. In: *Computer Software and Applications Conference*, pp. 620–625. IEEE (2014)
8. Le Thi, K.T., Dang, T.K., Kuonen, P., Drissi, H.C.: STRoBAC – spatial temporal role based access control. In: Nguyen, N.-T., Hoang, K., Jędrzejowicz, P. (eds.) *ICCCI 2012*. LNCS (LNAI), vol. 7654, pp. 201–211. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-34707-8\\_21](https://doi.org/10.1007/978-3-642-34707-8_21)
9. Liu, A.X., et al.: Xengine: a fast and scalable XACML policy evaluation engine. In: *ACM SIGMETRICS Performance Evaluation Review*, no. 1, pp. 265–276 (2008)
10. Mazurek, M.L., et al.: Exploring reactive access control. In: *Conference on Human Factors in Computing Systems*, pp. 2085–2094. ACM (2011)
11. Ngo, C., Makkes, M.X., et al.: Multi-data-types interval decision diagrams for XACML evaluation engine. In: *Privacy, Security and Trust*, pp. 257–266. IEEE (2013)

12. Pina Ros, S., Lischka, M., Gómez Mármol, F.: Graph-based XACML evaluation. In: Proceedings of the 17th ACM symposium on Access Control Models and Technologies, pp. 83–92. ACM (2012)
13. Rissanen, E.: Extensible access control markup language (XACML) version 3.0 (2013)
14. Sloman, M., Lupu, E.: Security and management policy specification. *IEEE Netw.* **16**(2), 10–19 (2002)
15. Son, H.X., Tran, L.K., Dang, T.K., Pham, Y.N.: Rew-XAC: an approach to rewriting request for elastic ABAC enforcement with dynamic policies. In: *Advanced Computing and Applications*, pp. 25–31. IEEE (2016)
16. Thi, Q.N.T., Dang, T.K.: X-STROWL: a generalized extension of XACML for context-aware spatio-temporal RBAC model with OWL. In: *Digital Information Management*, pp. 253–258. IEEE (2012)
17. Thi, Q.N.T., Si, T.T., Dang, T.K.: Fine grained attribute based access control model for privacy protection. In: Dang, T.K., Wagner, R., Küng, J., Thoai, N., Takizawa, M., Neuhold, E. (eds.) *FDSE 2016. LNCS*, vol. 10018, pp. 305–316. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-48057-2\\_21](https://doi.org/10.1007/978-3-319-48057-2_21)
18. Turkmen, F., Demchenko, Y.: On the use of SMT solving for XACML policy evaluation. In: *Cloud Computing Technology and Science*, pp. 539–544. IEEE (2016)
19. Turkmen, F., den Hartog, J., Ranise, S., Zannone, N.: Formal analysis of XACML policies using SMT. *Comput. Secur.* **66**, 185–203 (2017)

# Decoupling Security Services from IaaS Cloud Through Remote Virtual Machine Introspection

Huaizhe Zhou<sup>1</sup>(✉), Haihe Ba<sup>1</sup>, Jiangchun Ren<sup>1</sup>, Yongjun Wang<sup>1</sup>,  
Zhiying Wang<sup>1</sup>, and Yunshi Li<sup>2</sup>

<sup>1</sup> College of Computer, National University of Defense Technology,  
Changsha 410073, China

{huaizhezhou,haiheba,jiangchunren,yongjunwang,zhiyingwang}@nudt.edu.cn

<sup>2</sup> Northwest Institute of Eco-Environment and Resources,  
CAS, Lanzhou 73000, China  
liyunshi@lzb.ac.cn

**Abstract.** Security and privacy concern is still one of the major issues that prevent users from moving to public clouds. Introduction of security services based on virtual machine introspection into cloud does not relieve this situation, because these services are inflexible and untrusted by tenants. The root cause of the problem is that the cloud administrator has more privilege over the security services, which leaves no options for the tenants to protect their virtual machines. In this paper, we propose a technique to decouple security services from cloud platform via remote virtual machine introspection. It enables remote trusted managed security services to protect tenants' virtual machines stealthily. We have implemented a proof-of-concept prototype with Xen hypervisor, called SE-Cloud. With the separation of introspection and security-business code, the security services can not be abused by administrators and have little impact on the management virtual machine. Our preliminary experimental results show that SE-Cloud can provide more robust and flexible protections for tenant virtual machines with acceptable overhead.

**Keywords:** IaaS cloud · Managed security services  
Remote virtual machine introspection · Stealthy monitoring  
Security service management

## 1 Introduction

With the proliferation and popularity of cloud computing, security and privacy threats have been unfortunately emerging in endlessly. Infrastructure-as-a-Service (IaaS), one of the dominating service models, acts fundamentally in the cloud and risk challenges in IaaS layer can exercise a great influence on the rest. Tenants hesitate to move to cloud computing environments for the risk of security incident. It is urgent to provide tenants a practical method to protect their systems in the cloud.

One possible solution (seemingly straightforward) may be to deploy security services (e.g., IDS, anti-virus) by tenants in each of their VMs. Nevertheless, self-building security protection by tenants is inefficient for the physical machines beyond their control. The deployed security services are susceptible to be manipulated and controlled by sophisticated adversaries. Baliga et al. [3] demonstrated how easy to launch such an attack by manipulating Linux Netfilter to remove hook functions to packet filtering. Similarly, some hackers can disarm Windows Firewall by halting particular running system services normally. Further, such services inevitably push the service deployment and maintenance burden back to the tenants who have to deploy and manage the same set of services for every single VM.

Most of the previous works on securing VMs in the cloud tried to leverage Virtual Machine Introspection [8,9,12,14]. It enables security services to monitor and inspect guest VMs outside and makes security services more robust. Unfortunately, in modern cloud computing platforms like Xen<sup>1</sup>, KVM<sup>2</sup>, and VMware<sup>3</sup>, the Cloud Service Provider (CSP) owns the Virtual Machine Monitor (VMM) and the administrative VM (i.e. Dom0 in Xen). As a result, the adoption of VMI-based security services, which must have privilege to access VMM and Dom0, relies heavily on the willingness of cloud service providers to deploy them.

Even though the cloud platform could provide such services, there are still some problems that remain to be resolved:

- (1) VMI-based security services are always deployed in the administrative VM [8], and they may have resource competition with necessary management tasks, causing nontrivial performance impact. Additional codes of these services will also impose potential attack surface to intruders.
- (2) Tenants distrust such services for these privilege could be misused by administrators. The malicious or just curious administrators could inspect virtual memory and sensitive data and code may be leaked.
- (3) There are multiple VMs on a cloud platform and an “one size fits all” security service is unacceptable to all tenants. For example, an IDS service that checks network packets for malicious content using simple signatures in a VM is not applicable to another VM that receives encrypted packets [5]. However, tenants have little choice in the deployment or configuration of these services.

In this paper, we propose decoupling security services from IaaS cloud through remote VMI to address these problems. The proposed technique enables security service to deploy in a trusted place which beyond cloud provider’s control, and protect tenant VM on the target cloud platform. It leverages remote VMI to actively monitor VMs’ running states in a dedicated VM called DomS which is accessible to the security services. The integrity of the hypervisor and

<sup>1</sup> XEN: <http://www.xen.org>.

<sup>2</sup> KVM: [http://www.linux-kvm.org/page/Main\\_Page](http://www.linux-kvm.org/page/Main_Page).

<sup>3</sup> VMware: <http://www.vmware.com/>.



DomS can be guaranteed by various existing techniques. In our proposed technique, the security services are provided by a trusted Managed Security Services Provider (MSSP), with which tenants are able to custom individual security services depend on their requirements. We have implemented a proof-of-concept prototype of proposed technique with Xen hypervisor, called SE-Cloud. Our preliminary evaluation shows that the proposed technique is able to support security service to protect tenant VMs in the cloud. It will provide more robust and flexible protection for the tenants with acceptable performance impact.

To summarize, this paper makes the following contributions:

- We introduce a new approach for tenant VMs protection in the cloud, in which the cloud platform cooperates with the trusted Managed Security Service Providers to provide tenants more robust and flexible security protection in a cost-effective way.
- We present a novel framework to decouple security service from the IaaS cloud platform by utilizing virtual machine introspection. The proposed framework enhances the efficacy and trustworthiness of security services for they are difficult to be evaded by skilled attackers or misused by administrators.
- We present the prototype implementation of the framework with Xen hypervisor. Our evaluation demonstrates the efficacy of proposed framework.

## 2 Motivation and Background

Virtual Machine Introspection (VMI) [8] enables outside monitoring for guest VMs, with which we can move the security stack from VMs into regions protected by the hypervisor. But VMI has an inherent problem, semantic gap, which has been a main motivation for a significant portion of research over the last decade [7]. VMI relies on reconstructing high-level state information from low level data, such as the memory, virtual disk and vCPU registers. However, with recent advances in forensics tools, the semantic gap problem can be considered a solved engineering problem [11]. It is a powerful technique allowing the security related software to inspect VMs and their executions without guest interference or enforced guest VM cooperation. VMI has been used to support a wide range of use cases including: touchless resource usage tracking [18], intrusion detection [9], malware analysis [14] and etc.

VMI also enables isolating the security solution from other server workloads, by deploying the security solution in a dedicated VM that has the privilege to access the hardware through the hypervisor. This makes it much harder for hackers to detect the installed security software. With the proliferation of kernel mode rootkits, this elevated protection has become ever more important.

With the ubiquitous and convenient network, managed security services have become more and more popular to be provided in a more technically and economically flexible way called *Security as a Service* [19]. Such a model not only provides organisations of all sizes with access to the technology services that they need, it can also be a much more cost-effective way of accessing services than

performing functions in-house. By outsourcing security protections to a security service provider, users can gain consistent and cost-effective protection regardless of device types, users locations or operating systems [4, 10]. There are also some security services for the cloud platform, such as Intelligent Protection [6]. They could enhance cloud users' experience by offering more secure, flexible, automated security management for applications deployed on cloud infrastructures. These security services provide tenants the ability of monitoring the health and protection of their virtual machines. They can analyze their virtual networks, memory and applications for vulnerabilities; continuously monitor for attacks, intrusions, viruses, and application integrity. The tenants can subscribe different security services based on their system security requirements.

This paper aims at leveraging remote virtual machine introspection to support managed security service into the IaaS platform to enable more flexible and robust protection for tenant VMs. In this way can we decouple security services from cloud infrastructure to make them more effective. In order to achieve this goal, we need to address the following challenges:

- **Challenges in data collection.** In order to support various security services from different vendors, we need to collect adequate information of VMs runtime state, including active processes, system call, opening files, register values and so forth. It is necessary to get such data effectively and safely from VMs in time. It is crucial to security service.
- **Challenges in security service management.** In our proposal, security service protects tenant VMs without any agent assistant. We should make sure that the security service could get specified VM of its users correctly. The framework of our proposal should manage all the security service and coordinate with these services.
- **Challenges in minimizing overhead.** By using multiple different VMI-based profiling tools to monitor and collect guest VM state measurement can have non-negligible performance overhead. An important question to be addressed is how to extend VMI to minimize the performance overhead of supporting multiple security services from different vendors.

### 3 Decoupling of Security Services

#### 3.1 Threat Model and Assumptions

Similar to the adversary model assumed in SSC, we distinguish *cloud service provider* from *cloud administrators*. We assume that cloud services providers are trusted for reputation reason. They are willing to protect tenants' security and privacy and have no motivation to launch any attacks. Therefore we trust hardware and physical attacks or protection against hardware attacks are out of the scope of this paper. We assume that the physical hardware is equipped with a Trusted Platform Module (TPM) chip, using which the integrity of the hypervisor is guaranteed by various techniques [20].

On the other hand, we assume cloud administrators may be untrusted. The privilege of the management VM can be abused by malicious or honest but curious administrators. In our proposed technique, the trusted managed security services providers act as Trusted Third Party and are chosen by tenants.

### 3.2 Design Goals

To decouple security services from cloud and to provide tenant VMs more robust protection in a cost-effective way, we combine managed security services with remote virtual machine introspection. Our design principles of the proposed technique are:

- **Transparency and Tamper resistance of MSS.** It is important to make the MSS transparent when monitoring and analyzing malware, since malicious software often tries to detect such services to conceal its real purpose and behavior. In fact, the system and applications in it may not be aware of the presence of the security service. What's more, MSS should continue to function reliably even if an attacker gains entry into the monitored system. The privileged attacker can not disable or evade security service to protect VM.
- **Tenant-Configured Security Services.** Our technique provides tenants flexible security protection mechanism over their VMs in the cloud. They can decide protection level based on their requirements by subscribing related security services. This makes security protection flexibility for both tenants and cloud service providers.

### 3.3 Overview

Figure 1 illustrates the system architecture for decoupling security services from cloud. We use Xen hypervisor to illustrate our designed architecture. This architecture consists of three parts, including the cloud service node, the managed security services provider and the tenant.

**The tenant.** The tenant is the end-user of this framework. He rents VMs in the cloud and subscribes security services from the security service vendor to protect his VMs. He specifies his security requirements and submits them to security service vendor along with basic information of cloud services provider. He will also need to inform his cloud service provider of the security service vendor information he has subscribed. The tenant can get security state of his VMs from security service vendor after they successfully deploy corresponding service on the specified cloud platform.

**Trusted Managed Security Services Provider (MSSP).** The trusted security service provides various security services to tenants for protecting their VMs.

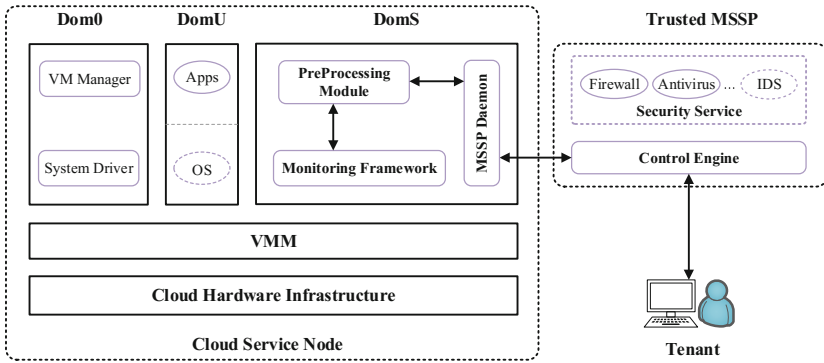


Fig. 1. Overview architecture of decoupling security services from cloud

After receiving tenant’s request for VM protection, the provider decide the protection mechanism by analyzing tenant’s requirements. He negotiates with the cloud service provider of the tenant to deploy security service. Once he was authenticated by the cloud service provider, he will be able to provide security protection for the tenant VMs.

**Cloud service node.** The cloud node is the physical server that runs the VMs of tenant. To support managed security services in the cloud node, a dedicated VM is coresident with management VM, named DomS. As show in Fig. 1, there are three main components in the DomS. The monitoring module acts as an agent for capturing various VM execution information. It leverages VMI to inspect specified VM for security service. The data analysis and processing module is used for analyzing collected data to get useful information for security service. The security service manager module helps the security service to access VMs state information.

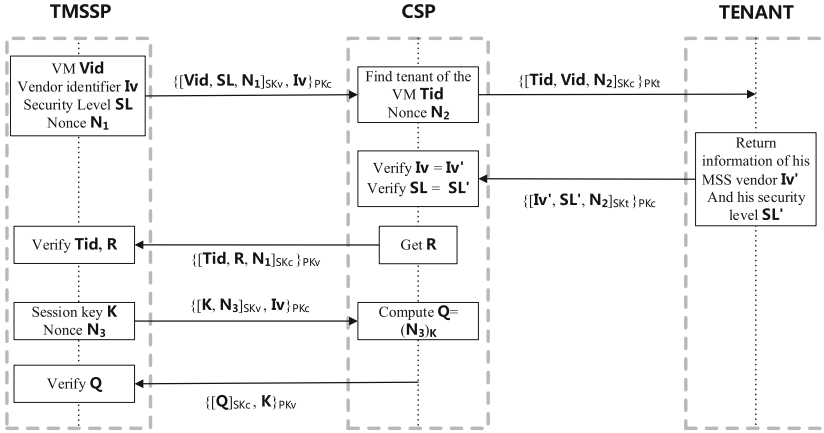
The VM manager module is responsible for managing VMs on behalf of tenant. It will shutdown or pause the VM in case of attacking accident was detected by the security service. The VM manager module is based on the existing management function of Dom0 and could be revoked by the security service.

### 3.4 Security Service Management

For tenant’s different choices, there are multiple managed security service providers need to be supported in our proposed architecture. It is crucial to ensure that the security services are authorized and constrained to the VMs which are associated with subscribing tenant. As described in Fig. 1, all the managed security services providers connect to the cloud platform with a MSSP daemon.

Each security service need to be authenticated before enabling it in the cloud platform. The MSSP daemon authenticates the vendor with corresponding tenant to establish trust between the security service vendor and the cloud provider.

Once the authentication succeed, the manager will authorize the security service to access VM related information and protect it. The authentication protocol is illustrated in Fig. 2.



**Fig. 2.** Authentication protocol for managed security service provider. We use the notation  $[M]_K$  for a private key operation with key K,  $M_K$  for a public key operation with key K, and  $(M)_K$  for a symmetric key operation with symmetric key K

In protocol description, we use notation ‘TMSSP’ for trusted managed security service provider, which is also called MSS vendor (to distinguish with cloud service provider) in this paper. Initially the security service vendor sends to CSP the protection requests including the VM identifier **Vid**, the security service vendor identifier **Iv**, the Security Level **SL** and a nonce **N<sub>1</sub>**. The CSP knows the tenant of requested VM and sends the tenant a verification request, including the tenant identifier **Tid**, the VM identifier **Vid** and a nonce **N<sub>2</sub>**. The tenant will return his security service vendor identifier **Iv’** and security level **SL’** related to **Vid** to the CSP. After verifying the consistency of security service vendor identifier and security level, the CSP generates a report **R** and sends it back to the vendor with tenant identifier **Tid** and the previous nonce **N<sub>1</sub>**. The trust relationship will be built if the vendor verifies the report successfully. Finally they should negotiate a session key for secure communications. When the security service is authenticated, the security service manager will create a monitoring instance associated with it. All the created instances will be recorded in a list, including its access rights derived from security level.

### 4 Implementation and Evaluation of SE-Cloud

We implement a prototype of our proposed architecture, named SE-Cloud, based on Xen and LibVMI [15]. LibVMI is an open-source implementation of VMI

supporting hypervisors such as Xen and KVM. It is a C library with Python bindings that makes it easy to monitor the low level details of a running virtual machine. The API provided by LibVMI support interacting with guest VM (pause/resume), inspecting guest memory, inspecting guest registers, and monitoring guest state. Various demonstrations are included with the software such as reading process lists from the guest memory, mapping symbol tables, and translating guest addresses.

#### 4.1 VM Monitoring and Date Capture

To assist security protection of hosted VMs, SE-Cloud should provide fine-grade and real time monitoring of VMs to the security service. Compared to traditional deployment of security service, SE-Cloud removes the security service agent from VMs by utilizing VMI. We monitor the guest VM in the DomS through hypervisor. The monitoring framework is showed in Fig. 3. The modules in the monitoring tools are used for network monitoring, memory inspecting and files monitoring. Those modules can be activated at runtime on demand.

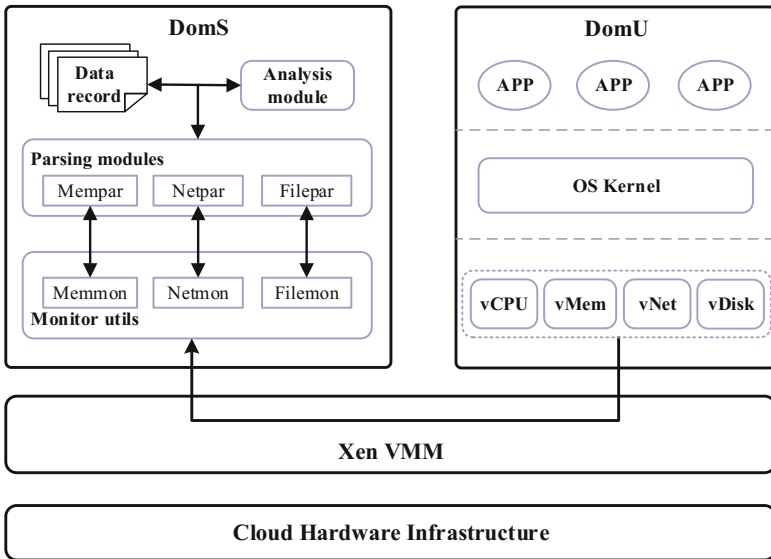


Fig. 3. Monitoring framework of SE-Cloud

**Memory Monitoring.** We implement memory monitoring by combining LibVMI with Volatility. LibVMI provides a set of functions to access virtual memory and vCPUs of each VM. What's more, it also provides a python interface named PyVMI which enables us to utilize Volatility<sup>4</sup> for memory inspection.

<sup>4</sup> Available online, <https://github.com/volatilityfoundation/volatility>.

It is possible to inspect and analyze Xen VMs' memory after we putting the `pyvmiaddressspace.py` in the `addrspaces` plugin directory of Volatility. In our present implementation of SE-Cloud, we used the plugins to capture information of process, network connections, loadable modules or DLLs for security service. For Extensibility, it is also convenient for security service vendors to develop their own plugins with the interfaces to access memory information provided by Volatility.

**Network Monitoring.** Xen provides three modes of network connection to VMs, namely bridge, NAT and route. As the default network connection mode, bridge mode is widely used in Xen VMs. Our implementation choose bridge mode network connection for VMs as well. Dom0 provides an Ethernet bridge connecting the physical network card to all virtual network devices provided by Xen to the DomU under bridge mode. The bridge forwards packets between physical network interface and virtual machine network interface. We intercept all the network packets just by hooking the bridge in the Dom0 kernel space. In our present implementation, we use a modified Ebttables<sup>5</sup> to intercept packets on the bridge. The Ebttables is a filtering tool for a Linux-based bridging firewall. It enables transparent filtering of network traffic passing through a Linux bridge.

**Files Monitoring.** We implement our file monitoring mechanism based on the assumption that file access event is more critical than the file content. As a result, our file monitoring focus on how to intercept and inspect all the file access requests instead of inspecting files content as Shi et al. [16] did. In our SE-Cloud, we use `filetracer` plugin of DRAKVUF [14] to monitor filesystem access. It use `FILE_OBJECT` to derive the access flags and file name, which can determine the full path of the file as well as the access privilege. All these data will be encapsulated in a vector along with the process information for security service analysis. In order to provide extensible function to advanced security service, our implementation is equipped with `libguestfs` for further analysis.

In order to bridge the semantic gap introduced by VMI, our architecture contains a parsing module to parse the raw data which get from guest VMs directly via monitoring framework. The parsing module incorporates some forensic tools to analysis captured data. All the data are collected in a central place denoted with data record in Fig. 3 and used for remote VMI. SE-Cloud preserves the integrity and confidentiality of introspected data between managed security services provider and cloud via secure communication.

## 4.2 Evaluation Results

In this section, we discuss the preliminary experiments on our present implementation of SE-Cloud. The experiments were performed on a physical DELL desktop (Intel Core i3-2130 CPU, with 4@3.40 GHz cores, 3M Cache, 12 GB memory,

<sup>5</sup> Available online, <http://ebtables.sourceforge.net/>.

Xen 4.8 and Linuxmint 18 are installed). We built tenant VM of WinXP with 1 GB virtual memory and 1 vcpu.

We first tested the function of monitoring tenant VM with remote VMI in SE-Cloud and then we traced a malware execution with our implemented analysis tools. At last, we evaluated the performance overhead of tenant VM in our framework.

**Tenant VM Monitoring with Remote VMI.** It is the foundation of our SE-Cloud to get runtime information of VMs effectively. To validate that our monitoring tools can capture the same data as what got by agent inside the VM, we used Windbg to get detail information of Windows process and compare with what we got out-of-VM with SE-Cloud. We took the *winlogon.exe* system kernel process as a example, because it is always infected by virus in Windows. We get detailed information of *winlogon.exe* running in the VM with windbg, which is showed in Table 1.

**Table 1.** Basic information of *winlogon.exe* in windbg

Item	Value	Description
PROCESS	85f2c9c8	The address of EPROCESS structure
Cid	01e0	Process id
ParentCid	0144	Parent process id
HandleCount	568	Number of handles number
Threads	20	Number of threads

Then we used monitoring tools in SE-Cloud to get the process information out of the VM. Figure 4 shows the screenshot of external review. The top screenshot shows the basic information of *winlogon.exe*, and the bottom screenshot lists the loaded DLLs for this process (only five modules for brevity).

To compare these two results of detailed information of *winlogon.exe*, it is easy to conclude that SE-Cloud could provide same view of the system as what we get inside VM. With the help of LibVMI, the semantic gap will be successfully bridged. To make deep analysis of this process, we dump it with monitoring tools and check the hash value of it. The result also confirms that we can get actual content of what is running inside VM.

### 4.3 Malware Detection

To verify that SE-Cloud is able to find suspicious activities in the VM which the security service focuses on, we used *hexdef100* to test the function of malware detection in SE-Cloud. Hacker defender (hxdef) is a Windows rootkit based on Windows NT 4.0. The main function of the rootkit is to rewrite the memory



```

Name                               Pid  PPid  Thds  Hnds  Time
-----
0x85cad020: explorer.exe           400   736   12    539  2017-04-08 04:45:57 UTC+0000
0x85b21da0: windbg.exe             1792  400   3    164  2017-04-08 18:17:25 UTC+0000
0x85b57020: IEXPLORE.EXE          1984  400   5    318  2017-04-08 17:07:47 UTC+0000
0x85b24020: notepad.exe           3096  400   1     52  2017-04-09 03:10:46 UTC+0000
0x85bb9020: taskmgr.exe           3152  400   3     82  2017-04-09 03:23:17 UTC+0000
0x85c92020: ctfmon.exe            344   400   1     82  2017-04-08 04:46:09 UTC+0000
0x85f7a00: System                   4     0    52   247  1970-01-01 00:00:00 UTC+0000
0x85e7fc00: smss.exe               324   4     2     23  2017-04-08 04:24:26 UTC+0000
0x85b7e778: csrss.exe             2564  324   0     ---  2017-04-09 03:58:27 UTC+0000
0x85e25750: csrss.exe             956   324   0     ---  2017-04-09 01:42:44 UTC+0000
0x85ed4350: csrss.exe             456   324  12    345  2017-04-08 04:24:27 UTC+0000
0x85e7e288: winlogon.exe          480   324  20    568  2017-04-08 04:24:28 UTC+0000
0x85e148b0: services.exe          524   480  15    244  2017-04-08 04:24:28 UTC+0000

*****
winlogon.exe pid: 480
Command line : winlogon.exe
Service Pack 3

Base      Size      LoadCount Path
-----
0x01000000 0x7d000 0xffff \\?\C:\WINDOWS\system32\winlogon.exe
0x7c920000 0x93000 0xffff C:\WINDOWS\system32\ntdll.dll
0x7c800000 0x11e000 0xffff C:\WINDOWS\system32\kernel32.dll
0x77da0000 0xa9000 0xffff C:\WINDOWS\system32\ADVAPI32.dll
0x77e50000 0x92000 0xffff C:\WINDOWS\system32\RPCRT4.dll
0x77fc0000 0x11000 0xffff C:\WINDOWS\system32\Secur32.dll
    
```

Fig. 4. Detailed information of *winlogon.exe* get in SE-Cloud

of running processes and hiding system component, such as the files, processes, system services, system drivers, the registry and opened port.

When we ran *hexdef100.exe* in WinXP VM, the program hid itself and we could not find it in the Windows taskmgr. But when we listed the running processes out of the VM, we found *hexdef100.exe* was running as a child process of *services.exe*, whose parent process was *winlogon.exe*. The screenshot in Fig. 5 shows the presence of *winlogon.exe* while it is hidden in guest VM.

```

0x85e7e288: winlogon.exe           480   324  20    564  2017-04-08 04:24:28 UTC+0000
0x85e148b0: services.exe           524   480  15    244  2017-04-08 04:24:28 UTC+0000
0x85d9b5f8: spoolsv.exe            1288  524  10    105  2017-04-08 04:25:44 UTC+0000
0x85dacb38: svchost.exe            1168  524  14    195  2017-04-08 04:25:44 UTC+0000
0x85dd0020: svchost.exe            856   524  21    204  2017-04-08 04:24:32 UTC+0000
0x85dd9da0: svchost.exe            824   524  81   3323  2017-04-08 04:24:32 UTC+0000
0x85cb6698: wscntfy.exe            416   824   1     37  2017-04-08 04:45:57 UTC+0000
0x85b98da0: wuauc1t.exe           2168  824   8    136  2017-04-09 04:02:18 UTC+0000
0x85d91da0: alg.exe                   1752  524   5    100  2017-04-08 04:25:59 UTC+0000
0x85db8020: svchost.exe            1088  524   5     84  2017-04-08 04:25:43 UTC+0000
0x85c12020: hxdef100.exe          3048  524   2     31  2017-04-09 03:14:14 UTC+0000
    
```

Fig. 5. Detection of *hexdef100.exe* in SE-Cloud

By deeply analysis with monitoring tools, we found that the *winlogon.exe* hooked itself and connected to the remote server with a hidden port.

#### 4.4 Performance Evaluation

We estimate overheads of guest VM introduced by SE-Cloud, we use the micro-benchmark of HardInfo for evaluation. We ran HardInfo benchmark in the guest VM without introspection to get baseline and then we ran HardInfo with different introspection intervals. We calculated the average of 10 times test value

as the result. The test result shows in Fig. 6. The higher the score is, the better performance the system gets. The first column in the left of the figure is the baseline. The other columns tagged with ‘interval-N’ mean that the introspection interval is N when we get the result. We can see that the overhead is related to the introspection interval. When the interval is larger than 8s, the computational overhead is negligible.

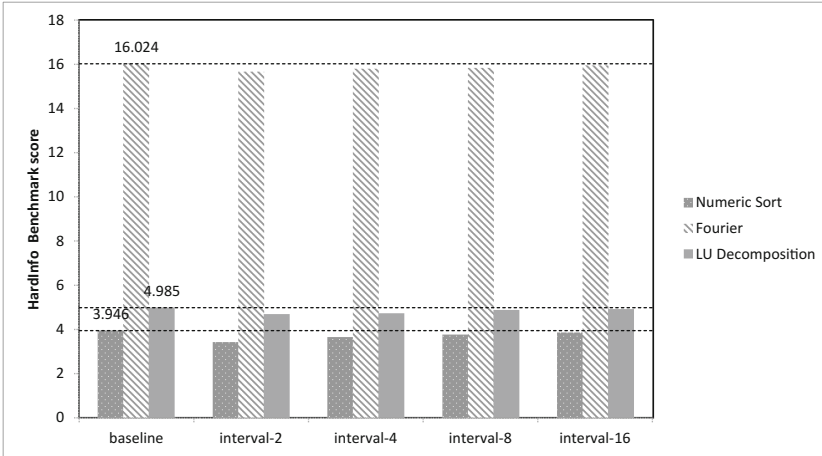


Fig. 6. Test result of benchmark with different intervals in the VM

## 5 Related Works

Researchers have investigated the use of virtual machine introspection in various security areas, such as touchless resource usage tracking [18], debugging [13], intrusion detection [8], malware detection [12] and so forth. Garfinkel and Rosenblum [8] are the first to propose the idea of virtual machine introspection for intrusion detection, which leverages strong isolation to protect the IDS from the VM being monitored. Srivastava and Giffin [17] proposed a white-list based application-level firewall in a virtualized environment which performs introspection for each new connection to identify the process bound to the connection. VMwatcher [12] is presented to enable “out-of-the-box” malware detection by addressing the semantic gap challenge. Ahmed et al. [1] shows HookLocator to real-time monitor the integrity of Windows kernel pools based entirely on virtual machine introspection.

Similar to SE-Cloud, Baek et al. [2] also aims at making privileged VMI as-a-service to cloud users and enables cloud centric introspection by allowing VMI actions to be performed across different physical machines. In contrast, SE-Cloud is a framework for supporting managed security services by leveraging remote VMI. It aims at providing tenants more robust and flexible security services.

## 6 Conclusion

In this paper, we propose decoupling security services from IaaS cloud by leveraging remote virtual machine introspection. By monitoring virtual memory, network and filesystem in a dedicated VM, the proposed technique enables security service deploy in a trusted place which beyond cloud provider's control, and protect tenant VM on the cloud platform. In our proposed technique, the security services are provided by a trusted MSSP, with which tenants are able to custom individual security services depend on their requirements. All the security services are authenticated before they are authorized to access monitoring information and protect tenant VMs. With some preliminary experiments, our present implementation with Xen hypervisor proves its ability to support remote managed security service into cloud to protect tenant VMs.

However, it is difficult to support legacy security service seamlessly in SE-Cloud at present. We need to improve our monitoring interface to be adapted to the agent way. In addition, we plan to further improve its performance by exploring hardware virtualization extensions in the future.

**Acknowledgments.** We would like to thank the anonymous reviewers for their insightful comments and suggestions on improving this paper. In this paper, the research was supported by the National Natural Science Foundation of China under Grant Nos. 61402508 and 61303191. The research also supported by National High Technology Research and Development Program of China (863) under Grant No. 2015AA016010.

## References

1. Ahmed, I., Richard, G.G., Zoranic, A., Roussev, V.: Integrity checking of function pointers in kernel pools via virtual machine introspection. In: Desmedt, Y. (ed.) ISC 2013. LNCS, vol. 7807, pp. 3–19. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-27659-5\\_1](https://doi.org/10.1007/978-3-319-27659-5_1)
2. Baek, H.W., Srivastava, A., Van der Merwe, J.: Cloudvmi: virtual machine introspection as a cloud service. In: Proceedings of the 2014 IEEE International Conference on Cloud Engineering (IC2E), pp. 153–158. IEEE (2014)
3. Baliga, A., Kamat, P., Iftode, L.: Lurking in the shadows: identifying systemic threats to kernel data. In: Proceedings of the IEEE Symposium on Security and Privacy (SP), pp. 246–251. IEEE (2007)
4. Bhattasali, T., Chaki, N.: Poster: exploring security as a service for IoT enabled remote application framework. In: Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services Companion (MobiSys Companion), p. 15. ACM (2016)
5. Butt, S., Lagar-Cavilla, H.A., Srivastava, A., Ganapathy, V.: Self-service cloud computing. In: Proceedings of the 2012 ACM Conference on Computer and Communications Security (CCS), pp. 253–264. ACM (2012)
6. Daniel, J., Dimitrakos, T., El-Moussa, F., Ducatel, G., Pawar, P., Sajjad, A.: Seamless enablement of intelligent protection for enterprise cloud applications through service store. In: Proceedings of the 2014 IEEE 6th International Conference on Cloud Computing Technology and Science (CloudCom), pp. 1021–1026. IEEE (2014)

7. Fu, Y., Lin, Z.: Bridging the semantic gap in virtual machine introspection via online kernel data redirection. *ACM Trans. Inf. Syst. Secur.* **16**(2), 1–29 (2013)
8. Garfinkel, T., Rosenblum, M., et al.: A virtual machine introspection based architecture for intrusion detection. In: *Proceedings of the Conference on Network and Distributed System Security Symposium (NDSS)*, pp. 191–206. Internet Society (2003)
9. Harrison, C., Cook, D., McGraw, R., Hamilton, J.: Constructing a cloud-based IDS by merging VMI with FMA. In: *Proceedings of the 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pp. 163–169. IEEE (2012)
10. Hurel, G., Badonnel, R., Lahmadi, A., Festor, O.: Outsourcing mobile security in the cloud. In: Sperotto, A., Doyen, G., Latré, S., Charalambides, M., Stiller, B. (eds.) *AIMS 2014. LNCS*, vol. 8508, pp. 69–73. Springer, Heidelberg (2014). [https://doi.org/10.1007/978-3-662-43862-6\\_9](https://doi.org/10.1007/978-3-662-43862-6_9)
11. Jain, B., Baig, M.B., Zhang, D., Porter, D.E., Sion, R.: SoK: introspections on trust and the semantic gap. In: *Proceedings of the 2014 IEEE Symposium on Security and Privacy (SP)*, pp. 605–620. IEEE (2014)
12. Jiang, X., Wang, X., Xu, D.: Stealthy malware detection through VMM-based “out-of-the-box” semantic view reconstruction. In: *Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS)*, pp. 128–138. ACM (2007)
13. King, S.T., Dunlap, G.W., Chen, P.M.: Debugging operating systems with time-traveling virtual machines. In: *Proceedings of the Annual Conference on USENIX Annual Technical Conference (ATEC)*, pp. 1–15. USENIX Association (2005)
14. Lengyel, T.K., Maresca, S., Payne, B.D., Webster, G.D., Vogl, S., Kiayias, A.: Scalability, fidelity and stealth in the DRAKVUF dynamic malware analysis system. In: *Proceedings of the 30th Annual Computer Security Applications Conference (ACSAC)*, pp. 386–395. ACM (2014)
15. Payne, B.D.: Simplifying virtual machine introspection using LibVMI. Technical report SAND2012-7818, Sandia National Laboratories (2012)
16. Shi, J., Yang, Y., He, J., Tang, C., Li, Q.: Design of a comprehensive virtual machine monitoring system. In: *Proceedings of the IEEE 3rd International Conference on Cloud Computing and Intelligence Systems (CCIS)*, pp. 510–513. IEEE (2014)
17. Srivastava, A., Giffin, J.: Tamper-resistant, application-aware blocking of malicious network connections. In: Lippmann, R., Kirda, E., Trachtenberg, A. (eds.) *RAID 2008. LNCS*, vol. 5230, pp. 39–58. Springer, Heidelberg (2008). [https://doi.org/10.1007/978-3-540-87403-4\\_3](https://doi.org/10.1007/978-3-540-87403-4_3)
18. Suneja, S., Isci, C., Bala, V., De Lara, E., Mummert, T.: Non-intrusive, out-of-band and out-of-the-box systems monitoring in the cloud. In: *Proceedings of the 2014 ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, pp. 249–261. ACM (2014)
19. Varadharajan, V., Tupakula, U.: Security as a service model for cloud environment. *IEEE Trans. Netw. Serv. Manage.* **11**(1), 60–75 (2014)
20. Wang, J., Stavrou, A., Ghosh, A.: HyperCheck: a hardware-assisted integrity monitor. In: Jha, S., Sommer, R., Kreibich, C. (eds.) *RAID 2010. LNCS*, vol. 6307, pp. 158–177. Springer, Heidelberg (2010). [https://doi.org/10.1007/978-3-642-15512-3\\_9](https://doi.org/10.1007/978-3-642-15512-3_9)

# Privacy Preserving Hierarchical Clustering over Multi-party Data Distribution

Mina Sheikhalishahi<sup>(✉)</sup> and Fabio Martinelli

Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche,  
56124 Pisa, Italy  
{mina.sheikhalishahi,fabio.martinelli}@iit.cnr.it

**Abstract.** This paper presents a framework for constructing a hierarchical categorical clustering algorithm on horizontal and vertical partitioned dataset. It is assumed that data is distributed between more than two parties, such that for general benefits all are willing to detect the clusters on whole dataset, but for privacy concerns, they avoid to share the original datasets. To this end, we propose algorithms based on *distributed secure sum* and *secure number comparison* protocols to securely compute the desired criteria in constructing clusters' scheme without revealing private data.

**Keywords:** Privacy · Hierarchical · Clustering · Distributed data  
Multi-party computation

## 1 Introduction

Facing the new challenges brought by a continuous evolving Information Technologies (IT) market, large companies and small-to-medium enterprises found in *Information Sharing* a valid instrument to improve their key performance indexes. Sharing data with partners, authorities for data collection and even competitors, may help in inferring additional intelligence through collaborative information analysis [10, 16]. Such an intelligence could be exploited to improve revenues, e.g. through best practice sharing [3], market basket analysis [12], or prevent loss coming from brand-new potential cyber-threats [6]. Other applications include analysis of medical data, provided by several hospitals and health centers for statistical analysis on patient records, useful, for example, to shape the causes and symptoms related to a new pathology [1].

Independently from the final goal, unfortunately information sharing brings issues and drawbacks which must be addressed. These issues are mainly related to the information privacy. Shared information might be sensitive, potentially harming the privacy of physical persons, such as employee records for business

---

This work has been supported by the H2020 EU funded project C3ISP [GA #700294].

applications, or patient records for medical ones. Hence, the most desirable strategy is the one which enables data sharing in secure environment, such that it preserves the individual privacy requirement while at the same time the data are still practically useful for analysis.

Clustering is a very well-known tool in unsupervised data analysis, which has been the focus of significant researches in different studies, spanning from information retrieval, text mining, data exploration, to medical diagnosis [2]. Clustering refers to the process of partitioning a set of data points into groups, in a way that the elements in the same group are more similar to each other rather than to the ones in other groups.

The problem of data clustering becomes challenging when data is distributed between two (or more) parties and for privacy concerns the data holders refuse to publish their own dataset, but still they are interested to shape more accurate clusters, identified on richer set of data.

To this end, we address the problem of securely constructing a hierarchical clustering algorithm, named CCTree, among several (more than two) parties. CCTree (Categorical Clustering Tree) [14] has a decision tree-like structure, which iteratively divides the data of a node on the base of an attribute, or domain of values, yielding the greatest entropy. The division of data is represented with edges coming out from a parent node to its children, where the edges are labeled with the associated values. A node which respects the specified stop conditions, i.e. either *pure enough* or containing few elements, is considered as a leaf. The leaves of the tree are the desired clusters.

In this study, we exploit *secure sum* and *secure number comparison* protocols to verify whether in each new node of CCTree (including the root) the stop conditions are hold or not; and if not, which attribute respects the highest entropy for data division. At the end, each data holder finds the structure of CCTree on whole data, without knowing the records of other parties.

In all this study it is assumed that clustering on joint datasets produces better result rather than clustering on individual dataset.

The contribution of this paper can be summarized as the following. A framework is proposed which serves as a tool for several parties to detect the structure of clusters (based on CCTree) on the whole dataset, without revealing their data. Distributed secure sum and secure number comparison protocols are exploited to propose new algorithms such that each party is able to verify *stop conditions* and (if not satisfy) find the *split attribute* for data division.

The rest of the paper is structured as follows. Related work on the two concepts of privacy preserving data clustering and secure decision tree construction is presented in Sect. 2. Section 3 presents some preliminary notations exploited in this study, including (1) CCTree (Categorical Clustering Tree), (2) distributed secure sum protocol, and (3) secure number comparison protocol. Section 4 describes the proposed framework, detailing the secure computation protocols for CCTree construction in both scenarios of data being either distributed horizontally or vertically among data holders. Finally, Sect. 5 briefly concludes and proposes the future research directions.

## 2 Related Work

The problem of privacy preserving data clustering is generally addressed for the specific case of  $k$ -means clustering, either when data is distributed between two parties [4, 8] or more than two parties [9]. Secure two-party  $k$ -means clustering is addressed in [4] with the use of *Paillier Homomorphic Encryption* scheme. Privacy preserving  $k$ -means clustering when data is distributed among more than two parties is addressed in [9], which proposes two protocols based on oblivious polynomial and homomorphic encryption, for computing cluster means. In [7], privacy preserving data clustering, when data is distributed horizontally among several parties is addressed through secure detection of *dissimilarity matrix*.

The more similar studies to what we proposed in present work can be found on secure decision tree construction, due to the fact that CCTree has decision tree like structure. In [18], the problem of secure *ID3* classification is addressed, when data are vertically distributed. The paper does not address the problem when data are partitioned horizontally. Moreover, addressing a classification problem, it is required to access labeled data, differently from what we addressed in present work. As a parallel solution, in [19], the problem of constructing secure *ID3* classifier is addressed when data are distributed horizontally. However, the problem is just discussed on horizontal partitioned data, and moreover, the data are labeled. In [5], the problem of secure *ID3* classification is studied when data is distributed among several parties. The proposed solution is based on secure sum protocols. However, the proposed approach can not be applicable when two parties are involved.

In [15], we proposed an approach to securely construct CCTree when data is distributed between two parties, differently from current study that the number of data holders are assumed to be more than two. In [17], CCTree is constructed formally among several parties.

In all aforementioned studies, differently from our approach, or data is distributed between two parties, or the problem is addressed for labeled data, or the number of clusters is known beforehand. To the best of our knowledge the problem of secure multi-party data clustering is a topic which is required to be explored deeper.

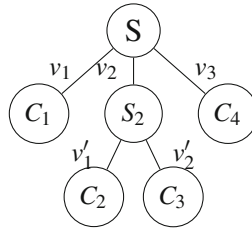
## 3 Preliminary Notations

### 3.1 Categorical Clustering Tree (CCTree) Construction

CCTree is a hierarchical categorical clustering algorithm, constructed iteratively through a decision tree-like structure, where the leaves of the tree are the desired clusters. In contrast to decision tree, CCTree is constructed on unlabeled data. A simple example of CCTree is depicted in Fig. 1.

The root of the CCTree contains all the elements desired to be clustered. Each element is described through a set of *categorical* attributes. Being categorical, each attribute may assume a finite set of discrete values, constituting its domain. For example the attribute *Language* may have its domain as  $\{English,$

*French, Spanish*}. At each new node of the tree (including the root), first two *stop conditions* are verified. If one of these stop conditions is hold, the node is labeled as *leaf*; Otherwise the best attribute is selected to generate new children nodes of the current node. The stop conditions for labeling a node as a leaf are as follows: (1) the number of elements in a node are less than a threshold, (2) the set of elements in a node are *homogeneous* enough. *Shannon Entropy* is the metric used both to define a homogeneity of a node, called *node purity*, and to select the best attribute to split a node. In particular non-leaf nodes are divided on the base of the attribute yielding the maximum value for Shannon entropy. The separation is represented through a branch for each possible outcome of the specific attribute. Each branch or edge extracted from parent node is labeled with the selected value which directs data to the child node. The process of CCTree construction can be formalized as follows.



**Fig. 1.** A small CCTree

**Input:** The input elements for constructing a CCTree are described in detail in the following:

- (1) *Attributes:* An ordered set of  $k$  attributes  $\mathcal{A} = \{A_1, A_2, \dots, A_k\}$  is given, where each attribute is an ordered set of mutually exclusive values. Thus, the  $i$ 'th attribute could be written as  $A_i = \{v_1 \cdot A_i, v_2 \cdot A_i, \dots, v_{|A_i|} \cdot A_i\}$ , where  $|A_i|$  is the number of values of attribute  $A_i$ , and  $v_r \cdot A_i$  ( $1 \leq r \leq |A_i|$ ) represents  $r$ 'th value of  $i$ 'th attribute.

Let  $n(v_r \cdot A_i, N_j)$  and  $n(N_j)$  represents the number of elements in node  $N_j$  that respect the  $r$ 'th value of  $i$ 'th attribute in node  $N_j$ . Considering  $p(v_r \cdot A_i, N_j) = \frac{n(v_r \cdot A_i, N_j)}{n(N_j)}$ , the *Entropy* of attribute  $A_i$  in node  $N_j$ , denoted by  $H(A_i, N_j)$ , is defined as:

$$H(A_i, N_j) = - \sum_{r=1}^{|A_i|} p(v_r \cdot A_i, N_j) \log_2 p(v_r \cdot A_i, N_j) \tag{1}$$

- (2) *Data Points:* A set  $\mathcal{D}$  of  $n$  data points is given, where each data point is a vector whose elements are the values of attributes, e.g.  $X_i = (v_{i_1}, v_{i_2}, \dots, v_{i_k})$ , where  $v_{i_r} \in A_r$ . For example we may have a record as “ID1 = (*Small, Red*)”.



(3) *Stop Conditions:* A set of stop conditions  $S = \{\mu, \varepsilon\}$  is given.  $\mu$  is the “*minimum number of elements in a node*”, i.e. when the number of elements in a node is less than  $\mu$ , then the node is not divided even if not pure enough.  $\varepsilon$  represents the “*minimum desired purity*” for each cluster, i.e. when the purity of a node is less or equal to  $\varepsilon$ , it will be considered as a leaf. To calculate the node purity, a function based on Shannon entropy is defined as follows:

Let  $n(v_r \cdot A_i, N_j)$  represents the number of elements having the  $r$ 'th value of  $i$ 'th attribute in node  $N_j$ , and  $n(N_j)$  be the number of elements in node  $N_j$ . Considering  $p(v_r \cdot A_i, N_j) = \frac{n(v_r \cdot A_i, N_j)}{n(N_j)}$ , the *purity* of node  $N_j$ , denoted by  $\rho(N_j)$ , is defined as:

$$\rho(N_j) = -\frac{1}{k} \sum_{i=1}^k \sum_{r=1}^{|A_i|} p(v_r \cdot A_i, N_j) \log_2(p(v_r \cdot A_i, N_j)) \quad (2)$$

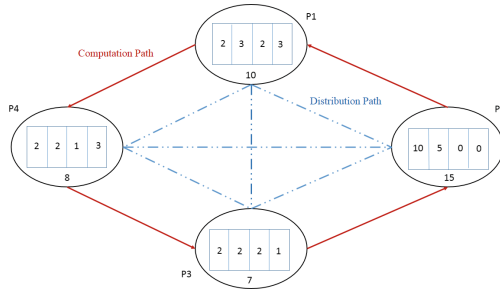
**Output:** The final output of the CCTree algorithm is a set of clusters, constituted by the leaves of the CCTree. For additional information on the CCTree algorithm we refer the reader to [14].

### 3.2 Distributed Secure Sum Protocol

In the following, we present the *distributed secure sum protocol* proposed in [13], proven to be resistant against colluding.

Assume that  $N$  parties  $P_1, P_2, \dots, P_N$  involve in a cooperative secure sum computation, where each party is able to break his private number into a fixed number of segments, such that the addition of segments is equal to her private number. In the proposed protocol the number of segments is equal to the number of parties ( $N$ ). The values of each segment is randomly selected by the associated party and it is secret from other parties. Then, each party holds one segment of her data and sends  $N - 1$  other segments to the other  $N - 1$  parties. In this way, at the end each party holds  $N$  segments, where only one belongs to the party and the others are collected from remaining parties, one from each. Now, the *secure sum protocol* can be applied to obtain the sum of all the segments. According to this protocol, one of the parties is required to be selected as the protocol initiator party that starts the computation by sending the data segment to the next party in the ring. The receiving party adds his data segment and to the received partial sum and then sends the result to the next party in the ring. This process is repeated till all the segments of all the parties are added and the sum is announced by the initiator party. For the sake of simplicity, in the rest of this paper, we represent the call of *distributed secure sum protocol* on  $N$  numbers  $x_i$  ( $1 \leq i \leq N$ ) as  $\mathcal{DSS}_{i=1}^N x_i$ , where without loss of generality  $x_1$  is the *initiator*.

In this scenario, even if two adjacent parties maliciously cooperate in order to discover the data of middle party, they just get some segments of the real data.



**Fig. 2.** Distributed secure sum among four parties (Color figure online)

Figure 2 shows the architecture for 4 parties before distribution of segments. Each party breaks the data block to four segments. There exist distribution paths for distributing the data segments to other parties before sum computation. The computation path is depicted with the solid red line, and the distribution path is demonstrated with dashed blue line.

### 3.3 Secure Numbers Comparison Protocol

Alice and Bob have natural numbers  $N_a$  and  $N_b$ , respectively. They want to know whose number is greater than the other one’s without revealing the numbers. To address this problem, we present and exploit the protocol proposed in [20]. Let  $1 < N_a, N_b < N$ ,  $\mathcal{M}$  be the set of all  $N$ -bit nonnegative integer numbers, and  $Q_N$  be the set of all one-to-one functions from  $\mathcal{M}$  to  $\mathcal{M}$ . Moreover, suppose  $E_a$  be the public key of Alice generated by a random variable from  $Q_N$ . Then, the protocol proceeds as the following:

- (Step 1): Bob selects a random  $N$ -bit integer, and computes the encrypted value  $x' = E_a(x)$
- (Step 2): Bob sends Alice the number  $x' - N_b + 1$
- (Step 3): Alice computes  $x^d = D_a(x' - N_b + u)$  for  $u = 1, \dots, N$ .
- (Step 4): Alice generates a random prime number  $p$  of  $\frac{N}{2}$ -bits and calculates  $Z_u = x^d \pmod p$  for all  $u$ ; if all  $z_u$  are different from each other by at least 2 in the  $\pmod p$  sense, stop; otherwise generates another random prime number and repeat the process until all  $z_u$  differ by at least 2. Without loss of generality, let  $p$  and  $z_u$  be the final results.
- (Step 5): Alice sends the prime number  $p$  and the following  $N$  numbers  $z_1, \dots, z_N$  to Bob.
- (Step 6): Bob looks the  $N_b$ ’th (not containing  $p$ ) sent from Alice, and decides that  $N_a \geq N_b$  if it is equal to  $x \pmod p$ , and  $N_a < N_b$  otherwise.
- (Step 7): Bob tell to Alice the conclusion.

In the rest of this study, we denote the call of *secure number comparison* protocol as a boolean function  $SNC$  which gets two natural numbers  $N_a$  and

$N_b$ , provided by *Alice* and *Bob* respectively; it returns 1 if  $N_a \geq N_b$ , and 0 otherwise. Formally, we have:

$$SNC(N_a, N_b) = \begin{cases} 1 & N_a \geq N_b \\ 0 & N_a < N_b \end{cases}$$

## 4 Framework Modeling

Suppose several (more than two) data holders are interested to detect the structure of clusters (through CCTree) on the whole of their datasets. However, for privacy concerns, they are not willing to publish or share the main dataset. As mentioned before, it is assumed that clustering on whole data (as in general cases) produces the better result comparing to clustering on individual dataset.

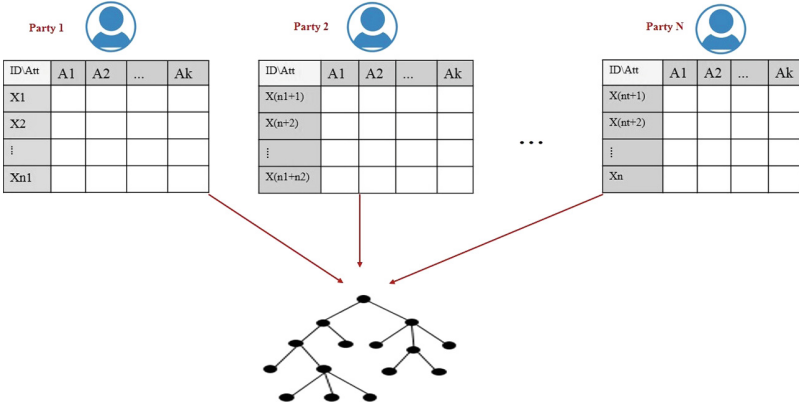
To this end, we propose algorithms, on the base of secure computation protocols, which verifies the stop conditions in each new node of CCTree, and if stop conditions are not hold, then the optimum attribute is found for data division. Basically each node of the CCTree is constructed through secure communications till all nodes are labeled as *leaf*. It is assumed that the participated parties set together the stop conditions, i.e. node purity ( $\epsilon$ ) and minimum number of elements ( $\mu$ ), before beginning the communication. Thence, the structure of CCTree is shaped on whole data, without revealing data to other parties. It is assumed that participated parties are *honest*, and they completely follow the protocols.

In Subsects. 4.1 and 4.2 secure CCTree construction over *horizontal* and *vertical* distributed data are addressed, respectively.

### 4.1 Horizontal Data Distribution

Let consider that  $N$  data holders are interested to model a CCTree clustering on whole of their data, when data is distributed horizontally among them. This means that each data holder has information about all the features but for different collection of objects. More precisely, let  $\mathcal{A} = \{A_1, A_2, \dots, A_k\}$  be the set of  $k$  categorical attributes all used to express each record of data, i.e. each record in each side is described as a  $k$  dimensional vector  $X_t = (v_{t_1}, v_{t_2}, \dots, v_{t_k})$ , where  $v_{t_i} \in A_i$ .

Figure 3 depicts a higher level representation of CCTree construction on horizontal distributed framework. *Party 1, Party 2, ...* and *Party N*, holding respectively datasets  $D_1, D_2, \dots, D_N$ , are  $N$  parties interested in constructing CCTree on  $\mathcal{D} = D_1 \cup D_2 \cup \dots \cup D_N$ , without knowing the data information of the other parties. As it can be observed, the  $N$  tables are described with the same set of attributes, but on different objects. To build the CCTree on whole data, participated parties communicate through upcoming secure algorithms, in each new node of the CCTree, to detect if the node satisfies the stop conditions (to be labeled as a leaf), and if not, which attribute respects the highest entropy for data division. At the end, all parties get the final CCTree structure identified on whole data, without revealing data to the other parties.



**Fig. 3.** CCtree construction on horizontal distributed data

**Attribute Entropy Computation.** In each new node of CCTree (including the root), it is required to securely compute the Shannon entropy of each attribute, for both verifying the node purity and for finding the split attribute (if it is not a leaf). To this end, all parties need to know the distribution of the values of each attribute on the whole data in a node without knowing data information.

To address it, first each party  $P_s$ , on her own dataset  $D_s$ , extracts the number of records respecting  $r$ 'th value of  $i$ 'th attribute, for all  $r, i$ , in node  $N_j$ . We denote this number as  $n_s(v_r \cdot A_i, N_j)$ . Moreover, each party  $P_s$  computes the number of elements of her dataset  $D_s$  which arrives in node  $N_j$ , denoted as  $n_s(N_j)$ .

Then, if all parties obtain *securely* the sum of  $n_s(v_r \cdot A_i, N_j)$  and  $n_s(N_j)$  over the range of  $s$ , then they are able to extract the entropy of attribute  $A_i$  in node  $N_j$  according to the following formula:

$$H(A_i, N_j) = - \sum_{r=1}^{|A_i|} \frac{Sum_s (n_s(v_r \cdot A_i, N_j))}{Sum_s (n_s(N_j))} \log_2 \left( \frac{Sum_s (n_s(v_r \cdot A_i, N_j))}{Sum_s (n_s(N_j))} \right)$$

where  $Sum_s (n_s(v_r \cdot A_i, N_j))$  and  $Sum_s (n_s(N_j))$  are the sum of  $n_s(v_r \cdot A_i, N_j)$  and  $n_s(N_j)$  over  $s$ . Algorithm 1 details the process of secure attribute entropy computation.

**Theorem 1.** *Algorithm 1 reveals nothing to any data holder except the total statistical information of the whole dataset.*

*Proof.* The only communications among parties occur at lines 9 and 10 which consists of a call to the distributed secure sum function ( $\mathcal{DSS}$ ) proven to be secure in [13]. □

**Stop condition (node purity).** In a node of CCTree, when the data constituting the node are *pure enough*, the node is labeled as a leaf (cluster). Before

---

**Algorithm 1.** *Att.Entropy()*: Secure computation of attribute entropy

---

**Data:** Each party  $P_s$  has the statistical information of attribute  $A_i$  in node  $N_j$  on her own dataset  $D_s$

**Result:** Entropy of  $A_i$  on whole data in node  $N_j$

```

1 initialization;
2 for 1 ≤ i ≤ k do
3   for 1 ≤ r ≤ |Ai| do
4     for 1 ≤ s ≤ N do
5       Ps: ns(vr · Ai, Nj) ← the number of records in dataset Ds
6         respecting r'th value of i'th attribute in node Nj
7       Ps: ns(Nj) ← the number of records of Ds in node Nj
8     end
9   end
10  Sums(ns(vr · Ai, Nj)) ← ∑s=1N ns(vr · Ai, Nj)
11  Sums(ns(Nj)) ← ∑s=1N ns(Nj)
12  return H(Ai, Nj) = - ∑r=1|Ai|  $\frac{Sum_s(n_s(v_r \cdot A_i, N_j))}{Sum(n_s(N_j))} \log_2\left(\frac{Sum_s(n_s(v_r \cdot A_i, N_j))}{Sum(n_s(N_j))}\right)$ 

```

---

CCTree construction, all parties set and fix together the required threshold of node purity, say  $\epsilon$ . The formula for *node purity* computation has been defined in Eq. 2. Algorithm 2 details the process.

---

**Algorithm 2.** *Node.purity()*: Secure Node Purity Computation

---

**Data:** Each party has a statistical information of node  $N_j$  in her own dataset.

**Result:** Each party securely computes the purity of node  $N_j$  on whole data of  $N_j$

```

1 initialization;
2 ρ(Nj) = 0
3 for 1 ≤ i ≤ k do
4   Att.Entropy(Ai, Nj)
5   ρ(Nj) =  $\frac{1}{k}(\rho(N_j) + Att.Entropy(A_i, N_j) )$ 
6 end
7 return ρ(Nj)

```

---

**Theorem 2.** *Algorithm 2 reveals only the entropy of attributes and not the counts of values in each party’s dataset.*

*Proof.* The only communication between parties occurs at line 4, which is a call to the Algorithm 1 proven to be secure in Theorem 1. □

If all parties find that the purity of a specific node is less than the purity threshold  $\epsilon$ , then that node is labeled as a leaf; Otherwise the other stop condition, i.e. number of elements in a node is verified through Algorithm 3. If both are not hold, the best attribute for data division is found through Algorithm 1, i.e. the attribute with the highest entropy.

**Stop condition (number of elements).** In CCTree construction, one of the stop conditions for labeling a node as a leaf, is that the number of elements in the node be less than a specified threshold. To this end, before constructing the CCTree, all  $N$  parties set together the minimum number of elements in a node to be identified as a leaf, say  $\mu$ . Now suppose that the number of elements of  $D_s$  which arrives to node  $N_j$  be denoted as  $n_s(N_j)$ , then the sum of  $n_s(N_j)$  over  $s$  returns the number of elements in  $N_j$ . After having the total number of elements which have been reached to node  $N_j$ , all parties know if it satisfies the stop condition. It is noticeable that the sum of the results is computed through distributed secure sum protocol. Algorithm 3 gives the details of secure verification of this stop condition.

---

**Algorithm 3.** *Num.elements()*: Secure number of elements validation

---

**Data:** Party  $P_s$  has private number  $n_s(N_j)$ ;  $\mu$  is a public number.

**Result:** All parties know securely whether  $\sum_s n_s(N_j) \leq \mu$  or not.

```

1 initialization;
2 for  $1 \leq s \leq N$  do
3   |  $P_s$ :  $n_s(N_j) \leftarrow$  the number of records in  $D_s$  reach to  $N_j$ 
4 end
5  $Sum(n_s(N_j)) \leftarrow \mathcal{DSS}_{s=1}^N(n_s(N_j))$ 
6 return  $Sum(n_s(N_j))$ 

```

---

## 4.2 Vertical Data Distribution

Now lets us consider that *Party 1, Party 2, ..., Party N* are interested to detect the clusters on the whole of their datasets, when data is partitioned *vertically* among  $N$  parties. This means that all data holders each holds the same set of objects, but described with different sets of attributes in each side. More precisely, let  $\mathcal{A} = \{A_1, A_2, \dots, A_k\}$  be the set of categorical attributes used to describe each record of data. However, having the same set of objects, each party holds excluded parts of attributes describing the data. Since in reality, it is possible they also have some common attributes, we assume that the common attributes to be shared with one of them. Without loss of generality, let consider *Party 1* has the description of data based on  $\mathcal{A}^1 = \{A_1, \dots, A_{k_1}\}$ , *Party 2* owns the set  $\mathcal{A}^2 = \{A_{k_1+1}, \dots, A_{k_1+k_2}\}$ , ..., and *Party N* has the information of attributes  $\mathcal{A}^N = \{A_{k(N-1)+1}, \dots, A_{k(N-1)+k_N}\}$  describing the same set of objects  $D$ .

Figure 4 depicts a higher level representation of CCTree construction when data is divided vertically among  $N$  parties. As it can be observed,  $N$  tables contain the same set of objects but expressed with different attributes. At the end all participated parties obtain CCTree structure on whole data without knowing each others' information. After obtaining the final CCTree structure, it is exploited in each site for data clustering.

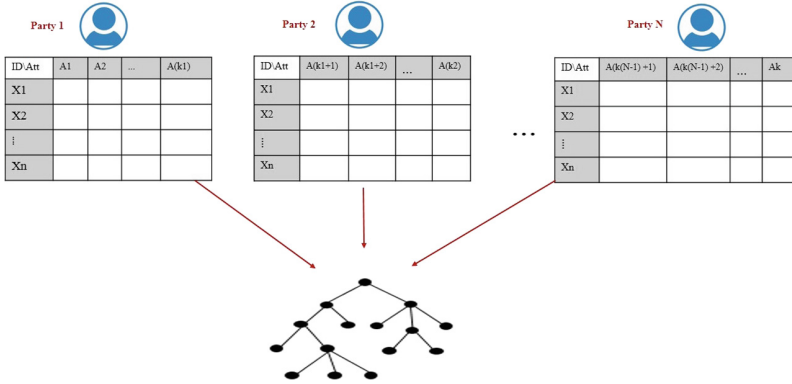


Fig. 4. CCtree construction on vertical distributed data

**Optimum Attribute Selection.** To find the optimum attribute for data division, first each party finds the best attribute, in terms of highest entropy, in her/his own dataset, say  $A_s^* \in P_s$ . Then it is required to verify which  $H(A_s^*)$  has the highest entropy among all. To run this verification securely, we plan to exploit the secure number comparison protocol presented in Sect. 3.3. However, secure number comparison protocol is only applicable on natural numbers, whilst the result of attribute entropy is not necessarily a natural number. To this end, suppose all attribute entropies are rounded to maximum  $l$  decimal numbers. Then, for comparing the attribute entropies, it is merely enough to multiply the results by  $10^l$ . Algorithm 4 details the process of finding the best split attribute.

---

**Algorithm 4.** *Att.selection()*: Secure Attribute Selection in Vertical Distributed Data

---

**Data:** Each party  $P_s$  finds the best attribute  $A_s^*$  in node  $N_j$ .  
 $l$  is the maximum length of decimal numbers.

**Result:** All parties *securely* obtain the optimum attribute  $A^* \in \mathcal{A}$  with the highest entropy.

```

1  $H(A_0^*) = 0$ 
2 initialization;
3 for  $1 \leq s \leq N$  do
4    $P_s : A_s^* \leftarrow \operatorname{argmax}_{A_{k(s-1)} \leq A_t \leq A_{k_s}} H(A_t)$ 
5   if  $\operatorname{SNC}(10^l \cdot H(A_s^*), 10^l \cdot H(A_{s-1}^*))$  then
6      $A^* \leftarrow A_s^*$ 
7   else
8      $A^* \leftarrow A_{s-1}^*$ 
9   end
10 end
11 return  $A^*$ 

```

---

**Theorem 3.** *Algorithm 4 only reveals the result of attribute entropy comparison but not the output of each party.*

*Proof.* The only communication among parties occurs at line 5, which is a call of secure number comparison protocol proven to be secure in [20]. □

**Stop condition (node purity).** To verify the satisfaction of *node purity*, as a stop condition, when data is distributed vertically, it is enough that each party  $P_s$  computes the purity of node  $N_j$  on her own side, say  $\rho_s(N_j)$ , and then multiplies the outcome by  $10^l$ . Then, they sum up the results securely with the use of distributed secure sum protocol. After having the sum of all results, the final result is divided to  $k \times 10^l$ . The rationale behind it comes from the fact that each party holds a set of attributes exclusively from others. Thence, by separately computing the purity of a node in each side, it is possible to obtain the purity of a node by summing up the results on all attributes. Algorithm 5 details the process.

---

**Algorithm 5.** *Node.purity.V():* Secure node purity computation in vertical distributed data

---

**Data:** Each party had data statistical information in node  $N_j$ ,  $l$  is the maximum length of decimal numbers.

**Result:** All parties securely obtain  $\rho(N_j) = \frac{1}{k} \sum_{i=1}^k H(A_i)$

1 initialization;

2 **for**  $1 \leq s \leq N$  **do**

3      $P_s : \rho_s(N_j) \leftarrow \sum_{t=k(s-1)}^{ks} H(A_t)$

4      $\rho(N_j) \leftarrow \mathcal{DSS}_{s=1}^N (\rho_s(N_j) \times 10^l)$

5 **end**

6 **return**  $\frac{1}{k \times 10^l} \rho(N_j)$

---

**Theorem 4.** *Algorithm 5 reveals nothing except the purity of a specific node.*

*Proof.* The only communication among parties occurs at line 4, which is a call of secure distributed sum protocol, proven to be secure in [13].

**Stop condition (number of elements).** In the case that the participating parties are willing to find that if the number of elements in a node is less than a threshold, first, before CCTree construction, all parties set together this threshold, say  $\mu$ . Then, since data is distributed vertically, in each new obtained node, *only* one of the party is required to declare if the condition of minimum number of elements is hold or not. The rationale behind it is that the division attribute is just selected from one side in vertical distribution scenario. Hence, suppose that the attribute having the highest entropy among all attributes in a node (without



loss of generality) belongs to the set of attributes of party  $P_s$ , say  $\mathcal{A}_s$ . Thence, after data division applying  $\mathcal{A}_s$ , party  $P_s$  says that whether in new obtained nodes the number of elements is less than  $\mu$  or not. However, she is not required to specify the number of elements. Algorithm 6 details the process.

---

**Algorithm 6.** *Num.elements.V()*: Secure number of elements comparison, vertical distribution

---

**Data:** Node  $N_j$ .

**Result:** Whether  $n(N_j) < \mu$  or not.

```

1 initialization;
2 for  $1 \leq s \leq N$  do
3   if father of node  $N_j$  is split with  $P_s$ 's attribute then
4     if  $n(N_j) < \mu$  then
5        $P_s : r \leftarrow 1$ 
6     else
7        $P_s : r \leftarrow 0$ 
8     end
9   end
10 end
11 return  $r$ 

```

---

Thence, if the result of Algorithm 6 equals to 1, then the node is labeled as a leaf, otherwise the optimum attribute for data division is selected through Algorithm 4.

**Theorem 5.** *Algorithm 6 reveals nothing about other party dataset information, except that if the number of elements in a node is less than specified threshold.*

*Proof.* The proof is resulted from the simple fact that no party specifies number of elements in a node and just determines if the stop condition of *number of elements* in a node satisfies or not.

## 5 Conclusion and Future Directions

In this work we proposed a framework which can be exploited for three (or more) parties to construct a hierarchical categorical clustering algorithm, named CCTree, on whole of their data, without revealing the original datasets. To this end, secure computation algorithms are proposed to obtain the required criteria for detecting the clusters on the whole data. Two scenarios of data being distributed either horizontally or vertically have been considered.

In the proposed approach, in this study, for constructing CCTree over horizontal and vertical distributed data, we proved that the communications among data holders are secure. This means that it is supposed that the participated parties will not reveal their own private information, but through secure communications they are able to obtain the final result of a specific function. This

function might return the result of number comparison or the summation of numbers. Although it is assumed that all participated parties are honest and they completely follow the protocols, it is still noticeable that the final result might leak some private information. For example, in computing the sum among  $N$  parties, if the result of the sum is equivalent to the *shared number* of one party, that party will notice that the private number of all other parties has been equal to *zero*. To avoid the privacy leakage of the output result, in future direction we plan to add systematic noise through  $\epsilon$  differential privacy according to the methodology proposed in [11]. In this way, data holders add some perturbation noise to their own real data, based on an *utility function*, which while at the end it does not reveal the precise private number, still it preserves the effectiveness of provided data.

Moreover, we plan to generalize the proposed framework to the wide range of hierarchical clustering algorithms. Furthermore, we plan to analyze the efficiency of proposed approach on clustering the benchmark datasets to evaluate communication cost in reality.

## References

1. Artoisenet, C., Roland, M., Closon, M.: Health networks: actors, professional relationships, and controversies. In: Collaborative Patient Centred eHealth, vol. 141. IOS Press (2013)
2. Berkhin, P.: A survey of clustering data mining techniques. In: Kogan, J., Nicholas, C., Teboulle, M. (eds.) Grouping Multidimensional Data, pp. 25–71. Springer, Heidelberg (2006)
3. Bogan, E., English, J.: Benchmarking for Best Practices: Winning Through Innovative Adaptation (1994)
4. Bunn, P., Ostrovsky, R.: Secure two-party k-means clustering. In: Proceedings of the 14th ACM Conference on Computer and Communications Security, CCS 2007, pp. 486–497. ACM, New York (2007)
5. Emekci, F., Sahin, O., Agrawal, D., Abbadi, A.E.: Privacy preserving decision tree learning over multiple parties. Data Knowl. Eng. **63**(2), 348–361 (2007)
6. Faiella, M.F., Marra, A.L., Martinelli, F., Mercaldo, F., Saracino, A., Shekhalishahi, M.: A distributed framework for collaborative and dynamic analysis of android malware. In: 25th Conference on Parallel, Distributed, and Network-Based Processing, St. Petersburg (2017)
7. Inan, A., Kaya, S.V., Saygn, Y., Savas, E., Hintoglu, A.A., Levi, A.: Privacy preserving clustering on horizontally partitioned data. Data Knowl. Eng. **63**(3), 646–666 (2007). 25th International Conference on Conceptual Modeling (ER 2006)
8. Jagannathan, G., Pillaipakkamnatt, K., Wright, R.N.: A new privacy-preserving distributed k-clustering algorithm. In: SDM, pp. 494–498. SIAM (2006)
9. Jha, S., Kruger, L., McDaniel, P.: Privacy preserving clustering. In: di Vimercati, S.C., Syverson, P., Gollmann, D. (eds.) ESORICS 2005. LNCS, vol. 3679, pp. 397–417. Springer, Heidelberg (2005). [https://doi.org/10.1007/11555827\\_23](https://doi.org/10.1007/11555827_23)
10. Martinelli, F., Saracino, A., Shekhalishahi, M.: Modeling privacy aware information sharing systems: a formal and general approach. In: 15th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (2016)

11. McSherry, F., Talwar, K.: Mechanism design via differential privacy. In: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2007, pp. 94–103 (2007)
12. Oliveira, S.R.M., Zaïane, O.R.: Privacy preserving frequent itemset mining. In: Proceedings of the IEEE International Conference on Privacy, Security and Data Mining, CRPIT 2014, vol. 14, pp. 43–54 (2002)
13. Sheikh, R., Kumar, B., Mishra, D.K.: A distributed k-secure sum protocol for secure multi-party computations. CoRR abs/1003.4071 (2010)
14. Sheikhalishahi, M., Saracino, A., Mejri, M., Tawbi, N., Martinelli, F.: Fast and effective clustering of spam emails based on structural similarity. In: Garcia-Alfaro, J., Kranakis, E., Bonfante, G. (eds.) FPS 2015. LNCS, vol. 9482, pp. 195–211. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-30303-1\\_12](https://doi.org/10.1007/978-3-319-30303-1_12)
15. Sheikhalishahi, M., Martinelli, F.: Privacy preserving clustering over horizontal and vertical partitioned data. In: 2017 IEEE Symposium on Computers and Communications, ISCC 2017, Heraklion, Greece, pp. 1237–1244, 3–6 July 2017
16. Sheikhalishahi, M., Martinelli, F.: Privacy-utility feature selection as a privacy mechanism in collaborative data classification. In: The 26th IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, Poznan, Poland (2017)
17. Sheikhalishahi, M., Mejri, M., Tawbi, N., Martinelli, F.: Privacy-aware data sharing in a tree-based categorical clustering algorithm. In: Cuppens, F., Wang, L., Cuppens-Boulahia, N., Tawbi, N., Garcia-Alfaro, J. (eds.) FPS 2016. LNCS, vol. 10128, pp. 161–178. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-51966-1\\_11](https://doi.org/10.1007/978-3-319-51966-1_11)
18. Vaidya, J., Clifton, C., Kantarcioglu, M., Patterson, A.S.: Privacy-preserving decision trees over vertically partitioned data. ACM Trans. Knowl. Discov. Data **2**(3), 14:1–14:27 (2008)
19. Xiao, M.J., Huang, L.S., Luo, Y.L., Shen, H.: Privacy preserving ID3 algorithm over horizontally partitioned data. In: Sixth International Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT 2005), pp. 239–243, December 2005
20. Yao, A.C.: Protocols for secure computations. In: Proceedings of the 23rd Annual Symposium on Foundations of Computer Science, SFCS 1982, pp. 160–164. IEEE Computer Society, Washington, D.C. (1982)

# Improving MQTT by Inclusion of Usage Control

Antonio La Marra<sup>1</sup>, Fabio Martinelli<sup>1</sup>, Paolo Mori<sup>1</sup>, Athanasios Rizos<sup>1,2(✉)</sup>,  
and Andrea Saracino<sup>1</sup>

<sup>1</sup> Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche, Pisa, Italy  
{antonio.lamarra,fabio.martinelli,paolo.mori,athanasios.rizos,  
andrea.saracino}@iit.cnr.it

<sup>2</sup> Department of Computer Science, University of Pisa, Pisa, Italy

**Abstract.** Due to the increasing pervasiveness of Internet of Things (IoT) and Internet of Everything (IoE) devices, securing both their communications and operations has become of capital importance. Among the several existing IoT protocols, Message Queue Telemetry Transport (MQTT) is a widely-used general purpose one, usable in both constrained and powerful devices, which coordinates data exchanges through a publish/subscribe approach. In this paper, we propose a methodology to increase the security of the MQTT protocol, by including Usage Control in its operative workflow. The inclusion of usage control enables a fine-grained dynamic control of the rights of subscribers to access data and data-streams over time, by monitoring mutable attributes related to the subscriber, the environment or data itself. We will present the architecture and workflow of MQTT enhanced through Usage Control, also presenting a real implementation on Raspberry Pi 3 for performance evaluation.

## 1 Introduction

Over the last years, Internet of Things (IoT) devices have become more and more pervasive to our daily life. As a matter of fact, we are currently using many connected objects, such as smart house appliances, connected cars, remote surveillance cameras, smart meters etc. According to Ericsson [5], in 2020 we should expect the total number of IoT devices to reach 50 billions, and this number becomes even more dramatic if we consider the Internet of Everything (IoE) paradigm, which also includes user devices such as smartphones, smartwatches, tablets, etc.

IoT devices could be very different, because they typically have different types of hardware, depending on the provided functionalities, and software applications to manage them. Hence, in order to have a unique application which eases the control of all the smart devices owned by the same user, a necessity has arisen to be able to easily communicate with a set of distinct IoT devices. To this aim, several application layer protocols have been proposed in the scientific literature, and among them, MQTT is one of the most widely used [1].

MQTT<sup>1</sup> is also recently standardized by OASIS<sup>2</sup> and works according to the Publish/Subscribe protocol pattern, where a central *Broker* handles the communications and data sharing, collecting data from a set of *Publishers* and redistributing them to a set of *Subscribers*, according to their specified interests.

According to [20], the MQTT standard and the existing implementations, provide support only for basic authentication and simple authorization policies, applied to Subscribers at subscription time. Since MQTT is based on HTTP functionalities, most of the MQTT security solutions seem to be either application specific, or just leveraging TLS/SSL protocols [1]. Currently, OASIS MQTT security subcommittee is working on a standard to secure MQTT messaging using MQTT Cybersecurity Framework [17]. Although the effort concerning security of MQTT protocol is rising, two main obstacles occur. The first one is that, although the protocol has the ability to deal with various components that become Publishers or Subscribers, the fact that they use different platforms makes it difficult to create and enforce a generic security policy. The second problem is that the current efforts are mainly directed to message communication security, to avoid eavesdropping, integrity violation and MITM attacks. Still no efforts have been done in the directions of supporting policy enforcement at Broker level, nor it has been considered the possibility of dynamically revoking subscriptions.

In this paper, we propose to enhance the security of the MQTT protocol by adding Usage Control (UCON) in the MQTT architecture and workflow. UCON is an extension of traditional access control which enforces continuity of access decision, by evaluating policies based on mutable attributes, i.e. attributes changing over time [10]. Adding Usage Control in MQTT we aim at enforcing dynamically fine grained policies, which do not only consider the identity of the Subscriber as a parameter for granting access to data, but also dynamic attributes such as Subscriber reputation, data reliability, or environmental conditions of a specific application. After surveying the main IoT application protocols, and motivating the choice of focusing on MQTT, this work will discuss the architecture and the workflow of the MQTT - UCON integration. The proposed framework is designed to be general, easy to integrate in the Broker component, remaining oblivious to both Publishers and Subscribers. The addition of UCON in fact, does not modify the MQTT protocol, enforcing the policies independently from the implementation of Publisher and Subscriber, which allows the proposed solution to be compatible with any Off-the-Shelf MQTT Publisher/Subscriber application. Furthermore, we will demonstrate the viability of the approach by presenting a real implementation of the framework on both general purpose and performance constrained devices, discussing also the performance measured on two Raspberry Pi 3 model b<sup>3</sup>, used respectively as Broker and Subscriber.

The rest of the paper is organized as follows: In Sect. 2 a comparison between the main IoT application protocols is reported, detailing afterward the MQTT

---

<sup>1</sup> <http://mqtt.org>.

<sup>2</sup> <https://www.oasis-open.org/>.

<sup>3</sup> <https://www.raspberrypi.org/products/raspberry-pi-3-model-b/>.

protocol and motivating our choice to focus on it. Furthermore, some background information about usage control are reported. Section 3 describes the integration of UCON and MQTT detailing the architecture and the operative workflow. Section 4 details the results of the performance analysis. In Sect. 5 are reported a set of related works about security in MQTT and application of UCON in IoT. Finally, Sect. 6 concludes by proposing future directions which stem from this preliminary work.

## 2 Background

In this section we will survey the main protocols for IoT, motivate the choice to focus on MQTT, briefly describing the protocol and we will recall some background notions on the concept of usage control.

### 2.1 IoT Application Protocols

The most known application layer protocols in IoT are CoAP, MQTT, XMPP, HTTP, AMQP and WebSocket. In [7], the authors claim that CoAP is more Resource-friendly than MQTT but in terms of Message Oriented Approach (MOA), MQTT stands out. They report also that MQTT needs less RAM but more CPU load than CoAP. All the protocols mentioned above use TCP as transport layer. Only CoAP uses UDP. The same happens as for the security layer. All protocols use TLS/SSL except from CoAP that uses DTLS. In fact, CoAP targets to very constrain environments.

Furthermore, according to [8] MQTT provides the smallest header size of two bytes, although it is based on TCP. Moreover, it provides three levels of QoS which puts this protocol in the first place in terms of QoS, even though it needs extra load in the network for message retransmission. On the other hand, XMPP requires processing and storing XML data, which necessitates memory space too large for most IoT devices. In addition, HTTP performs better in non constrained environments when PC, Laptop and Servers are used. It is generally not applicable in IoT devices due to its high overhead. AMQP [14], is more suitable for server-to-server communication than device-to-device communication. WebSocket is neither a request/response nor a Publish/Subscribe protocol. In WebSocket, a client initializes a handshake with a server to establish a WebSocket session. The handshake process is intended to be compatible with HTTP-based server-side software so that a single port can be used by both HTTP and WebSocket clients [4]. According to [21], MQTT messages experience lower delays than CoAP for lower packet loss and vice versa. When the message size is small the loss rate is equal. AllJoyn [22], is a full stack of protocols intended for IoT. Though quite popular, the main disadvantage of AllJoyn is that the application protocol cannot be separated from the rest of the protocol stack. Due to this fact, Alljoyn is a complete framework and not only an application layer protocol. Thus, it is not taken into consideration in this study. As a synopsis to the basis, reader can also consult Table 1. This comparison gives the details about

the existence of Quality of Service (QoS). Also it refers to the communication pattern which in the case of MQTT is the Publish/Subscribe. The most significant column is the third. In this column, we identified that MQTT is more general purpose.

**Table 1.** Application layer protocol comparison

Protocol	QoS	Communication pattern	Target devices
CoAP	YES	Req/Resp	Very constrained
MQTT	YES	Pub/Sub	Generic
XMPP	NO	Req/Resp Pub/Sub	High memory consumption
HTTP	NO	Req/Resp	High performance
AMQP	YES	Pub/Sub	Ser-2-Ser communication
Web socket	NO	Client/Server Pub/Sub	Needs less power than HTTP still need high power
AllJoyn	NO	Client/Server Pub/Sub	High computational power

## 2.2 The MQTT Protocol

MQTT is an open pub/sub protocol designed for constrained devices used in telemetry applications. MQTT is designed to be very simple on the client side either this is the Subscriber or the Publisher. Hence, all of the system complexities reside on the Broker which performs all the necessary actions for the MQTT functionality. MQTT is independent from the routing or networking specific algorithms and techniques. However, it assumes that the underlying network provides a point-to-point, session-oriented, auto-segmenting data transport service with in-order delivery (such as TCP/IP) and employs this service for the exchange of messages.

MQTT is a topic-based Publish/Subscribe protocol that uses character strings to provide support of hierarchical topics. This means that there is the ability to create and control the hierarchy of the topics. There is also the opportunity to the subscription to multiple topics. In Fig. 1, we can see the topology of the protocol. It consists of the Publisher(s) that send the data to the Broker for publishing. A Subscriber authenticates and subscribes to the Broker for certain topics. Moreover, the Broker sends the data to the specific Subscriber(s) that are subscribed to the specific topic of the message. The Broker is responsible to distribute them to the related Subscribers correctly. The Publishers and the Subscribers can be very constrained devices, especially in the case of Publishers that can be even a sensor. On the other hand, though, the Broker has to have enough computational power so as to be able to handle the amount of data being distributed. MQTT supports basic end-to-end Quality of Service (QoS) [3]. Depending on how reliably messages should be delivered to their receivers, MQTT provides three QoS levels. QoS level 0 only offers a best-effort delivery

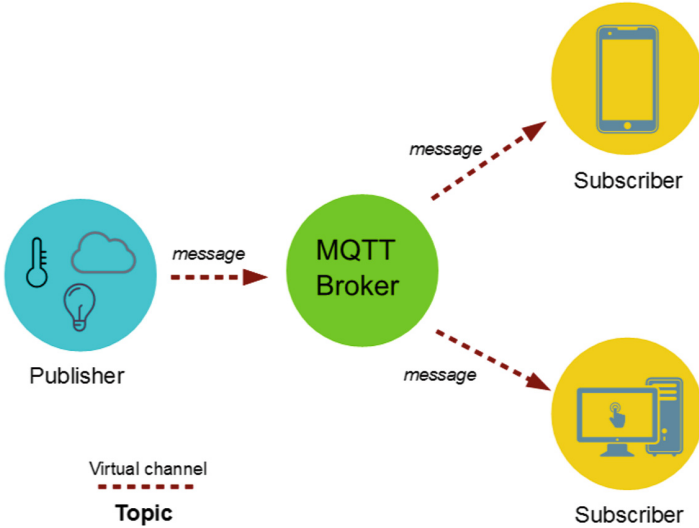


Fig. 1. MQTT topology diagram.

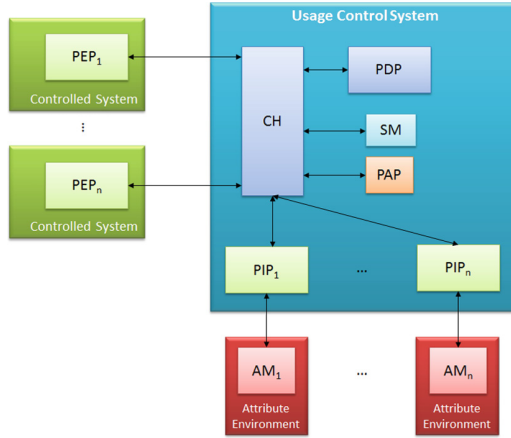
service, in which messages are delivered either once or not at all to their destination. No retransmission or acknowledgment is defined. QoS level 1 retransmits messages until they are acknowledged by the receivers. Hence, QoS level 1 messages may arrive multiple times at the destination because of the retransmissions, still multiple copies are not natively handled. QoS level 2 ensures not only the reception of the messages, but also that they are delivered only once.

We focus on the MQTT protocol since it is the most generic among the IoT protocols described, and libraries are available for all major IoT development platforms, like Arduino, for several programming languages (C, Java, PHP, Python, Ruby, Javascript) and for the two major mobile platforms (iOS and Android) [5]. The authentication to the Broker can be done by providing the following credentials [13]: Topic to be Subscribed on, Username and Password. The most known effort to add more security features in MQTT is SMQTT [18], but no solution is given to the policies that are followed by the information after it is delivered to the Subscribers. Our proposal addresses this problem, as long as the continuous control of Publishers/Subscribers on both authentication and access.

### 2.3 Usage Control

The UCON model extends traditional access control models. It introduces mutable attributes and new decision factors besides authorizations; these are obligations and conditions. Mutable attributes represent features of subjects, object, and environment that can change their values as a consequence of the operation of the system [6].





**Fig. 2.** Usage control framework diagram.

Since mutable attributes change their values during the usage of an object, UCON model allows to define policies which are evaluated before and continuously during the access. In particular, a usage control policy consists of three components: authorizations, conditions and obligations. Authorizations are predicates which evaluate subject and object attributes, and also the actions that the subject requested to perform on the object. Pre-Authorizations are evaluated when the subject requests to access the object, while Ongoing-Authorizations are predicates which are continuously evaluated while the access is in progress. Obligations are predicates which define requirements that must be fulfilled before the access (Pre-Obligations), or that must be continuously fulfilled while the access is in progress (Ongoing-Obligations). Conditions are requirements that evaluate the attributes of the environment. In this case too, Pre-Conditions are evaluated when the subject requests to access the object, while Ongoing-conditions are continuously evaluated while the access is in progress.

The continuous evaluation of the policy when the access is in progress is aimed at interrupting the access when the execution right is no more valid, in order to reduce the risk of misuse of resources. Hence, in the Usage Control model it is crucial to be able to continuously retrieve the updated values of the mutable attributes, in order to perform the continuous evaluation of the policy and to promptly react to the attributes change by interrupting those ongoing accesses which are no longer authorized.

The main blocks of UCON is the *Usage Control System (UCS)* surrounded by the *Controlled Systems* and the *Attribute Environment* are shown in the Fig. 2. The Controlled Systems are those components on which the UCON policy can be enforced. Each Controlled System communicates with the UCS issuing the request to access a resource by performing a specific operation on it. These components are the *Policy Enforcement Points (PEPs)*. For more information about UCON, readers can refer to [10].

UCS has its own components which are the following [16]:

*Policy Decision Point (PDP)*: This component takes as an input an access (usage) request and an access (usage) policy returning one of the following decisions: *Permit*, *Deny*, *Undetermined*.

PIPs communicate with the Attribute Environment through Attribute Managers (AMs) which are not part of the UCS [2].

*Policy Information Points (PIPs)*: These components retrieve attributes related to subject, object and environment of received access requests. Each PIP acts as the interface between the UCS and a specific Attribute Manager. Each PIP has custom implementation for each specific application, AM and the kind of attribute that should be retrieved.

*Session Manager (SM)*: This component is a database which stores all the active sessions, with the necessary information to perform policy reevaluations.

*Context Handler (CH)*: This component is the main core of the UCS, where it is responsible of routing messages among the various components. Firstly, it has to forward the access request to the various PIPs for attribute retrieval, then the complete access to the PDP and as a result to return the decision to the PEP. Finally, it receives notification from PIPs when the value of an attribute changes, forwarding to the PDP the new value for policy reevaluation. UCON framework consists of the following actions [9]:

**TryAccess**: This function is invoked by the PEP to send to the UCS the request to perform an action or access a resource, to be evaluated against a policy. The UCS will respond with a Permit or Deny decision, eventually collecting the needed attributes from the PIPs. If the answer is Permit, this response is also containing the Session ID for the session that is about to start.

**StartAccess**: This function is invoked by the PEP having the SessionID as a parameter. This is the actual start of using the service requested. There is again evaluation from the PDP and after an affirmative response the CH confirms the session to the SM as active.

**RevokeAccess**: If a mutable attribute changes its value, the PIP sends it to the CH for reevaluation because it might change the policy decision. If this event occurs, the usage has to be revoked. The CH informs both PEP and SM that this session is revoked. On one hand, the SM keeps the session recorded but in an inactive state, whereas on the other hand the PEP blocks the usage to the resource.

**EndAccess**: This function is invoked when the usage of the resource terminates. When received by the UCS, it deletes the session details from the SM and communicates to the PIPs that the attributes related to that policies are not needed anymore, unless other sessions are using it.

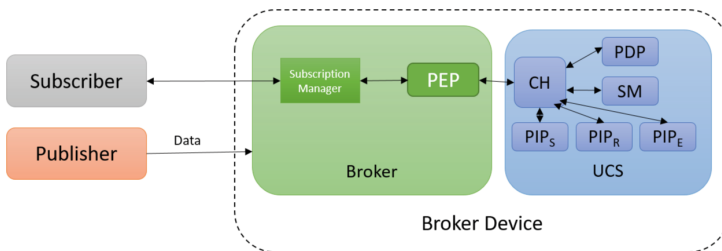
### 3 Introducing UCON in MQTT

In this section we present the proposed architecture, presenting first the model, then the operative workflow and the performed implementation.

#### 3.1 System Model

As previously mentioned, MQTT protocol is based on the Publish/Subscribe model, thus the entities participating to the protocol can act either as *Publishers* or *Subscribers*. Publishers could be sensors or other devices which collect and provide specific data, when available, periodically or even as a stream. Subscribers are instead entities that register to the broker to receive, when available, specific data or set of information grouped under a *Topic*. The *Broker* acts as middleware and coordinator, managing the subscription requests and dispatching data to Subscribers, when made available by prosumers.

The MQTT protocol supports ID and password-based authentication for both Publishers and Subscribers. The enforcement is performed on Broker’s side, which keeps track of the ID and authentication password of authorized Publishers and Subscribers. However, we argue that this authentication model is too simplistic and coarse grained, making impossible to check the right to access information over time. In fact, once a Subscriber has been authorized, the subscription remains valid until the Subscriber explicitly invokes an `unsubscribe` for the topic(s) it was registered for. The same goes for Publishers which keeps the right to publish continuously or on demand, till they have valid credentials. In real applications, several features might imply a condition for a subscription to decay, or for a publication to be denied. Detected Publisher malfunction or corruption, conditions on time spans in which a subscription should be allowed, and Subscriber reputation, are just few examples of aspects on which a more complex policy should be enforced. To be able to enforce policies with similar conditions to the aforementioned ones, and to have the possibility of revoking a subscription, usage control has been added to the MQTT logical architecture.



**Fig. 3.** UCON implementation in MQTT.

In Fig. 3 we depict the logical architecture of the proposed framework. As shown, the UCS is physically integrated in the Broker Device, i.e. the physical

machine that is hosting the Broker software, which enables the MQTT protocol. It is worth noting that we consider in this example three abstract PIPs, which are conceptually grouping the PIPs reading attributes related to the subject ( $PIP_S$ ), to the resource ( $PIP_R$ ) and to the environment ( $PIP_E$ ). The PEP is (partially) embedded in the broker, to dynamically control the subscription events. In particular, the PEP will intercept the subscription events and interact directly with the Broker subscription manager, deleting and inserting the entries for Subscribers from the list of authorized ones, according on the UCS decision. In such a way, the PEP ensures that no Subscribers can register by avoiding the enforcement of the usage control policy. Since the PEP is embedded in the Broker, the proposed architecture remains compatible with any implementation of MQTT Subscribers. The only requirement is that the Subscriber is configured to access with username and password, otherwise the connection will be refused by the Broker.

### 3.2 Operative Workflow

In Fig. 4, we report the envisioned workflow. For the sake of simplicity, we will consider a simple system made out of a Broker and a single Publisher and Subscriber.

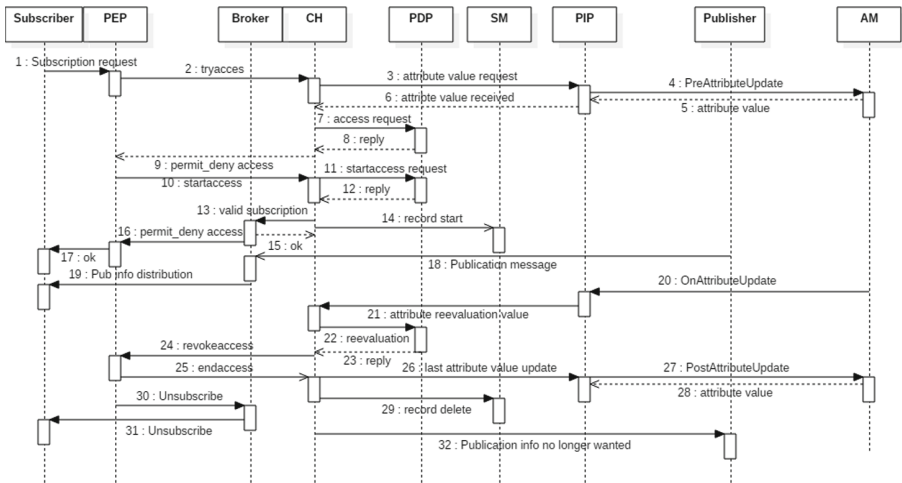


Fig. 4. Full workflow sequence diagram.

The workflow is initiated by a subscription request from the Subscriber to the Broker. This request is intercepted by the PEP, which interprets it, so as to take the credentials of the Subscriber that are needed in order to create and send the request to the UCS for evaluation. Hence the PEP invokes the `TryAccess` sending to the CH request and policy. The request is eventually filled

by attributes retrieved through the PIP, then is sent, together with the policy, to the PDP for evaluation, which should return a Permit or Deny decision. In case of Deny, the subscription request is dropped and the Subscriber will be notified, as if a wrong username/password has been inserted. In case of Permit, the Session Manager (SM) creates the session and sends its ID to the PEP (via the CH) which is informed about this decision and performs the **StartAccess**. Supposing a permit decision has been received, the Broker informs the Subscriber about the successful subscription and starts to send data related to the topic when available, eventually stimulating Publishers in an idle state.

To illustrate the **revoke** workflow, we suppose that one of the attributes relevant for the Subscriber policy changes its value (**OnAttributeUpdate**). This causes the PIP to send this new attribute to the CH that forwards it to the PDP for reevaluation. Supposing that the value of this attribute leads to a conclusion that this session must be revoked (Deny decision), the CH invokes the **RevokeAccess** on the PEP, also informing the Subscriber that the access is no longer granted (**RevokeAccess**). The termination of the access could happen also if the Subscriber is no longer interested to the data, invoking the **Unsubscribe**. The unsubscribe triggers the PEP to send an **EndAccess** to the CH. The latter informs the PIP to take the last value of the attribute (**PostAttributeUpdate**). Also the UCS informs the Broker that the Subscriber is no longer subscribed and forces the unsubscription. Moreover, the SM is also informed that this session is over so that the record should be archived or deleted. Finally, if this Subscriber is assumed to be the only one that was interested to the Publisher, the Broker informs him to stop data publication due to fact that there is no more any interest from any Subscriber.

We point out that the simplification of considering a single Publisher/Subscriber does not harm generality. In fact, the protocol is not modified and multiple Subscribers/Publishers do not introduce any additional criticalities, since concurrency is natively supported in both the UCS and MQTT.

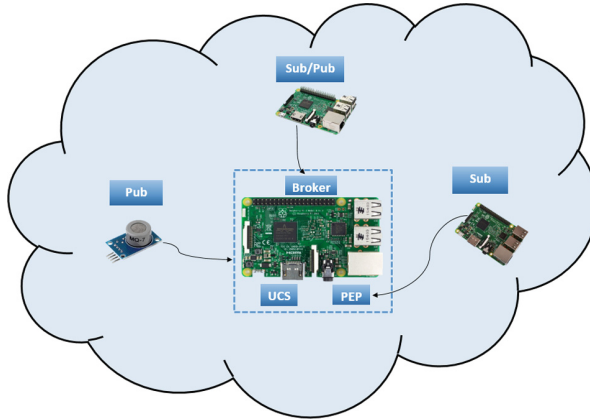
### 3.3 Implementation

As previously mentioned the UCS is a Java-based configurable framework, easy to integrate in any system with a Java runtime environment. The software used to implement the Broker is the open source MQTT Broker Moquette<sup>4</sup>. Though not largely used as the Mosquitto<sup>5</sup> Broker, Moquette is easier to integrate with the UCS framework, since they are based on the same programming language. The Broker has been partially modified to include in it the PEP functionalities. In particular, the subscription request is intercepted by hooking the subscription handling method, as to invoke **TryAccess** and **StartAccess** and waiting results before allowing or denying the subscription. If a Deny decision is received, the Broker will return a *wrong user/password* message to the subscriber.

<sup>4</sup> <https://github.com/andsel/moquette>.

<sup>5</sup> <https://mosquitto.org>.

If there is a policy violation, the `RevokeAccess` is invoked. Hence, the PEP calls the `Unsubscribe` function so as to prevent the Subscriber from receiving messages, while the `EndAccess` is invoked to remove the session details on the UCS side.



**Fig. 5.** Testbed logical representation.

In Fig. 5, is depicted the architecture of our testbed. In one Raspberry (central in Fig. 5) we run the Broker which includes the PEP, and the UCS as JARs. The code of the Subscriber<sup>6</sup> and the Publisher<sup>7</sup> were running unmodified in different Raspberries. Furthermore, additional tests have been performed by having the Subscriber host in an Android application called *MyMQTT*, which can be accessed through Google Play. Hence, the Subscriber code can be almost completely executed in the same device of the Publisher. Moreover, it is or the latter can be a small sensor that gives the data to the Broker as shown in Fig. 5. Since the framework is general, none of these configurations affects the functionality or requires any modifications to the framework.

## 4 Results

To demonstrate the viability of the proposed approach, the overhead introduced by usage control has been measured in a simulated and in a real environment. The framework has been tested in two different environments: the first one is a virtual machine installing Ubuntu 16.04 64-bit, equipped with an Intel i7-6700HQ with 8 cores enabled, 8 GB-DDR4 RAM running in 2133 MHz, the second one is a

<sup>6</sup> [https://github.com/pradeesi/MQTT\\_Broker\\_On\\_Raspberry\\_Pi/blob/master/subscriber.py](https://github.com/pradeesi/MQTT_Broker_On_Raspberry_Pi/blob/master/subscriber.py).

<sup>7</sup> [https://github.com/pradeesi/MQTT\\_Broke\\_On\\_Raspberry\\_Pi/blob/master/publisher.py](https://github.com/pradeesi/MQTT_Broke_On_Raspberry_Pi/blob/master/publisher.py).

Raspberry Pi 3 with a Broadcom ARMMv7 Quad Core Processor running on 1.2 GHz and 1 GB of LPDDR2 RAM on 900 MHz, running official Raspbian as operative system. The Publisher and the Subscriber were installed in two other Raspberries.

The complete set of results is reported in Table 2. All values have been extracted as the average times computed on 10 runs of the framework in every setting. The first column describes the title of the timings which are all described in milliseconds. The second column describes the timings when the Raspberries are used, and the third one the scenario where we used the Desktop-PC.

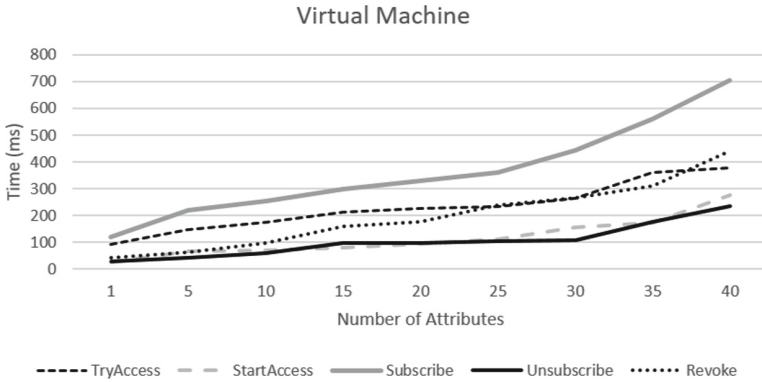
**Table 2.** Result comparison

Event timings (ms)	Raspberry	Desktop
Total tryaccess time	770	91
Total startaccess time	169	26
Total subscription time	969	121
UCON subscription part	939	118
MQTT subscription part	30	3
Total endaccess time	211	27
Unsubscribe time with UCON	213	27
Unsubscribe time without UCON	2	0
Revoke duration in broker	216	27
Revoke duration on UCON	455	41

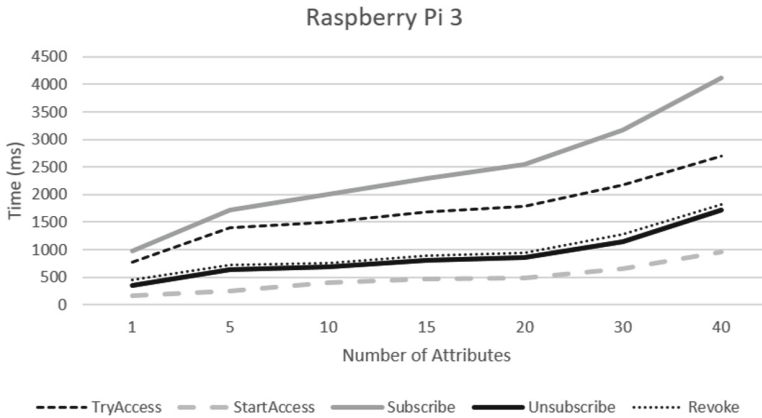
In Table 2, there are reported the detailed timings, considering a policy with a single attribute. In Figs. 6 and 7 are reported the performance variation at the increase of the number of attributes used in the policy. As shown, the timing behavior is almost linear to the amount of attributes, which is expected, due to the longer time needed to collect a larger number of attributes and for the evaluation performed by the PDP. However, in the real case, even considering 40 attributes, the timings are still acceptable for most of applications. Moreover, it is worth noting that policies with a large number of attributes such as 40 are quite unusual [15].

As expected, the low computational power of the Raspberry alongside the existence of a real network among the MQTT components, explains the longer timings than in the simulated environment. However, also considering a limited amount of attributes which is usual as mentioned above the overhead is slightly bigger than 1 s.

Considering the subscription time, we see that there is some overhead caused by UCS. This is not considered as a constraint because, since the Broker provides a buffer, we can still send all the published messages between the time of the request and the actual acceptance of the Subscriber. This causes no packet loss



**Fig. 6.** Timings on the simulated testbed.



**Fig. 7.** Performance on the real testbed.

to the Subscriber and high QoS. Furthermore, the most significant time is the one of the revocation. This time is in fact the actual time in which the policy is violated and should be minimized. As shown, this time is equivalent to 216 ms in the real use case and 27 ms in the Virtual Machine, considering a policy with a single attribute. For several applications, this time can be considered as negligible. As shown, the time between a non-valid value is taken and revocation of the access is very small.

Finally, it is worth mentioning that in the ongoing phase, i.e. after a successful `StartAccess`, no delay is introduced by UCON while delivering messages to the Subscribers independently also of the number of attributes.



## 5 Related Work

IoT is a paradigm which includes applications spanning from e-health to industrial controls. IoT architectures are distributed targeting on constrained devices. The different nature of these devices makes introduction of security mechanisms very difficult, especially when there exists the requirement of dynamic policy (re)evaluation. Although there exist applications of UCON in GRID [16] and Cloud [2] systems, alongside another one Android mobile devices [11], there is only one targeting on IoT [15], where the authors present a modified version of the standard usage control framework, called *UCIoT* that aims to bring UCON on IoT architectures. The architecture is designed to be seamless, configurable and dynamic. However, the authors did not consider any specific IoT protocol. The integration with the MQTT extension we are proposing and further evaluation, could be considered a valuable extension.

In [19], the authors present EventGuard in order to secure generally Publish/Subscribe overlay services. EventGuard is a dependable framework and a set of defense mechanisms for securing such a service. It comprises of a suite of guards to enhance security. But their solution does not target on MQTT but general in these type of protocols which means it is not targeting on constrained devices and protocols for IoT.

The authors of [12], propose a solution to securing Smart Maintenance Services. Their goal is to proactively predict and optimize the Maintenance, Repair and Operations (MRO) processes carried out by a device maintainer for industrial devices deployed at the customer side. They focus on the MQTT routing information asset and they define two elementary security goals regarding the client authentication. Their solution is based on Transport Layer Security (TLS) which is already a basic feature of the protocol. They proposed on how to use it more efficiently as a hardware element. Although they claim that the performance impact is not significant, the adoption of an extra hardware component might be critical in the constrained environment of IoT.

The most significant effort that targets in securing MQTT is a variation of it, called SMQTT [18]. It adds a security feature that is augmented to the existing MQTT protocol based on Key/Ciphertext Policy-Attribute Based Encryption (KP/CP-ABE) using lightweight Elliptic Curve Cryptography. This type of lightweight Attribute Based Encryption, produces extra overhead caused by the time and the computational power and is significant in the constrained environment of IoT. Moreover, our solution, has the advantage of using computational power only in the Broker which by default has sufficient computational power compared to Publishers and Subscribers. However, their implementation needs specific Publishers and Subscribers in order to decrypt the data whereas our solution works silently without being noticed by them and can work with any type of Publishers or Subscribers.

## 6 Conclusion

Current security mechanisms for IoT protocol are mainly focused on ensuring standard security properties such as message confidentiality and integrity, together with authentication. To the best of our knowledge, up to now the efforts for policies enforcement, which would ensure much more flexible, expressive and effective properties, are still quite limited. In this paper we have presented a first preliminary effort to increase the security of the MQTT protocol, by enabling the dynamic enforcement of usage control policies. We have presented a general methodology which allows to integrate UCON in a seamless way, without requiring protocol modifications. A real implementation has been presented, with performance evaluation to demonstrate the viability of the approach.

As future work, we plan to test the presented framework on a larger testbed with a larger number of attributes for the definition and enforcement of more complex policies, with a possible evaluation in a real applicative setting. Furthermore, we point out that the applied methodology can be easily extended to other IoT application protocols, where the benefits of integration are worth to be investigated in future works.

**Acknowledgments.** This work has been partially funded by EU Funded projects H2020 C3ISP, GA #700294, H2020 NeCS, GA #675320 and EIT Digital HII on Trusted Cloud Management.

## References

1. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., Ayyash, M.: Internet of things: a survey on enabling technologies, protocols, and applications. *IEEE Commun. Surv. Tutorials* **17**(4), 2347–2376 (2015, fourthquarter)
2. Carniani, E., D’Arenzo, D., Lazouski, A., Martinelli, F., Mori, P.: Usage control on cloud systems. *Future Gener. Comput. Syst.* **63**(C), 37–55 (2016)
3. Chen, D., Varshney, P.K.: QoS support in wireless sensor networks: a survey (2004)
4. Colitti, W., Steenhaut, K., De Caro, N., Buta, B., Dobrota, V.: Evaluation of constrained application protocol for wireless sensor networks. In: 2011 18th IEEE Workshop on Local Metropolitan Area Networks (LANMAN), pp. 1–6, October 2011
5. Collina, M., Corazza, G.E., Vanelli-Coralli, A.: Introducing the QEST broker: scaling the IoT by bridging MQTT and REST. In: 2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 36–41, September 2012
6. Faiella, M., Martinelli, F., Mori, P., Saracino, A., Sheikhalishahi, M.: Collaborative attribute retrieval in environment with faulty attribute managers. In: 2016 11th International Conference on Availability, Reliability and Security (ARES), pp. 296–303, August 2016
7. Fysarakis, K., Askoxylakis, I., Soultatos, O., Papaefstathiou, I., Manifavas, C., Katos, V.: Which IoT protocol? Comparing standardized approaches over a common M2M application. In: 2016 IEEE Global Communications Conference (GLOBECOM), pp. 1–7. IEEE (2016)

8. Karagiannis, V., Chatzimisios, P., Vázquez-Gallego, F., Alonso-Zrate, J.: A survey on application layer protocols for the internet of things. *Trans. IoT Cloud Comput.* **1**(1), 11–17 (2015)
9. Karopoulos, G., Mori, P., Martinelli, F.: Usage control in SIP-based multimedia delivery. *Comput. Secur.* **39**, 406–418 (2013)
10. Lazowski, A., Martinelli, F., Mori, P.: Survey: usage control in computer security: a survey. *Comput. Sci. Rev.* **4**(2), 81–99 (2010)
11. Lazowski, A., Martinelli, F., Mori, P., Saracino, A.: Stateful data usage control for android mobile devices. *Int. J. Inf. Secur.* **16**(4), 345–369 (2017)
12. Lesjak, C., Hein, D., Hofmann, M., Maritsch, M., Aldrian, A., Priller, P., Ebner, T., Rupprechter, T., Pregartner, G.: Securing smart maintenance services: hardware-security and TLS for MQTT. In: 2015 IEEE 13th International Conference on Industrial Informatics (INDIN), pp. 1243–1250, July 2015
13. Locke, D.: MQ telemetry transport (MQTT) v3. 1 protocol specification. IBM developerWorks Technical Library (2010)
14. Luzuriaga, J.E., Perez, M., Boronat, P., Cano, J.C., Calafate, C., Manzoni, P.: A comparative evaluation of AMQP and MQTT protocols over unstable and mobile networks. In: 2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC), pp. 931–936, January 2015
15. La Marra, A., Martinelli, F., Mori, P., Saracino, A.: Implementing usage control in internet of things: a smart home use case. In: 2017 IEEE Trustcom/BigDataSE/ICCESS, Sydney, Australia, 1–4 August 2017, pp. 1056–1063 (2017)
16. Martinelli, F., Mori, P.: On usage control for GRID systems. *Future Gener. Comput. Syst.* **26**(7), 1032–1042 (2010)
17. NIST: MQTT and the NIST Cybersecurity Framework Version 1.0 (2014). <http://docs.oasis-open.org/mqtt/mqtt-nist-cybersecurity/v1.0/cn01/mqtt-nist-cybersecurity-v1.0-cn01.pdf>. Accessed 22 Jan 2017
18. Singh, M., Rajan, M.A., Shivraj, V.L., Balamuralidhar, P.: Secure MQTT for internet of things (IoT). In: 2015 Fifth International Conference on Communication Systems and Network Technologies, pp. 746–751, April 2015
19. Srivatsa, M., Liu, L.: Securing publish-subscribe overlay services with EventGuard. In: Proceedings of the 12th ACM Conference on Computer and Communications Security, CCS 2005, pp. 289–298. ACM, New York (2005)
20. Talaminos-Barroso, A., Estudillo-Valderrama, M.A., Roa, L.M., Reina-Tosina, J., Ortega-Ruiz, F.: A machine-to-machine protocol benchmark for eHealth applications use case: respiratory rehabilitation. *Comput. Methods Programs Biomed.* **129**, 1–11 (2016)
21. Thangavel, D., Ma, X., Valera, A., Tan, H.-X., Tan, C.K.-Y.: Performance evaluation of MQTT and CoAP via a common middleware. In: 2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), pp. 1–6. IEEE (2014)
22. Villari, M., Celesti, A., Fazio, M., Puliafito, A.: AllJoyn Lambda: an architecture for the management of smart environments in IoT. In: 2014 International Conference on Smart Computing Workshops, pp. 9–14, November 2014

# Using JSON to Specify Privacy Preserving-Enabled Attribute-Based Access Control Policies

Que Nguyet Tran Thi<sup>(✉)</sup>, Tran Khanh Dang, Huy Luong Van,  
and Ha Xuan Son

Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam  
{ttnguyet, khanh, 51301464}@hcmut.edu.vn,  
hxsonhxson@ctuet.edu.vn

**Abstract.** With the growth of big data systems and ubiquitous computing, privacy has become a critical issue in security research. Purpose based access control model is a common approach for privacy preserving in data access for database management systems. However, previous works that are based on purposes ranging from the table level to the data cell level and that are extended with role-based access control model inherently suffer from the problem of role explosion and cumbersomeness in context aware policy specification. Besides, NoSQL databases have recently become increasingly popular as data platforms for big data and real-time web applications. Due to the simplicity in design but effectiveness in horizontal scaling and performance, using NoSQL databases are a better alternative approach in comparison with traditional relational databases. However, the lack of a fine-grained access control system with data privacy protection is one of the most important considerations in NoSQL databases. In this paper, we address this issue by proposing and implementing a comprehensive mechanism for enforcing document store attribute-based security policies together with an improved data privacy protection mechanism in the fine-grained level. We use Polish notation for modeling conditional expressions which are the combination of subject, resource, and environment attributes so that privacy policies are flexible, dynamic and fine grained. Furthermore, privacy rules are constrained not only by access and intended purposes but also by subject, resource, and environment attributes as well as levels of data disclosure. The experiments have been carried out to illustrate the execution time of evaluating access control policies and privacy policies in various data sizes.

**Keywords:** Attribute based access control model  
Purpose based access control model · Privacy protection  
Privacy preserving · JSON

## 1 Introduction

Since the rapid development of large scale, open and dynamic systems, the shortcomings of traditional access control models (e.g. Discretionary Access Control (DAC), Mandatory Access Control (MAC), Role-based Access Control (RBAC)) have gradually revealed such as only applied for closed systems, role explosion, complexity

in compulsory assignments between users, roles, and permissions, and inflexibility in specifying dynamic policies and contextual conditions [1]. Attribute-based Access Control (ABAC) models have been recently investigated and considered as one of three mandatory features for future access control systems [2]. By National Institute of Standards and Technology (NIST), ABAC is defined as an access decision is permitted only if the request satisfies conditions on attributes of subject, resource and environment specified in policies [3].

However access control systems are successful in preventing unauthorized accesses, they are ineffective in privacy protection for a large, decentralized system like Internet of Things or distributed systems. In these days, it is possible to infer sensitive information from publicly available information. Privacy is a major concern in both of research and industrial fields due to dissemination of personal and sensitive data without user control, especially in mobile and ubiquitous computing applications and systems. In [4], privacy is defined as the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others. Most previous studies have considered privacy protection in access control models as constraints on purpose of data usage. Although it has developed since 2000s, this research area has recently drawn many interests due to the importance of privacy preserving in the new era of Internet of Things or Big data [5–9]. However, to the best of our knowledge, no research has allowed for setting levels of data disclosure (from the most generalization to the highest precision) associated with purposes and contextual conditions and addressing with NoSQL data model.

Besides, integrating NIST ABAC into purpose-based access control model has not been examined. Based on these considerations, we propose an access control model with 2-step authorization consisting of access authorization and privacy authorization. Access policies are designated for controlling subjects to access data and specified by security administration. Privacy policies are designated for protecting privacy of objects and specified by data owners, data providers, or security administrators depending on applications. All policies can contain conditional expressions on subject, action, resource, and environment attributes so that we can specify privacy policies more dynamically and flexibly. Besides, they are specified by JavaScript Object Notation which is a lightweight data interchange format and popularly used in document store NoSQL database such as MongoDB.

The rest of the paper is organized as follows. Section 2 gives a brief survey of related works. Section 3 presents the policy structure and policy decision mechanism. Section 4 discusses our experimenting results. Concluding remarks and future work are given in Sect. 5.

## 2 Related Work

The development of Information Technology, especially in the age of Big Data and the Internet of Thing environments causes the role explosion problem and increases the complexity in permission management in the Role-based Access Control (RBAC) models, which have been dominant for a long time [10, 11]. An emerging interest in addressing these problems is Attribute-based Access Control (ABAC) models, which

can be adaptable with large, open and dynamic environments [2]. In the common approach of ABAC, according to NIST standard [3], authorization decision is based on rules that simultaneously specify a set of conditions on numerous attributes such as subjects, objects, actions and environments for a certain valid permission. Extensible Access Control Markup Language 3.0 (XACML) [12] for enforcing security policies in advanced access control models such as [13, 14, 22] has been considered as a predecessor of ABAC. However, in XACML policies, every operation on attributes even trivial conditions such as comparison requires function and data type definitions. This has caused the verbosity and difficulty in the specification of policies. Moreover, it obtains disadvantages of XML data interchange format when integrating into current and future applications and systems [15]. Nowadays, JavaScript Object Notation (JSON), which is a more light-weight data interchange format than XML, has been widely used in Web 2.0 applications and NoSQL databases [15]. Therefore, in our approach, we use JSON to express attribute based security policies and build a new access control mechanism to enforcing JSON security policies.

In another aspect, policy decision point in privacy preserving access control models, namely purpose based access control models (PBAC), depends on the relationship between access purpose and intended purposes of data objects ranging from the level of tables to the data cells [5, 6, 16, 17]. In the beginning, Byun et al. [5, 6] proposed the model with two types of allowable and prohibited intended purposes. It was then extended with an additional purpose, i.e. conditional intended purpose [17]. Several works have been conducted on enhancing this model by combining RBAC (e.g., [17–21]), that implements with relational database management systems (DBMSs) with the technique of SQL query rewriting [7] and integrating with MongoDB [9]. Recently, action-aware with indirect access and direct access has also been considered in policies [8]. Nevertheless, in the literature, data privacy through allowable, conditional and prohibited purposes is still limited and not dynamic. Moreover, to provide fine grained and flexible privacy protection, instead of expressing as additional fields of data, purposes should be modelled in a new kind of policy, called as privacy policy. Our proposed model supports these privacy policies not only with context aware conditions but also with purposes associated with various data disclosure levels.

In summary, our work gives a novel approach by using JSON to specify attribute based policies and proposing a new access control mechanism that can enforce security policies according to the principle of ABAC and PBAC models for privacy preserving at the more fine grained levels.

### 3 Using JSON to Specify Privacy Preserving Enabled Attribute-Based Access Control Policies

In this section, we describe the structure of policies and action steps in our mechanism. When subject accesses an object, the authorization process is carried out through two stages called as 2-stage authorization:

- First stage: the authorization system verifies that the request is legitimate with rights for the subject to access data by *access control policies*.

- Second stage: the request is transfer to this stage for checking privacy compliance based on privacy policies. *Privacy policies* indicate how much to hide data or generalize data.

**Access Control Policies:** Contain policies which are used to determine whether a subject can access resources. The decision is made based on rules inside policies which are the Boolean expressions evaluated by user’s defined function, subject, resource, environment attribute.

**Privacy Policies:** Contain policies which are used to determine whether access data should be shown, hidden or generalized. The privacy protection engine is also based on rules which are the Boolean expressions evaluated by user’s defined function, subject, resource, environment attribute.

### 3.1 Conditional Expression

Modeling rule expressions in policies is one of the most challenge things because rules ensure that every access to a system and its resource must be valid. To deal with complex conditions, we consider each single expression and its parameter as a function structure and combine them by using the tree structure.

Using this approach, we visualize a complex conditional expression by a tree function structure where each parent node is the name of function and child node is a constant value, resource identifier or another function name.

For example, consider the following rule:

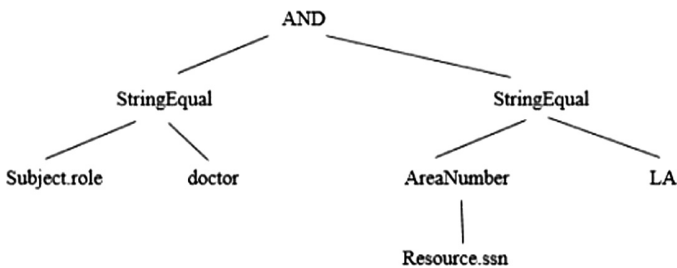
*“A user with role doctor cannot view those users who has Area Number of SSN field which equals LA.”*

The conditional expression is demonstrated as below.

Exp = *“StringEqual (Subject.role, ‘doctor’) AND StringEqual (AreaNumber (Subject.ssn), ‘LA’)”*.

It is visualized as the expression tree in Fig. 1 and easily converted into JSON document according to this tree structure.

Both of access control policy and privacy policy can contain the element of conditional expression written in JSON.



**Fig. 1.** An example for expressing conditional expression in tree structure

### 3.2 Policy Structure

**Access Control Policy Structure.** In our system, an access control policy includes rules. Each rule defines a conditional expression that is modeled in the above function tree structure. The rule returns a value specified in the element *Effect* if the condition is true. To avoid conflicts between policies and rules, the combining rule algorithms such as permit override, deny override, etc. are applied into the policy set and policies. The solution for using combining rule algorithm is inherited from XACML. The relationship diagram between policies and rules are illustrated in the Fig. 2.

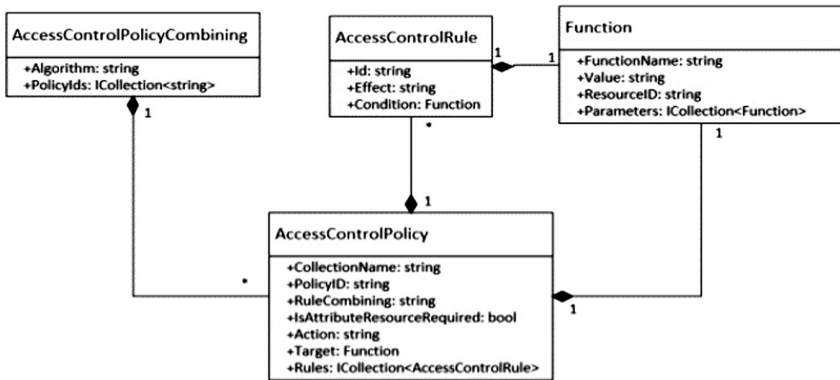


Fig. 2. Access control policy class diagram

**Privacy Policy Structure.** To specify policies for privacy protection, access purpose is modeled as an attribute of environment and intended purpose is considered as an attribute of data objects in the resource content. Moreover, in our solution, privacy policies have the same structure of access control policies which can express complex attribute based policies in the simple way. Besides, in privacy rules, a special field “*field\_effects*”, which has an array type, describes the list of data disclosure levels for each field of JSON data constrained by these rules (Fig. 3). For example, if we want to hide the city of address, the address data field will be specified in the component “*field\_effects*” of the corresponding privacy policy.

Each element in “*field\_effects*” has the following components:

“name”: the path to the single value field.

“effect\_function”: This field contains “X.Y” value where X is privacy domain, and Y is name of privacy function in that domain. The default value is DefaultDomain. Show.

To demonstrate the privacy policy structure, the below example indicates Policy\_1 has a rule 10001. This rule describes the body data field of email documents is processed by the privacy function SubHide200 which removes the body value from the 200th character.



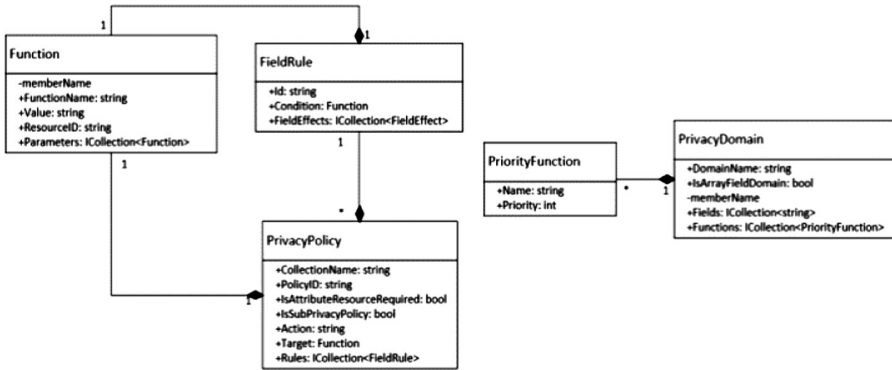


Fig. 3. Privacy policy class diagram

```

_id": "Policy 1",
"Collection_name" : "emails",
"description" : "",
"target" : null,
"is_attribute_resource_required" : false,
"rules" : [{
  "identifier" : "1001",
  "condition" : null,
  "field_effects" : [
    {
      "name": "body",
      "effect_function": "DefaultDomain.SubHide200"
    }
  ]
}]
}]

```

By this approach, it can happen conflicts that many privacy rules are applied on the same data. The system can define many privacy functions for various domains which are stored in the PrivacyDomain collection. There is a default domain that contains two basic privacy functions: *Show* and *Hide*. The below example illustrates the privacy functions for the default domain is Hide and Show; and the DateTimeDomain contains four privacy functions: Hide, ShowYear, ShowMonthAndYear and Show. Each privacy function is assigned by a *priority value* to solve conflicts. The function, which has the highest priority value, is used to apply to the data item.

```

{
  "domain_name": "DefaultDomain",
  "fields": [],
  "is_sub_policy": false,
  "hierarchy": [
    {"name": "Hide", "priority": 100},
    {"name": "Show", "priority": 0}
  ]
}
{
  "domain_name": "DateTimeDomain",
  "fields": ["Employee.personal"],
  "is_sub_policy": false,
  "hierarchy": [
    {"name": "Hide", "priority": 100},
    {"name": "ShowYear", "priority": 10},
    {"name": "ShowMonthAndYear", "priority": 1},
    {"name": "Show", "priority": 0}
  ]
}

```

In assumption, the *personal* field of Employee contains three subfields: name, birthdate, and ssn. In the privacy policies, the system applies privacy functions to these fields. Thus, we receive the conflict privacy functions on each subfield which is retrieved from the applicable privacy policies and the corresponding final privacy functions processed based on their priorities (Table 1).

**Table 1.** An example of conflict privacy functions

Fields	Conflict privacy functions	Result
personal.name	DefaultDomain.Show	DefaultDomain. Show
personal. birth_date	DateTimeDomain.ShowMonthAndYear, DateTimeDomain.ShowYear, DefaultDomain.Show	DateTimeDomain. ShowYear
personal.ssn	SsnDomain.AreaNumber, SsnDomain.SerialNumber	SsnDomain. AreaNumber.

## 4 Experiment

In this section, we evaluate the processing time of the policy decision mechanism for both of access control and privacy policies. We have used the Enron Corpus dataset with 1.5 GB of MongoDB data comprising with 517,425 emails. In our experiments, the prototype has implemented on the platform.NET core 2.0 and MongoDB 3.2 and been deployed on an Intel Xeon E5-2620 2 GHz with 8 GB of RAM. We build the scenarios on the various size of dataset such as 50 K, 100 K, 200 K, 300 K, and 400 K emails and enforce that the processes must read data items in the dataset according to the selectivity values 75% and 100% (defined as the percentage of documents and fields can be accessed). Firstly, we update the documents in the datasets by replacing the *From* field with “*Alice@gmail.com*” and allowing the *filename* field to be null with the ratio 25%. Then, the experiments have been carried out with the request “*find all emails with the sender from Alice@gmail.com*”. Therefore, all data items have been selected and browsed in the cases of evaluating field level policies. We analyze the processing time of the policy decision mechanism with policies specified in the following situations (1) one access control policy AP at the field level with the selectivity 75%, (2) one privacy policy P01 at the field level with the selectivity 75% and 100%, and (3) one privacy policy P02 with the default function of hiding the data value at the field level in the embedded document with the selectivity 75% and 100%. We divided into two experiments: (1) each policy separately and (2) the combination between AP and P01, AP and P02, and all policies.

In the remainder section, we illustrate our policies in the above scenarios and present the experimenting results.

### 4.1 Policy Specification

In our sample policies, the first one is an access control policy which indicates the right of Alice which is reading emails if the file name is not null. Meanwhile, the privacy

policies indicate that some data field will be generalized in the result set however the access request is permitted. They are specified as below:

- AP: Alice can read emails if the file name is not null.
- P01 with the selectivity 75%: All body of emails is hidden completely if the file name is not null.
- P01 with the selectivity 100%: All body of emails is hidden from the 1st character.
- P02 with the selectivity 75%: All names of To address of emails is replaced by x. Example, abc@gmail.com becomes x@gmail.com.
- P02 with the selectivity 100%: All names of To address of emails is hidden completely if the filename is not null.

### 4.2 Results

Figure 4 indicates the processing time in the cases of a single policy in the database with the difference selectivity values (0.75 and 1). Meanwhile, Fig. 5 illustrates the processing time in the cases of the combination of policies. The results show that there

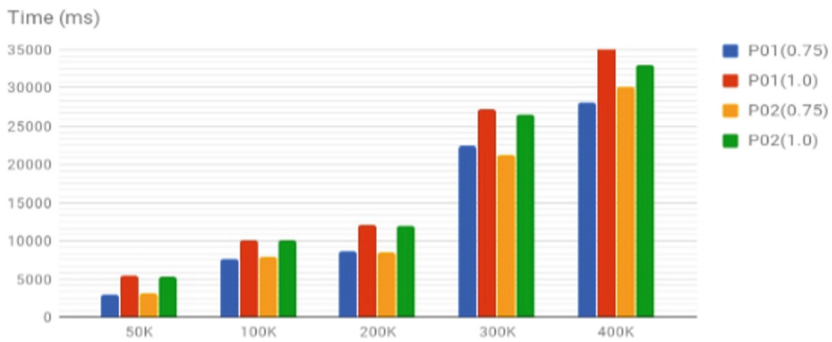


Fig. 4. The processing time in the single policy cases

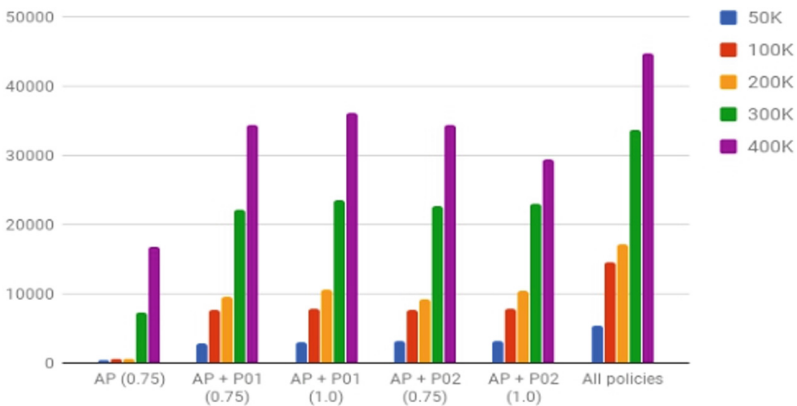


Fig. 5. The processing time in the combination cases

has been a significant change when the data size increases more than 300 K. Besides, there are no much difference in processing time between the various kinds of policies. Moreover, Fig. 5 also shows that there is no much increase from the combination of two policies to all policies.

## 5 Conclusion

In this research, we have proposed a comprehensive mechanism for enforcing attribute-based security policies stored in JSON document together with the feature of data privacy protection in the fine-grained level. We have used Polish notation for modeling conditional expressions which are the combination of subject, resource, and environment attributes so that the policies are flexible, dynamic and fine grained. Through the proposed flexible structure for privacy policies, it can be evaluated not only by access and intended purpose but also by subject, resource, environment attributes. We also support to define data disclosure levels for privacy protection constrained at various scopes in JSON document such as collection, document, field and embedded document. Due to the size of the paper, our architecture and prototype cannot be described in details. In future, we will improve our solution to work with other NoSQL database document stores. Besides that, we will improve the performance by applying heuristic functions when evaluating conditional expressions.

**Acknowledgements.** This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number C2017-20-11.

## References

1. Bertino, E., Ghinita, G., Kamra, A.: Access control for databases: concepts and systems. *Found. Trends® Databases* **3**(1–2), 1–148 (2011)
2. Sandhu, R.: The future of access control: attributes, automation, and adaptation. In: Krishnan, G., Anitha, R., Lekshmi, R., Kumar, M., Bonato, A., Graña, M. (eds.) *Computational Intelligence, Cyber Security and Computational Models. AISC*, vol. 246, p. 45. Springer, New Delhi (2014). [https://doi.org/10.1007/978-81-322-1680-3\\_5](https://doi.org/10.1007/978-81-322-1680-3_5)
3. Hu, V.C., Ferraiolo, D., Kuhn, R., Friedman, A.R., Lang, A.J., Cogdell, M.M., Schnitzer, A., Sandlin, K., Miller, R., Scarfone, K.: Guide to attribute based access control (ABAC) definition and considerations (draft). NIST Special Publication **800**(162) (2013)
4. Westin, A.F.: Privacy and freedom. *Washington Lee Law Rev.* **25**(1), 166 (1968)
5. Byun, J.W., Bertino, E., Li, N.: Purpose based access control of complex data for privacy protection. In: *Proceedings of the Tenth ACM Symposium on Access Control Models and Technologies*, pp. 102–110. ACM, June 2005
6. Byun, J.W., Li, N.: Purpose based access control for privacy protection in relational database systems. *VLDB J.* **17**(4), 603–619 (2008). <https://doi.org/10.1007/s00778-006-0023-0>
7. Colombo, P., Ferrari, E.: Enforcement of purpose based access control within relational database management systems. *IEEE Trans. Knowl. Data Eng.* **26**(11), 2703–2716 (2014)
8. Colombo, P., Ferrari, E.: Efficient enforcement of action-aware purpose-based access control within relational database management systems. *IEEE Trans. Knowl. Data Eng.* **27**(8), 2134–2147 (2015)

9. Colombo, P., Ferrari, E.: Enhancing MongoDB with purpose based access control. In: IEEE Transactions on Dependable and Secure Computing (2015)
10. Ferraiolo, D.F., Sandhu, R., Gavrila, S., Kuhn, D.R., Chandramouli, R.: Proposed NIST standard for role-based access control. *ACM Trans. Inf. Syst. Secur. (TISSEC)* **4**(3), 224–274 (2001)
11. Fuchs, L., Pernul, G., Sandhu, R.: Roles in information security—a survey and classification of the research area. *Comput. Secur.* **30**(8), 748–769 (2011)
12. Parducci, B., Lockhart, H., Rissanen, E.: Extensible access control markup language (XACML) version 3.0. OASIS Standard, pp. 1–154 (2013)
13. Le Thi, K.T., Dang, T.K., Kuonen, P., Drissi, H.C.: STRoBAC – spatial temporal role based access control. In: Nguyen, N.-T., Hoang, K., Jędrzejowicz, P. (eds.) ICCCI 2012. LNCS (LNAI), vol. 7654, pp. 201–211. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-34707-8\\_21](https://doi.org/10.1007/978-3-642-34707-8_21)
14. Thi, Q.N.T., Dang, T.K.: X-STROWL: a generalized extension of XACML for context-aware spatio-temporal RBAC model with OWL. In: 2012 Seventh International Conference on Digital Information Management (ICDIM), pp. 253–258. IEEE, August 2012
15. Nurseitov, N., Paulson, M., Reynolds, R., Izurieta, C.: Comparison of JSON and XML data interchange formats: a case study. *Caine* **2009**, 157–162 (2009)
16. Wang, H., Sun, L., Bertino, E.: Building access control policy model for privacy preserving and testing policy conflicting problems. *J. Comput. Syst. Sci.* **80**(8), 1493–1503 (2014)
17. Kabir, M.E., Wang, H.: Conditional purpose based access control model for privacy protection. In: Proceedings of the Twentieth Australasian Conference on Australasian Database, vol. 92, pp. 135–142. Australian Computer Society Inc., January 2009
18. Kabir, M.E., Wang, H., Bertino, E.: A role-involved conditional purpose-based access control model. In: Janssen, M., Lamersdorf, W., Pries-Heje, J., Rosemann, M. (eds.) EGES/GISP -2010. IAICT, vol. 334, pp. 167–180. Springer, Heidelberg (2010). [https://doi.org/10.1007/978-3-642-15346-4\\_13](https://doi.org/10.1007/978-3-642-15346-4_13)
19. Kabir, M.E., Wang, H., Bertino, E.: A conditional purpose-based access control model with dynamic roles. *Expert Syst. with Appl.* **38**(3), 1482–1489 (2011)
20. Ni, Q., Bertino, E., Lobo, J., Brodie, C., Karat, C.M., Karat, J., Trombeta, A.: Privacy-aware role-based access control. *ACM Trans. Inf. Syst. Secur. (TISSEC)* **13**(3), 24 (2010)
21. Ni, Q., Lin, D., Bertino, E., Lobo, J.: Conditional privacy-aware role based access control. In: Biskup, J., López, J. (eds.) ESORICS 2007. LNCS, vol. 4734, pp. 72–89. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-74835-9\\_6](https://doi.org/10.1007/978-3-540-74835-9_6)
22. Son, H.X., Tran, L.K., Dang, T.K., Pham, Y.N.: Rew-XAC: an approach to rewriting request for elastic ABAC enforcement with dynamic policies. In: 2016 International Conference on Advanced Computing and Applications (ACOMP), pp. 25–31. IEEE, November 2016

# Comprehensive Diversity in Recommender Systems

Tranos Zuva<sup>(✉)</sup> and Raoul Kwuimi

Department of ICT, Vaal University of Technology,  
Vanderbijlpark, South Africa  
tranosz@vut.ac.za

**Abstract.** The present age of digital information has presented a heterogeneous online environment which makes it a formidable mission for a noble user to search and locate the required online resources timely. Recommender systems were implemented to rescue this information overload issue. However, majority of recommender systems focus on the accuracy of the recommendations, leaving out other important aspects in the definition of good recommendation such as diversity and serendipity. This results in low coverage and long-tail items are often left out in the recommendations. In this paper, we present and explore a recommendation technique that ensures that comprehensive diversity is also factored-in in the recommendations. The algorithm adopts the second line of recommendation improvement whereby a recommendation list is re-ranked in such a way that it would include long-tail items. The results showed that the proposed algorithm is capable of giving a balanced list of recommendations in terms of accuracy and diversity.

**Keywords:** Recommender system · Re-ranking · Diversity · Techniques

## 1 Introduction

While accuracy is an important aspect in recommendations, diversity is as much important as well and it needs to be factored in also. Diversity ensures that there is coverage in recommended items in order to recommend even the long-tail items that might be of interest to the users. Diversity can be divided into two: individual and aggregate diversity. Individual diversity is the measure of average dissimilarity between all pairs of items recommended to individual user while aggregate diversity is the total number of distinct items recommended across all users [1]. High individual diversity does not necessarily imply high aggregate diversity [2].

Recommender Systems with higher aggregate diversity can benefit a number of business models [3]. This paper addresses both the types of diversities under the same umbrella “comprehensive diversity”. It will be referred to in short as a diversity unless otherwise explicitly stated. We present an algorithm that will ensure that while we still keep an eye on the accuracy we also ensure that diversity is enhanced.

## 2 Related Work

The classical approach of Recommender Systems to the problem domain is divided into two; the rating of unrated items in the whole item space and the ranking of candidate items to be recommended [1]. The rating of items is achieved by using some recommendation algorithms on the users' available information. There are a lot of algorithms employed by recommender systems to recommend items to the users [4, 5]. However, there are two most popular algorithms namely; Collaborative Filtering (CF) and Content-based algorithms (CB). These two algorithms are the foundation on which other algorithms are build. There are also hybrid algorithms build from both the Collaborative Filtering and Content-based algorithms [6] (Burke, 2007). The ranking is processed by finding the items that maximize the user's utility based on the predicted ratings and recommends them to the user. In most cases these recommender systems are mostly interested in providing accurate recommendations without taking into consideration users who are first movers thus consider only popular items for recommendation are mostly recommended. In order to balance the list of recommendations there is need to diversify the items.

### 2.1 Diversity in Recommender Systems

Diversity in recommender systems is an important factor which ensures that even the long-tail items with less information data which might be interesting to the active user are included in the recommendation list [2]. These can be very helpful in some online business models because they provide wider range of items other than only bestsellers which users are often capable of getting them by themselves [3]. There are two types of diversity that is individual diversity and aggregate diversity.

### 2.2 Individual Diversity

Individual diversity is the measure of average dissimilarity of items recommended to an individual [1]. It is measured from each user's recommendation list. The intension is to ensure that not too similar items are recommended to an individual user. If we take a music domain for an example, we would not want all the recommended songs to be from one artist.

Intra-list similarity metric is used to evaluate the list diversity. The average dissimilarity of all pairs of items is computed by the use of distance function  $d : I \times I \rightarrow R$  such that  $d(i, j)$  represents the dissimilarity between elements  $i$  and  $j: j \in I$ . This is illustrated in formula 1.

$$diversity - in - topN = \frac{2}{p(p-1)} \sum \sum d(i, j) \quad (1)$$

where  $i$  and  $j$  are the items in the recommendation list and  $i \neq j$  and  $p$  is the number of items in the recommendation list.

The assumption is that the distance function is symmetric, meaning  $d(i,j) = d(j,i)$  and  $p$  is the number of items in the recommendation list (Top-N). The function  $d(i,j)$  is derived from the similarity metric  $s(i,j)$  such that  $d(i,j) = 1 - s(i,j)$ .

### 2.3 Aggregate Diversity

Aggregate diversity is the total number of distinct items recommended across all users [1]. There are two lines of research attempting to enhance aggregate diversity. The first one computes the rating predictions based on the discussed filtering techniques such as Collaborative Filtering and then the recommendations are re-ranked using a specific technique to make way for long-tail items. The second approach which is rarely used targets the estimation process to try and come with correct predictions. The discovered algorithms to enhance diversity follow either the first line of approach or the second line. There a number of algorithms to address the diversity namely; Ranking-based techniques, graph theoretic approach and Hybrid User-based and Item-based CF.

### 2.4 Ranking-Based Techniques

Ranking based techniques were proposed by Adomavicius and Kwon [2] to improve the aggregate diversity in recommender systems. These techniques employ the Collaborative Filtering as their foundation on which they build their recommendations. The items ratings are predicted using CF and the recommendation list is ranked according to Standard ranking approach and then Item-popularity and finally parameterised ranking. These techniques work together and complement each other to improve aggregate diversity of recommended items.

### 2.5 Standard Ranking Approach

In this approach, the first step is to predict the unknown ratings using traditional techniques, then the predicted ratings are used to support the recommendation process. The user gets recommended a list of top N items selected according to some ranking criteria. The items with highest predicted rating are the ones being recommended to the user. This criterion of recommending highly rated items improves the accuracy of the predictions, but the diversity is compromised. The need to balance accuracy and diversity led to popularity based approach to compliment the standard ranking approach. The Item-popularity-based approach was proposed.

### 2.6 Item-Popularity-Based Approach

Item popularity based ranking works exactly like standard approach in prediction stage. They only differ when it comes to the recommendation stage. Item-popularity as the name suggests, considers the popularity of items before recommending them. That is, it ranks items according to their popularity from less popular to more popular. The popularity of an item is deduced from the number of total ratings the item has. The higher number of ratings is perceived to indicate that the item is known to a number of users.



This approach proved to improve the aggregate diversity of recommended items. However, the approach significantly compromises the accuracy. That is why another technique is needed to address the trade-off between accuracy and diversity. The parameterized ranking approach was introduced as a result.

### 2.7 Parameterized Ranking Approach

Parameterized ranking approach parameterize the other ranking approaches by introducing a ranking threshold  $T_R \in [T_H, T_{max}]$  (where  $T_{max}$  is the largest possible rating on the rating scale, e.g.,  $T_{max} = 5$  and  $T_H$  is the minimum acceptable threshold value). This is to offer the user a flexibility to choose a certain level of recommendation accuracy and diversity. In general, for any given ranking function  $rank_x(i)$ , this threshold  $T_R$  is used to create a parameterized version of that function  $rank_x(i, T_R)$ . The formal representation is illustrated in formula 2.

$$rank_x(i, T_R) = \left\{ \begin{array}{l} rank_x(i), \text{ if } R * (u, i) \in [T_R, T_{max}] \\ \alpha_u + rank_{s \text{ tan dard}}(i), \text{ if } R * (u, i) \in [T_H, T_R] \end{array} \right\} \tag{2}$$

where  $\alpha_u = \max rank_x(i)$ . Items that are predicted above  $T_R$  are ranked according to  $rank_x(i)$ , while items that are below  $T_R$  are ranked according to the standard ranking approach. All items that are above  $T_R$  are ranked ahead of all items that are below  $T_R$ .

### 2.8 Graph Theoretic Approach

Adomavicius and Kwon [7] introduced another approach to address the aggregate diversity in Recommender Systems. They named it a graph-theoretic approach. The recommendation step is carried out by using the standard ranking approach discussed in the previous section.

The approach formulates the problem of maximizing diversity as a well-known *max-flow problem* in graphs [8]. It translates users and items to vertices or nodes and an association of user and item to an edge. An edge from user to item exists if and only if item ( $i$ ) has been predicted to be relevant for user ( $u$ ). Each edge is assumed to have a capacity  $c(e) = 1$  and can be assigned an integer flow of 1 only if the item ( $i$ ) is actually recommended to user ( $u$ ) as part of top-N recommendations and 0 otherwise. The concept is illustrated in Fig. 1. The concept results in a situation where maximum flow value will be equal to the largest possible number of recommendations that can be made from among available items. In this case, no user can be recommended more than maximum capacity of an edge and no item can be counted more than once.

The concept precisely defines the *diversity-in-top-N* metric. Therefore, finding the maximum flow will be indeed finding the recommendations that yield maximum diversity. The graph theoretic approach algorithm yields better accuracy-diversity results as compared to item re-ranking approaches, however this improvement come at the cost of computational complexity.

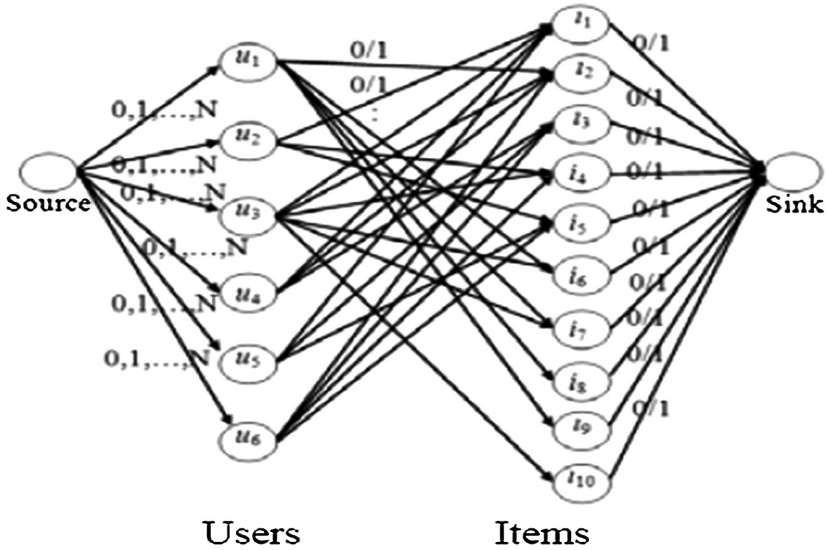


Fig. 1. Max flow problem [7]

### 2.9 Comprehensive Diversity

It was mentioned that diversity of recommendations can be measured in two ways: individual and aggregate. Individual diversity has been explored in a number of studies [9–13]. It simply focuses on providing dissimilar items to an individual user. The paper also indicated that an intra-list similarity is employed to monitor the individual diversity [9].

The aggregate diversity has also been discussed extensively with the techniques employed in the previous section. It has been apparent that more work still has to be done to put these recommendation aspects together; diversity without the significant loss of accuracy. Comprehensive diversity will bridge the gap between aggregate and individual diversity putting them together and treating them as unit entity. The proposed approach is discussed in the next section.

### 3 Proposed Approach: Collaborative Filtering with Clustering Re-Ranking Technique (CFCRT)

Collaborative Filtering with Clustering Re-Ranking Technique is proposed in this paper in order to enhance diversity in recommender systems. This method seeks to optimize the performance of recommender systems. It improves the recommendation quality by re-ranking the recommended list in such a way that long-tail items are included.

Basically, CFCRT uses Item-based collaborative filtering for the predicting of items’ ratings and then it shuffles the recommendations in the manner illustrated by the following steps:

1. Compute similarities between items (Pearson Correlation)
2. Predict ratings to unrated items
3. Use standard ranking to create a recommendation list with items above the pre-set threshold value (ensures accuracy)
4. Divide the list into equal clusters
5. Create the final list with the items selected randomly from each cluster (ensures diversity)
6. Send the final list to the active user.
7. This can be illustrated graphically in Fig. 2.

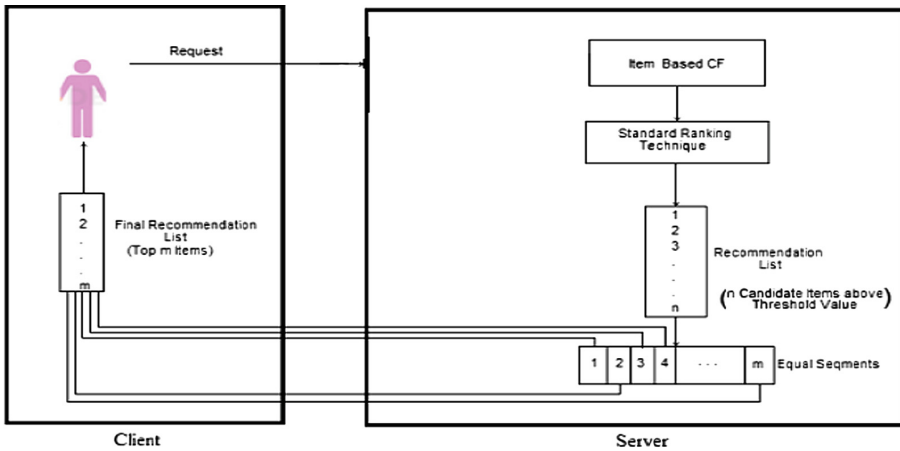


Fig. 2. Recommendation mechanism for user request

### 3.1 Similarity Computations

We employed Pearson Correlation similarity to compute similarities between items. From statistics background, Pearson Correlation of two series is taken to be the ratio of their covariance to the product of their variance. Formula 3 is used for this purpose.

$$s(i, j) = \frac{\sum_{u \in U} (R(u, i) - \overline{R(i)})(R(u, j) - \overline{R(j)})}{\sqrt{\sum_{u \in U} (R(u, i) - \overline{R(i)})^2} \sqrt{\sum_{u \in U} (R(u, j) - \overline{R(j)})^2}} \quad (3)$$

### 3.2 Predicting Ratings

To predict the rating that an active user would give to an item he/she has never rated, we used the weighted sum technique. The predicted rating of the item was computed by taking the quotient of the sum of the ratings given by the user on the items which were found to be similar to the item in question weighted by their corresponding similarity

values to the item and the sum of these similarity values. The similarity sum term ensures that the predictions are within the range. This is formalized in Formula 4.

$$R^*(u, i) = \frac{\sum_{n=1}^m (R(u, i_n) \times s(i, i_n))}{\sum_{n=1}^m s(i, i_n)} \tag{4}$$

### 3.3 Standard Ranking of Items

In standard ranking, the items predicted above the pre-set threshold value ( $T_H$ ) which was 3.5 in our case, were ranked in descending order, which means the highly predicted item comes first in the list and the lowest predicted at the bottom of the list. This is to ensure accuracy of the recommended items.

### 3.4 Dividing the List into Equal Clusters

The list created by using the standard rating of items was then divided into equal number of segments (clusters) based on the expected number of recommendations. For example, in our case, the number of recommended items was 5 and there were 20 items in the list, the list is divided into 5 clusters each carrying 4 items. This is illustrated in Fig. 3.

### 3.5 The Final Recommendation

The final recommendation list was processed by randomizing the items in different clusters and picking an item from each cluster to make the final list. This was to ensure and enhance diversity. From Fig. 3, if diversity was to be ignored, items 1, 2, 3, 4 and 5 would have been recommended to the user and as far as accuracy is concerned it would be correct. However, long-tail items like 9, 15 and 19 would have been left out yet the user might have been interested in them.

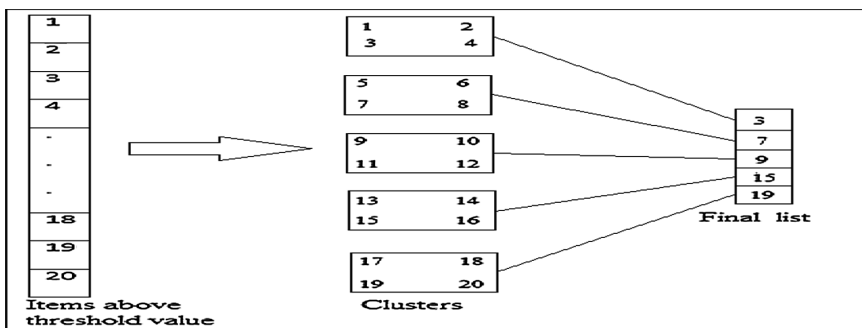


Fig. 3. Clustering and re-ranking of the list

### 3.6 Evaluating Accuracy

Accuracy was evaluated as a percentage of truly highly ranked ratings among those that were predicted to be  $N$  most relevant highly-ranked ratings (precision). The highly-ranked items were denoted as  $correct(L_N(u))$ .

$$precision - in - topN = \frac{\sum_{u \in U} |correct(L_N(u))|}{\sum_{u \in U} |L_N(u)|} \quad (5)$$

where  $L_N(u)$  is the recommendation list and  $correct(L_N(u)) = \{i \in L_N(u) | R(u, i) \geq T_H\}$ .

### 3.7 Evaluating Individual Diversity

Intra-list similarity metric was used to evaluate the individual list diversity. The average dissimilarity of all pairs of items was modelled as a distance function  $d : I \times I \rightarrow R$  such that  $d(i, j)$  represents dissimilarity between elements  $i, j \in I$ . The diversity was measured as:

$$diversity - in - topN = \frac{2}{p(p-1)} \sum \sum d(i, j) \quad (6)$$

where  $i$  and  $j$  are the items in the recommendation list and  $i \neq j$ . The assumption is that the distance function is symmetric, meaning  $d(i, j) = d(j, i)$  and  $p$  is the number of items in the recommendation list (Top-N). The function  $d(i, j)$  is derived from the similarity metric  $s(i, j)$  such that  $d(i, j) = 1 - s(i, j)$ .

### 3.8 Evaluating Aggregate Diversity

We measured the performance based on the lists of recommended items, therefore aggregate diversity was measured as the total number of distinct items recommended across all the users. The formal definition can be illustrated as:

$$Agg - div - in - topN = |\cup L_N(u)| \quad (7)$$

where  $L_N(u)$  is the recommendation list.

### 3.9 Evaluating Comprehensive Diversity

Comprehensive diversity is the observation that both individual and aggregate diversity are maximized with the minimum or no loss of precision.

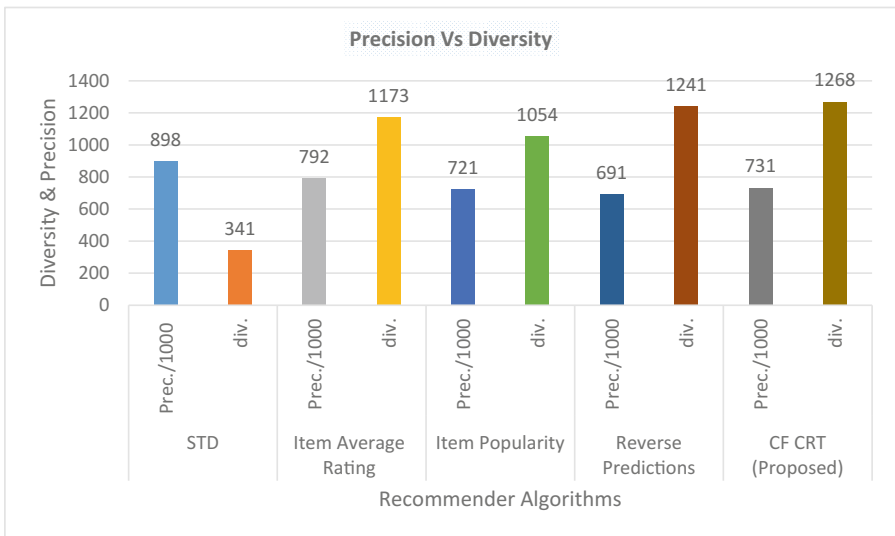
## 4 Experiments and Results

The experiments were conducted using one set of data. The dataset was MovieLens dataset (<http://grouplens.org/datasets/movielens/>). The dataset has 100,000 ratings on 1,700 movies by 1,000 users. The basic statistical information for the dataset is shown in Table 1.

**Table 1.** Statistical information about the dataset.

	Movie lens
Number of users	1,000
Number of movies	1,700
Number of ratings	100,000
Data sparsity	5.88%
Avg. # of common movies between two users	57
Avg. # of common users between two movies	30
Avg. # of users per movie	143
Avg. # of movies per user	243
Test data	40%
Training data	60%

The performance of the proposed algorithm was benchmarked against the existing algorithms such as Item Average Rating, Item popularity and Reverse Predictions [2]. The diversity and accuracy of the proposed algorithm was tested together with that of the other algorithms mentioned as well as the standard ranking algorithm.



**Fig. 4.** Precision and aggregate diversity at 3.5 threshold

The experiment was conducted with the threshold fixed at 3.5 and the results are shown in Fig. 4.

When the threshold was at 3.5 the precision and diversity of different algorithms are shown on Fig. 4. It can be noticed from Fig. 4 that the diversity enhancing algorithms including the proposed algorithm can indeed gain diversity on the recommendation list but with significant accuracy loss as compared to standard rating. The standard ranking algorithm reached .898 precision with only 341 different items while other algorithms had a lower precision but with more different items. The objective was to obtain maximum possible values for both accuracy and diversity. The proposed method had the highest diversity and had third best precision.

Varying threshold value towards the maximum rating value (which is 5 in this case), significantly increases accuracy but on the other hand diversity is much affected. Varying it towards the minimum value enhances diversity but accuracy is lost, hence the trade-off. So, we varied threshold value towards the maximum rating value (from 3.5 towards 5) while the other parameters were fixed to obtain different values of

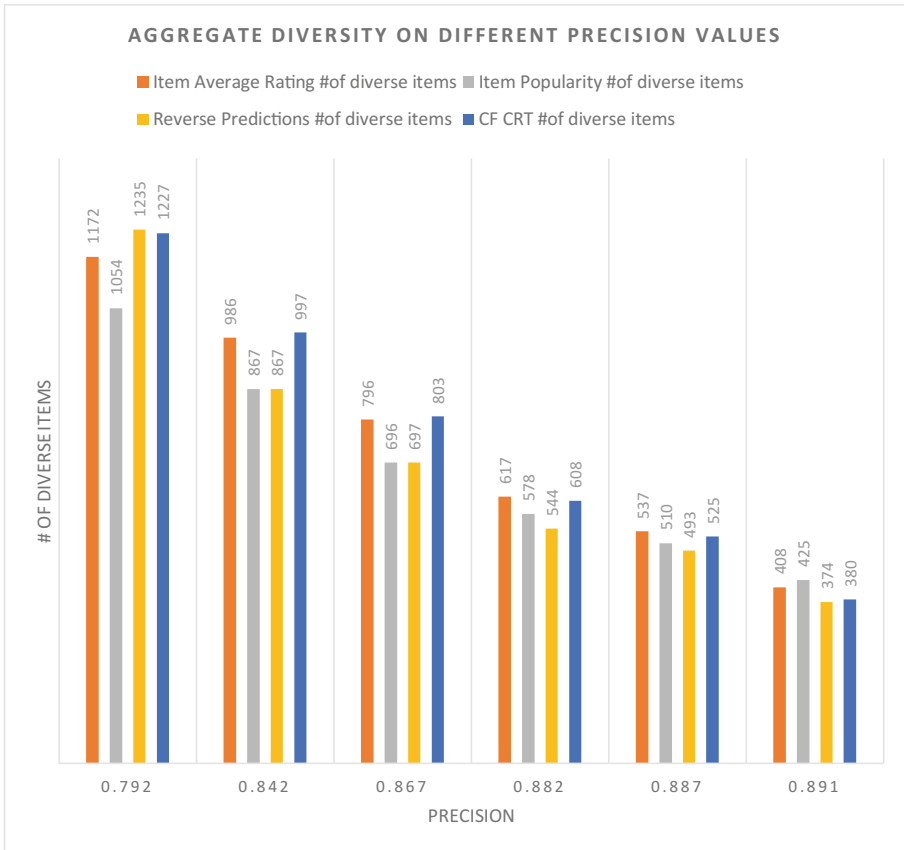


Fig. 5. Aggregate diversity on different precision values

accuracy in terms of precision from which different algorithms were compared. The number of items in the recommendation list was fixed at 5 and the number of neighbours at 20. The results are shown in Fig. 5. The threshold value for each and every algorithm was adjusted to get the common precision values from which diversity gain was measured.

The observation from Fig. 5 is that irrespective of the ranking technique used, an increase in precision is antagonistic to diversity. Varying the threshold value from 3.5 towards the maximum prediction value improved accuracy which was hurt by the diversity enhancing algorithms. If we look at the proposed algorithm, we can notice that with precision improvement from 0.792 to 0.842 we still have the significant diversity of 997. Most of the time the proposed algorithm performed better than the other algorithms.

The results obtained from individual user perspective (Individual Diversity) are illustrated in Table 2.

**Table 2.** Individual diversity on different precision values

Precision	Item average rating (intra-list dissimilarity)	Item popularity (intra-list dissimilarity)	Reverse predictions (intra-list dissimilarity)	CF CRT (intra-list dissimilarity)
0.792	0.798	0.81	0.884	0.891
0.842	0.774	0.732	0.831	0.857
0.867	0.729	0.69	0.796	0.812
0.882	0.701	0.624	0.752	0.79
0.887	0.689	0.582	0.721	0.714
0.891	0.654	0.55	0.671	0.601

From Table 2, it can be noticed again that the intra-list dissimilarity is minimized by increasing precision. If we take the proposed algorithm for example, at the precision of 0.792, the average dissimilarity was 0.891 which means that there was a high difference between the items recommended to a single user because the precision was low. Increasing the precision lowered the dissimilarity, meaning that the items' set was now becoming full of similar items. It can also be noticed from the table that the proposed algorithm performs better at the precision of 0.882 although it severely drops below the Item Average Rating and Reverse Predictions algorithms with the further increase in precision.

From the results, it is eminent that adjusting the threshold value has a significant impact in recommendation list. It was also observed that with the same precision loss in all algorithms, the diversity of the proposed algorithm is better as compared with other algorithms.

Adjusting the threshold value towards the maximum rating number results in loss of diversity because most of the times the highly rated items are the well-known items so those items tends to have high similarity values from the collaborative effect.



The proposed algorithm achieves a better balance because it first ensures accuracy by employing standard ranking and then mix evenly the items from all the sectors of the list which results in a better diversity gain.

From the results, it can be noticed that there is diversity enhancement coming with the proposed CF CRT algorithm. It can also be notice that there is minimal accuracy loss on average. The proposed algorithm still gets a challenge from the existing algorithms which we used for comparison but in general it looks the best in most cases.

## 5 Conclusion and Future Work

The trade-off issue between diversity and accuracy can be balanced by varying the threshold value. For every algorithm, it was noted that varying the value towards the maximum rating value significantly increases accuracy. While accuracy is an important metric, the problem that surfaced was that only well-known movies were recommended to users and most probably they were the movies which almost every-body who is into movies has already watched or knows about.

Varying the threshold value towards the minimum rating value increased diversity by recommending even the long tail movies which did not have enough information. That is again risky because it recommended movies which users have shown no interest in them. So it was only the balancing of the extreme ends that proved to provide the optimum results with the mixture of well-known and lesser known movies.


The future work that still warrants attention is to try and get the system with 100% accuracy and 100% user satisfaction all the time, until then the field of Recommender Systems is still open and interesting.

## References

1. Niemann, K., Wolpers, M.: A new collaborative filtering approach for increasing the aggregate diversity of recommender systems. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, Illinois, USA (2013)
2. Adomavicius, G., Kwon, Y.O.: Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Trans. Knowl. Data Eng.* **24**, 896–911 (2012)
3. Brynjolfsson, E., Hu, Y., Simester, D.: Goodbye pareto principle, hello long tail: the effect of search costs on the concentration of product sales. *Manage. Sci.* **57**, 1373–1386 (2011)
4. Ricci, F., Rokach, L., Shapira, B., Kantor, P.B.: *Recommender Systems Handbook*. Springer, Heidelberg (2011)
5. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th International Conference on World Wide Web, Hong Kong, pp. 285–295 (2001)
6. Burke, R.: Hybrid web recommender systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) *The Adaptive Web*. LNCS, vol. 4321, pp. 377–408. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-72079-9\\_12](https://doi.org/10.1007/978-3-540-72079-9_12)

7. Adomavicius, G., Kwon, Y.O.: Maximizing aggregate recommendation diversity: a graph-theoretic approach. In: Workshop on Novelty and Diversity in Recommender Systems, Held in Conjunction with ACM RecSys, 23 October 2011
8. Dong, J., Wei, L., Congbo, C., Zhong, C.: Draining algorithm for the maximum flow problem. In: WRI International Conference on Communications and Mobile Computing, CMC 2009, pp. 197–200 (2009)
9. Bradley, K., Smyth, B.: Improving recommendation diversity. In: Proceedings of the 12th Irish Conference on Artificial Intelligence and Cognitive Science (2001)
10. Smyth, B., McClave, P.: Similarity vs. diversity. In: Aha, D.W., Watson, I. (eds.) ICCBR 2001. LNCS (LNAI), vol. 2080, pp. 347–361. Springer, Heidelberg (2001). [https://doi.org/10.1007/3-540-44593-5\\_25](https://doi.org/10.1007/3-540-44593-5_25)
11. McSherry, D.: Diversity-conscious retrieval. In: Craw, S., Preece, A. (eds.) ECCBR 2002. LNCS (LNAI), vol. 2416, pp. 219–233. Springer, Heidelberg (2002). [https://doi.org/10.1007/3-540-46119-1\\_17](https://doi.org/10.1007/3-540-46119-1_17)
12. Zhang, M., Hurley, N.: Avoiding monotony: improving the diversity of recommendation lists. In: Proceedings of the 2008 ACM Conference on Recommender Systems, pp. 123–130 (2008)
13. Ziegler, C.N., McNee, S.M., Konstan, J.A., Lausen, G.: Improving recommendation lists through topic diversification. In: Proceedings of the 14th International Conference on World Wide Web, pp. 22–32 (2005)

# Towards Intelligent System Wide Information Management for Air Traffic Management

Li Weigang<sup>1</sup> , Alessandro F. Leite<sup>1</sup>, Vitor F. Ribeiro<sup>1</sup>,  
Jose A. Fregnani<sup>2</sup>, and Italo R. de Oliveira<sup>2</sup>

<sup>1</sup> TransLab, Department of Computer Science, University of Brasilia,  
Brasilia-DF, Brazil  
weigang@unb.br

<sup>2</sup> Boeing Research and Technology/Brazil, Av. Dr. Altino Bondesan,  
SJCamos-SP, Brazil

**Abstract.** This paper briefly reviews the state-of-the-art in Artificial Intelligence (AI) applied to Air Traffic Management (ATM). The research topics include the application of semantic ontology, multi-agent systems, reinforcement learning (RL), and game theory in ATM. Likewise, this paper also highlights our research advances in this area. In this case, we describe a new Probabilistic Web Ontology Language (PR-OWL) algorithm to enable the reasoning on big datasets in polynomial time. Then, we present the use of both Particle Swarm Optimization (PSO) and Simulated Annealing (SA) algorithms in 4D trajectory management. Next, we describe the usage of Multi-agent Planning (MAP) theory on airport ground handling management. Finally, this paper envisions some research and development directions of AI applied to ATM. It includes: (a) mapping and reducing the gaps between advanced AI technologies and ATM; (b) considering uncertainty in Semantic Ontology for SWIM data exchanging models in ATM; (c) using big data analytics in SWIM; and (d) integrating collaborative ATM technologies towards intelligent SWIM (I-SWIM).

**Keywords:** Air Traffic Management · Artificial Intelligence  
Semantic Ontology · Swarm Optimization  
System Wide Information Management

## 1 Introduction

As the successful development of the System Wide Information Management (SWIM) in Air Traffic Management (ATM) all over the world [1], challenges have emerged for every related partner, especially for the developing countries to harmonize and secure ATM modernization. A good example is the collaboration progress toward achieving the necessary level of interoperability between the Next Generation Air Transportation System (NextGen) of the United States and the Single European Sky ATM Research (SESAR) Program of the European Union [2].

SWIM is a new technological framework to address the communications and interoperability requirements of highly distributed, loosely coupled and platform-independent components by consistently applying the principals of Service-Oriented

Architecture (SOA) for ATM [3]. By Federal Aviation Administration (FAA), SWIM is the Information Technology (IT) infrastructure to offer the information access to transform the aviation community [4]. At least 117 subsystems to form Aeronautical, Flight (Planning and Tactical) and Weather Information resources were mapped in the implementation of SWIM. Its complex, dynamic and intrinsically heterogeneous nature involves multi-disciplines solutions from various fields such as Mathematics, Computer Science, Statistics, Information Theory, among others.

In October 2016, the National Science and Technology Council (NSTC) of United States released “*The National Artificial Intelligence Research and Development Strategic Plan*” [5]. As cited in this plan, AI can improve the efficiency of transportation to materially impact safety for all types of travel. These facts show that the application of AI in SWIM is a new challenge in the modernization of Air Transportation. The ICAO Report on its 39<sup>th</sup> assembly session outlines the requirements for new approaches to improve all aviation perspectives, through an environment that enables interoperability between different stakeholders for an efficient decision-making process [6].

In this context, this paper briefly reviews the state-of-the-art in Artificial Intelligence (AI) applied to ATM. From literature, the applications include semantic ontology for aviation data, multi-agent systems for ATM, reinforcement learning (RL) to improve the performance of controllers, and game theory in Collaborative Decision Making (CDM). Likewise, this paper also describes recent research developments on these topics. First, we describe a new Probabilistic Web Ontology Language (PR-OWL) approach to enable reasoning on big datasets in polynomial time. Second, we present the use of Particle Swarm Optimization (PSO) and Simulated Annealing (SA) algorithms in 4D trajectory management. Third, an application of Multi-Agent Planning (MAP) applied to airport management operation is described. Based on the understanding of the National AI R&D strategy plan [5] and on our experience, this paper envisions several research and development (R&D) trends in AI for ATM, especially for the System Wide Information Management (SWIM).

With this concerning, the main contribution of this paper is to highlight AI R&D trends for ATM, including the following directions: (a) mapping and reducing the gaps between advanced AI technologies and ATM; (b) considering uncertainty in Semantic Ontology for SWIM data exchanging models in ATM; (c) using big data analytics in SWIM; and (d) integrating collaborative ATM technologies with SWIM. The attention of these challenges by ATM community will make a significant step for implementing Intelligent System Wide Information Management (I-SWIM).

## 2 AI R&D State of the Art in ATM

### 2.1 Semantic Aviation Information Management

Currently, SWIM focuses on information exchange based on a service-oriented architecture. In this case, SWIM offers a set of infrastructure services, including messaging capabilities, enterprise service management, and structured data format, as depicted in Fig. 1 [1, 7, 8]. As the air traffic volume has increased significantly over the

world, the great mass of traffic management data, named as Big Data, have also accumulated day by day. This factor presents more opportunities and also challenges as well in the study and development of ATM.

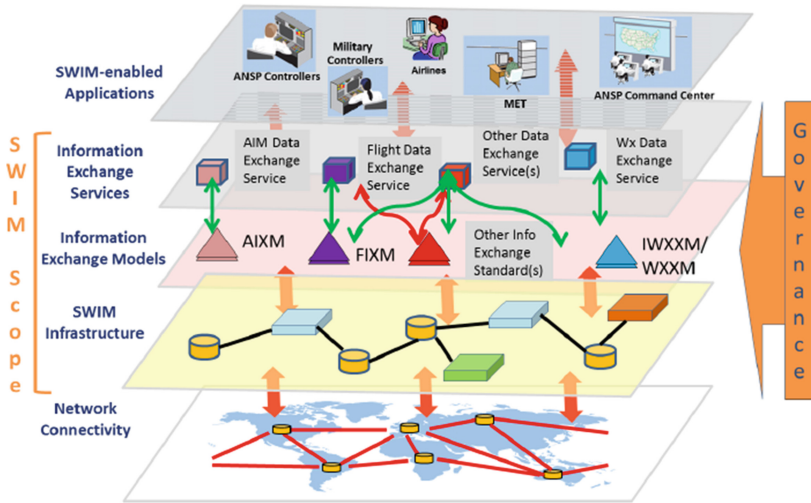


Fig. 1. SWIM global interoperability framework [1].

Using Unified Modelling Language (UML) in the solution of SWIM, the management of information is to establish the standards that define the information content, format and rules for data exchange [1]. The information exchange standards are applicable to aeronautical information by Aeronautical Information Exchange Model (AIXM), flight information by Flight Information Exchange Model (FIXM), and meteorological information by Weather Information Exchange Model (WXXM), ICAO Meteorological Information by Exchange Model (IWXXM), and aviation information by Aviation Information Data Exchange (AIDX), see Fig. 1.

Intelligent System Division of NASA Ames Research Center has been leading some research projects to implement a system of semantic ontology, in prototype level, for combining heterogeneous air traffic management data using semantic integration techniques [9]. The system transforms data from its original disparate source formats into a unified semantic representation within an ontology-based store. More important observations of the research revealed the potential of the application of ontology approaches to keep the data model and the data representation synchronized with AIXM, FIXM, WXXM, IWXXM and AIDX [9] to improve the SWIM performance.

## 2.2 Multi-agent System in ATFM

To manage the air traffic flow efficiently, the methods of multi-agent systems from AI have been successfully introduced to ATFM. For instance, Nguyen-Duc and colleagues [10] have proposed an approach to deal with cooperation and negotiation among agents

using grid computing in a real time traffic synchronization problem. Collaborative ATFM using multi-agent simulation has also been developed by NASA, where some route selection strategies for airline operations were implemented [11]. Tumer and Agogino [12] also developed a distributed agent-based Air Traffic Flow Management, and their research presented an interesting and contemporary approach based on multi-agent and reinforced learning applied to ATM considering the human-in-the-loop. Meanwhile, a model for a multi-agent system for ATFM in grid computing was developed by Dib and colleagues [13].

### 2.3 Learning to Improve the Performance of Controllers

To Agogino and Tumer [14] developed metrics to evaluate the efficiency of traffic flow control measures generated by reinforcement learning agents. As a result, it was possible to improve the actions taken by an agent conceived to simulate the behavior of air traffic controllers, given the suggestions by the ATFM agent and as such suggestions were accepted by the air traffic controller.

It should be noted that, in the system proposed by Agogino and Tumer [14], the ATFM agent is a reinforcement learning algorithm that seeks to maximize the reward that derives from the generated scenarios. The scenarios will be evaluated based on the total resulting delays and the number of aircraft in the ATC sectors – both resulting from the flow control policy.

Crespo and colleagues [15] used reinforcement learning to implement a software agent system to support the decision-making process in ATFM. Their system takes into the experiences of the decision makers. Hence, a device capable of generating flow adjustment policies were developed. The policies comprise a set of traffic flow control measures that are suggested to the traffic flow managers.

### 2.4 Game Theory for Collaborative Decision Making Environment

Air Traffic Flow Management (ATFM) is a critical activity due to the strict requirements by safety and efficiency factors, involving many partners in the system such as passengers, airports, airlines, air traffic control authorities, among others. Collaborative Decision Making (CDM) with Ground Delay Program (GDP) is one of the more important management processes in ATFM. Schummer and Rakesh [16] develop a game-theoretical model for airport landing slots reassignment. Their model is also a generalization of the top trading cycle (TTC) algorithm, and it aims to enforce airlines' ownership of assigned landing slots during a ground delay program, in a way that the airlines are free to use their slots as they wish or to trade them to meet their objectives.

To address the solution of airport slot allocation in CDM, especially in slot compression step, the matching method with Deferred Acceptance (DA) algorithm of Game Theory is applied in various previously research. Considering airport management as an effective stakeholder in CDM, Arruda and colleagues [17] propose a new kind of two-sided matching market, in which one side is with the flights belonging to several airlines, and another side is with the runway utility slots belonging to airport management. The solution for this special market, named as DA-SLOT, obtains the stability and optimization of Compression results comparing to the classic CDM.

Ribeiro and colleagues [18] also developed a game-theoretical approach for air traffic management. In their research, they implemented a collaborative departure management (CoDMAN) framework. The theoretic foundation relies on an adapted Rubinstein protocol for departure sequencing at an airport. Several demand scenarios were simulated, and the results show that the intelligent, cooperative behavior toward a common goal can take advantage of individual interests without compromising the safe operations during the departure management.

### 3 Recent AI Research and Application in ATM

#### 3.1 A New PR-OWL Algorithm for Big Datasets

Semantic Web (SW) adds semantic information to the traditional Web, allowing computers to understand the contents of that Web page, before it is accessible only by human beings. The Web Ontology Language (OWL) is the main language for building ontologies in SW, and it allows a formal modeling of a knowledge domain based on description logics. OWL, however, does not support the process of the uncertainty information [19]. This restriction motivated the creation of several extensions of this language, such as Probabilistic OWL (PR-OWL) which improves OWL with the ability to treat uncertainty using Multi-Entity Bayesian Networks Using the first-order probabilistic logic, the inference of MEBN consists of generating a Situation Specific Bayesian Network (SSBN) [20]. PR-OWL 2 extends the PR-OWL offering a better integration with OWL and its underlying logic, allowing the creation of ontologies with deterministic and probabilistic parts.

Santos and colleagues [19] proposed PR-OWL 2 RL, a scalable version of PR-OWL based on OWL 2 RL profile and on triple stores. OWL 2 RL allows reasoning in polynomial time for the main reasoning tasks. Triple stores can store RDF (Resource Description Framework) triples in databases optimized to work with graphs. A new algorithm was developed to allow the generation of SSBNs for databases with large evidence base. It is scalable because it instantiates an evidence node only if it influences a target node. A case study over frauds into procurements showed the efficiency of the developed approach to deal with big datasets.

#### 3.2 Local Search Approaches in CDA in ATM

The operation of Continuous Descent Arrivals (CDA) is an important part of 4D trajectory management in ATM. The main goal in CDA operations is to manage the flight trajectories in order to optimize the operational capability of arriving aircraft in a Terminal Maneuvering Area (TMA) by reducing the fuel burn and emissions during the approach/descent phase. The traffic coordination upon arrival can be understood as a scenario in which several aircraft are scheduled to cross a confluence point within optimum time windows under the coordination of the air traffic controller (ATC). Safety constraints for the airspace are also a matter of concern, thus the ATC agent must eliminate en-route conflicts and provide an adequate arrival sequence that is fair and conflict-free [21].

Ribeiro and colleagues [21] propose the usage of local search algorithms applied to arrival coordination under the 4D paradigm. The selected algorithms are Particle Swarm Optimization (PSO) and Simulated Annealing (SA). These are heuristic algorithms that provide fast conversions towards global minima, frequently used on non-linear constrained single objective problems, appropriate for CDA sequencing modeling and dynamic simulations.

### 3.3 Multi-agent Planning in Airport Operation Management

Airport Collaborative Decision Making (A-CDM) is a wide accepted concept which creates a common ground management procedure for the different components of the Air Transportation System (ATS). This concept is based on an improved communication between the different stakeholders of the airport (e.g., Air Traffic Control, Airport Authority, and Airlines). Nevertheless, within the turnaround process of aircraft at airports, Ground Handling Management (GHM) of aircraft has not been well developed specifically in the A-CDM procedure, even if it takes an important role in the fluidity of the aircraft ground movements at airports [22].

GHM comprises various services required by airplanes while they are on the ground, parked at a terminal gate or at a remote position at an airport. This includes the processing of boarding/de-boarding passengers, baggage and freight, as well as the maintenance of the aircraft itself (e.g., fueling, cleaning, and sanitation, among others) [22].

Using Multi-Agent Planning (MAP) method [23], Kabongo and colleagues [24] developed a computation framework as well as a management system to improve airport ground handling management (GHM). With the identification of the services and resources related to GHM, the forward MAP approach is applied to coordinate the tasks and plans in order to reduce both the delays and the operating cost. In this case, the key contribution includes MAP model for airport ground handling operations under a unified framework compatible with the airport collaborative decision making (A-CDM) strategy.

## 4 AI R&D Trends in SWIM for ATM

### 4.1 Reducing the Gaps Between Advanced AI and ATM

During the last decades, AI has matured from the advances in many fields. The factor of the quick development of intelligent information technology left some significant gaps between AI and other fields such as air transportation. Following research topics are listed to show the potential AI technologies in ATM in order to reduce these gaps based on the National AI R&D strategy plan [5, 25].

**Machine Learning for ATM.** Over the past decade, the machine learning as an important subfield of AI enables computers to learn from experience or examples and has demonstrated increasingly accurate results [5]. In ATM study, especially in the study of the new process of 4D trajectory management and SWIM, the theory and application of machine learning are not formally established. There is no specification of AI related to SWIM in official documents of Eurocontrol, FAA and ICAO [1–4].



Overall speaking, machine learning systems are in their infancy in ATM, but they are appealing potential applications [14, 15, 21].

**Autonomous learning from Information Exchanging Models in SWIM.** The ATM systems are exposed to real time with uncertainties, i.e. many facts were not predicted at design time and can lead the system to failure. It is necessary to get the information at right time, in case of SWIM. The challenge is to develop a mechanism using autonomous learning from Information Exchanging Models in SWIM at early stages in order to reduce runtime uncertainty in decision-making process of self-adaptive systems. In fact, autonomous learning has already been applied to deep learning, but mostly to show the advantage of unlabeled examples and is far from enough to achieve satisfactory performance [25].

**Effective methods for human-AI collaboration.** Reinforcement learning algorithms have been applied to seek the maximization of using reward [14] and to improve the controller's performance [15] in ATM process. There is much space for improvement to develop effective methods for human-AI collaboration in automation pilot and assistance controller. Rather than replace humans, most AI systems will collaborate with humans to achieve optimal performance [5] in ATM.

**Intelligente perception methods for ATM.** By introducing new surveillance technologies such as Automatic dependent surveillance — broadcast (ADS-B) and Automatic dependent surveillance — contract (ADS-C), the methods and procedures of air traffic control and management will be significantly modified. Advanced intelligent perception can be successfully applied to generate a uniform semantic representation of objects, scenes, behavior and events in ATM scenario. At the same time, the solution for complex ATM problems can breakthrough intelligent perception methods and technologies by the development of new machine learning algorithms and methods for large-scale perception data [25].

## 4.2 Considering Uncertainty by Semantic Ontology in SWIM with Big Data

In the development of SWIM, there are some challenges to take into consideration. The solution will also direct the trends of AI R&D in ATM. It introduces Semantic Ontology approaches to SWIM. For better understanding of context, Table 1 shows the evolution of OWL related languages for treatment of the uncertainty on big datasets.

**Table 1.** Recent OWL forms for uncertainty and big datasets

Semantic form	Profiles		
	Logic form	Probability support	Scale support
PR-OWL	OWL	MEBN	–
PR-OWL 2	OWL 2	MEBN	Bottom-up
PR-OWL 2 RL	OWL 2 RL	MEBN	Bayes-ball

**Interface to synchronize the information from SWIM.** It is necessary to construct an interface to synchronize the useful information from the exchange models of AIXM, FIXM, WXXM, IWXXM and AIDX. As mentioned by Keller et al. [9], the application of OWL is a considerable solution. With the implementation of OWL in SWIM by triple store, the information from different communities can be better harmonized to supply to the SWIM users.

**Uncertainty treatment of the information from SWIM.** Even OWL has a potential to be applied in the harmonization of the information from different exchange models in SWIM. In case of ATM scenario, uncertainty is a very common property of the flight schedule in practice, subject to airline operational difficulties, weather and other factors. PR-OWL was developed to deal with the uncertainty information by MEBN [20].

**Big dataset processing approaches for SWIM.** For large scale database, especially in ATM case, PR-OWL 2 RL has been developed [19]. This OWL approach was modified by adapting Bayes-Ball algorithm to process large scale Bayesian Networks considering uncertainty. The next step is to construct the triple store by the translation of the information from the exchange models of AIXM, FIXM, WXXM, IWXXM and AIDX. And then to apply PR-OWL 2 RL in SWIM environment, this maybe the highlighted research as a new AI R&D trend in ATM.

### 4.3 Integrating Collaborative ATM Technologies with SWIM

One success progress in ATM is the development of Collaborative Air Traffic Management Technologies (CATMT). The main technologies of CATMT include Collaborative Decision Making (CDM) [16–18, 26, 27], Collaborative Trajectory Options Program (CTOP) [28, 29] and others. These technologies are well established in ATM applications. As mentioned in the NextGen program by FAA, it is necessary to construct a mechanism for Collaborative Information Exchange (CIX) in order to increase situational awareness and improved constraint prediction by the incorporation of data made available via SWIM. The core problem is in the following two aspects that have not been clearly specified by community yet:

**The right information from SWIM to support CATMT.** To realize processes of CDM or CTOP, vast information should be available for related stakeholders in ATM. With the implementation of SWIM, there will be sufficient information available to support CATMT. The challenge is how to integrate the systems of CDM or CTOP with SWIM effectively to get the right information for decision making.

**The result information from CATMT back to SWIM.** The decision results by CATMT are very rich and important for ATM stakeholders. Another challenge for SWIM is to feedback these results for more related users of SWIM.

The perspective of this integration may also include the implementation of a semantic ontology based framework with an interface to enable the communication between CATMT and SWIM and improve the operational performance of the air transportation.

## 5 Conclusion

The Artificial Intelligence as an emerging technology for air transportation has demonstrated the great potential to address some important applications in air traffic management. This was demonstrated by the US National AI R&D strategy plan and the state of the art of application AI in ATM. From the literature review, it was possible to identify the great challenges that the ATM community faces and that can be solved with the help of advanced AI techniques, in the various aspects of its daily activities.

The AI R&D directions in ATM listed in this article should be seriously considered for further development, as means to obtain substantial improvements in ATM, especially with regards to SWIM. The gaps between advanced AI technologies and ATM applications here presented are based on our particular perspective, so we know that there are more benefits and challenges related to the influence of AI in Air Transportation. The proposed strategy of applying Semantic Ontology considering uncertainty for large-scale ATM data is just one way of exploiting the benefits of information exchange models available from SWIM. The next steps in our research will be devoted to the consolidation of AI tools in SWIM (I-SWIM) to render the advantages of the evolution that were envisioned in this paper.

## References

1. International Civil Aviation Organization (ICAO): Doc 10039 - Manual on System Wide Information Management (SWIM) Concept. Technical report of ICAO, Montreal (2015)
2. NextGen and SESAR: State of Harmonisation Document. MG-04-15-043-EN-N (2015). <https://doi.org/10.2829/572729>
3. Fernandez-Sancho, P., Kaplun, M., Roelants, E., Uri, C.: SWIM common registry: concept, architecture, and implementation. In: 4th ATIEC, Maryland, USA (2015)
4. Federal Aviation Administration (FAA): System Wide Information Management (SWIM) Product Portfolio (PDF) (2016). <https://www.faa.gov/nextgen/programs/swim/>
5. The Subcommittee on Networking and Information Technology Research and Development (NITRD): The national artificial intelligence research and development strategic plan. Technical report (2016). [https://www.nitrd.gov/PUBS/national\\_ai\\_rd\\_strategic\\_plan.pdf](https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf)
6. International Civil Aviation Organization (ICAO): Aviation security – policy. Technical report of ICAO, Montreal (2016)
7. Porosnicu, E.: Towards a global digital NOTAM specification. In: Proceedings of 5th Annual Air Transportation Information Exchange Conference (ATIEC), Maryland, USA, (2016)
8. Matthews, M., Pressler, C.: The SWIM PMO: utilizing data today for better situational awareness tomorrow. In: 5th ATIEC, Maryland, USA (2016)
9. Keller, R., Ranjan, S., Wei, M.Y., Eshow, M.M.: Semantic representation and scale-up of integrated air traffic management data. In: International Workshop on Semantic Big Data. ACM (2016)
10. Nguyen-Duc, M., Briot, J.P., Drogoul, A., Duong, V.: An application of multi-agent coordination techniques in air traffic management. In: Proceedings of the IEEE/WIC International Conference on Intelligent Agent Technology, Canada, pp. 622–628 (2003)

11. Wolfe, S.R., Jarvis, P.A., Enomoto, F.Y., Sierhuis, M.: Comparing route selection strategies in collaborative traffic flow management. In: Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology, pp. 59–62. IEEE Press (2007)
12. Tumer, K., Agogino, A.: Distributed agent-based air traffic flow management. In: Proceedings of the 6th AAMAS, Honolulu, USA, pp. 330–337 (2007)
13. Dib, M.V., Weigang, L., Melo, A.C.: Approach of balancing of the negotiation among agents in traffic synchronization. *IEEE Lat. Am. Trans.* **5**(5), 338–345 (2007)
14. Agogino, A., Tumer, K.: Learning indirect actions in complex domains: action suggestions for air traffic control. *Adv. Complex Syst.* **12**(4–5), 493–512 (2009)
15. Crespo, A.M.F., Weigang, L., de Barros, A.: Reinforcement learning agents to tactical air traffic flow management. *Int. J. Aviat. Manag.* **1**(3), 145–161 (2012)
16. Schummer, J., Rakesh, R.: Assignment of arrival slots. *Am. Econ. J.: Microeconomics* **5**(2), 164–185 (2013)
17. Arruda, A.C., Weigang, L., Milea, V.: A new airport collaborative decision making algorithm based on deferred acceptance in a two-sided market. *Expert Syst. Appl.* **42**(7), 3539–3550 (2015)
18. Ribeiro, V.F., Weigang, L., Milea, V., Yamashita, Y., Uden, L.: Collaborative decision making in departure sequencing with an adapted rubinstein protocol. *IEEE Trans. Syst. Man Cybern.: Syst.* **46**(2), 248–259 (2016)
19. Santos, L.L., Carvalho, R.N., Ladeira, M., Weigang, L.: A new algorithm for generating situation-specific bayesian networks using bayes-ball method. In: Proceedings of the 12th International Workshop on Uncertainty Reasoning for the Semantic Web, Japan, pp. 36–48 (2016)
20. Laskey, K.B.: MEBN: a language for first-order Bayesian knowledge bases. *Artif. Intell.* **172**(2–3), 140–178 (2008)
21. Ribeiro, V.F., Pamplona, D.A., Fregnani, J.A., Oliveira, I.R., Weigang, L.: Modeling the swarm optimization to build effective continuous descent arrival sequences. In: 19th IEEE ITSC, Rio de Janeiro, Brazil, pp. 760–765 (2016)
22. Fitouri-Trabelsi, S., Mora-Camino, F., Nunes-Cosenza, C.A., Weigang, L.: Integrated decision making for ground handling management. *Global J. Sci. Front. Res.: Math. Decis. Sci.* **15**(1), 17–31 (2015)
23. De Weerd, M., Ter Mors, A., Witteveen, C.: Multi-agent planning: an introduction to planning and coordination. In: Handouts of the European Agent Summer (2005)
24. Kabongo, P.C., Ramos, T.M.F., Leite, A.F., Ralha, C.G., Weigang, L.: A multi-agent planning model for airport ground handling management. In: 19th IEEE ITSC, pp. 2354–2359, Rio de Janeiro, Brazil (2016)
25. Tian, Y., et al.: Towards human-like and transhuman perception in AI 2.0: a review. *Front. Inf. Technol Electron. Eng.* **18**(1), 58–67 (2017)
26. Ball, M.O., Chen, C.Y., Hoffman, R., Vossen, T.: Collaborative decision making in air traffic management: current and future research directions. In: Bianco, L., Dell’Olmo, P., Odoni, A. R. (eds.) *New Concepts and Methods in Air Traffic Management*. Transportation Analysis. Springer, Berlin, Heidelberg (2001). [https://doi.org/10.1007/978-3-662-04632-6\\_2](https://doi.org/10.1007/978-3-662-04632-6_2)
27. Bertsimas, D., Gupta, S.: Fairness and collaboration in network air traffic flow management: an optimization approach. *Transp. Sci.* **50**(1), 57–76 (2015)
28. Bosung, K., Clarke, J.-P.: Optimal airline actions during collaborative trajectory options programs. In: Proceedings of 54th AGIFORS, Dubai, UAE (2014)
29. Cruciol, L., Clarke, J.-P., Weigang, L.: Trajectory option set planning optimization under uncertainty in CTOP. In: Proceedings of 18th IEEE ITSC, Spain, pp. 2084–2089 (2015)

# A Security Risk Management Model for Cloud Computing Systems: Infrastructure as a Service

Mouna Jouini<sup>(✉)</sup> and Latifa Ben Arfa Rabai

Laboratoire SMART, Institut Supérieur de Gestion,  
Université de Tunis, Tunis, Tunisie

jouini.mouna@yahoo.fr, latifa.rabai@gmail.com

**Abstract.** Cloud Computing represents a new computing way that increases dynamically capabilities without investing new infrastructure. It become much adopted today thanks to many advantages like distributed computing, scalability and performance, multi-tenancy and pay per use services. However, it poses many serious security issues at all cloud delivery models. Software, Platform, and Infrastructure as a Service are the three main service delivery models for Cloud Computing. Infrastructure as a Service (IaaS) serves as the basis layer for the other delivery models, and a lack of security in this layer will affect the other delivery models. This paper presents a detailed study of IaaS components' security and determines vulnerabilities and security solutions. Finally, to combat security repose, we present a security risk management framework for Cloud system to threats and vulnerabilities reduction security risks mitigation. The proposed security risk management framework is based on a quantitative security risk assessment model to evaluate risks for this system.

**Keywords:** Cloud computing · Risk management framework · Virtualization Infrastructure as a service layer · Metrics

## 1 Introduction

Individual or enterprise users expect information systems to be secured and able to predict their risk and their strategies in reducing these risks. Security risks arise mainly from the loss of confidentiality, integrity, or availability of information or information systems and reflect the potential adverse impacts to organizational operations (like missions, functions, reputation), organizational assets and individuals.

Although the benefits of Cloud Computing (CC) are clear, so is the need to develop proper security for cloud implementations. In fact, the major concerns of adopting this new technology by organizations and individuals are security [10, 11, 20–22, 25–29]. This added level of security risk is because CC essential services and data are often outsourced to a third party which makes it harder to maintain data integrity, availability and privacy, support data and service availability, and demonstrate compliance [11, 25, 27]. Risk management framework is one of security assessment tool like metrics to reduction of threats and vulnerabilities and mitigates security risks.

The drive of secure organizational information has initiated the need to develop better metrics for understanding the state of the organization's security attitude [5, 6]. Wang states [7] "It is widely recognized that metrics are important to information security because metrics can be an effective tool for information security professionals to measure the security strength and levels of their systems, products, processes, and readiness to address security issues they are facing". On the other hand, risk assessment is one of the fundamental components of an organizational risk management process [10]. It is based on security metrics to assess security risks.

The National Institute of Standards and Technology (NIST) defines risk management as "the process of identifying risk, assessing risk, and taking steps to reduce risk to an acceptable level" [4]. The NIST defines or describes risk assessment as the process of identifying, estimating, and prioritizing information security risks which requires a careful analysis of threat and vulnerability information to determine the extent to which circumstances or events could adversely impact an organization and the likelihood that such circumstances or events will occur [4].

Risk is presented as a probability of an event and its impact or a consequence of the event when a threat was materialized. There are two types of security assessment approaches that estimate risks: quantitative and qualitative methods. Quantitative methods are based on security metrics or measures.

Security metrics are an effective tool to measure and assess the security levels of their systems, products, processes, and readiness to address the security issues they are facing. It help to identify system vulnerabilities and provide guidance in prioritizing corrective actions. Moreover, metrics can be used to justify and direct future security investment [7].

On the other hand, Boehm et al. argue that all dilemmas that arise in software engineering are of an economic rather than of a technical nature, and that all decisions ought to be modeled in economic terms: maximizing benefit; minimizing cost and risk [1–3]. They propose a new research axis, named Value Based Software Engineering, which aims to create "Good" software products and services in an economic standpoint. Quantitative security metrics represent a means of quantifying the risks in monetary terms in such a way as to enable rational decision making. Therefore, we propose in this article a security risk management model based on quantitative security model for Cloud Computing environment.

The aim of this paper is to present an information risk management approach for better understanding security issues in Cloud Computing environment and to identifying security threats and vulnerabilities. This approach is covering all of cloud infrastructure models. Cloud provider can be applied this framework to organizations to do risk mitigation using quantitative security risk assessment model.

The remainder of the paper is organized as follows. In Sect. 2, we review basic existing security risk assessment models for information system. We present these descriptions to separate methods in order to get better appreciation of the review results. Section 3 review security challenges and components for the Infrastructure as a Service layer (IaaS) in CC systems. Then, in Sect. 4 we describe the review methodology for Cloud Computing security risk management. Then, we provide an illustration of the application of this proposed framework in order to reduce security threats risk in the Infrastructure as a Service layer for CC environment. We draw some concluding remarks in Sect. 5.

## 2 Related Work

A range of general risk assessment methodologies and standards are used in literature [8–10, 14, 25, 28]. In this section, we discuss the most used methodologies and guidelines attending to security risk management for information systems.

### 2.1 ISO Security Risk Management Guideline

This International Standard ISO provides guidelines for information security risk management in an organization. It does not provide any specific method for information security risk management. It is up to the organization to define their approach to risk management, depending for example on the scope of the information security management system (ISMS), context of risk management, or industry sector. It is based on several activities mainly [8, 9]:

- Establish the risk management context (e.g. the scope, compliance obligations, approaches/methods to be used and relevant policies and criteria such as the organization's risk tolerance or appetite)
- Quantitatively or qualitatively assess risks, taking into account the information assets, threats, existing controls and vulnerabilities to determine the likelihood of incidents or incident scenarios, and the predicted business consequences if they were to occur, to determine a 'level of risk'
- Threat identification
- Monitor and review risks, risk treatments, obligations and criteria on an ongoing basis, identifying and responding appropriately to significant changes.

### 2.2 NIST Model

The National Institute of Standards and Technology (NIST) is a guide that describes the risk management methodology, and how the risk management process is tied to the process of system authorization.

The NIST was developed an information security for information systems. The framework aim's is to improve information security, strengthen risk management processes, and encourage reciprocity among federal agencies. The model encompasses three processes [10] risk assessment, risk mitigation, and evaluation and assessment:

- The identification and evaluation of risks and risk impacts and likelihood, and recommendation of risk-reducing measures. The assessment of security risk will be done using quantitative or qualitative approaches.
- The risk mitigation involves prioritizing, implementing, and maintaining the appropriate risk-reducing measures recommended from the risk assessment process. This step aims to implement the most appropriate controls to decrease mission risk to an acceptable level, with minimal adverse impact on the organization's resources and mission.
- The risk evaluation process to ensure a successful risk management program.

### **2.3 Operationally Critical Threat, Asset, and Vulnerability Evaluation Method (OCTAVE)**

OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation) is a risk-based strategic assessment and planning technique for security which was developed by Software Engineering Institute of Carnegie Mellon University in USA [14]. The method aims at examining organizational and technological issues as well as defining an organization's security strategy and plan. It consists of three steps: making a list of threat scenarios based on assets, recognizing the vulnerabilities about major facilities, and assessing the risk and developing security strategies.

The first step allows identifying assets of the system, security requirements (confidentiality, integrity and availability ... etc.), threat profiles and main vulnerabilities by interviewing some people during workshops. The second step identifies vulnerabilities that expose those threats and creates risks to the organization. The last step develops a practice-based protection strategy and risk mitigation plans to support the organization's missions and priorities.

### **2.4 Information Security Risk Management Framework for the Cloud Computing Environments**

The work presents a qualitative information risk management framework for better understanding critical areas of focus in Cloud Computing environment and identifying threats and vulnerabilities. The qualitative risk analysis proposed method is used to approach risk assessment and rank severity of threats by using classes such as low, medium and high of probabilities and damages for cloud providers. That is, to help controlling their security position and then to proceed to risk mitigation [11].

The framework has seven processes including: processes selecting relevant critical areas, strategy and planning, risk analysis, risk assessment, risk mitigation, assessing and monitoring program, and risk management review. Each process will be necessary to clarify specific roles, responsibilities, and accountability for each major step in the process [11].

At first, the method highlights the area of concern for Cloud Computing environment. For example if you are a SaaS provider, you may select application security, identify access management or assess threats and risks of vulnerabilities to organization. After proposing a strategy and planning process, the risk analysis step allows identifying threat sources and essential vulnerability in Cloud Computing in order to protect hosts, network devices, and applications from attacks against known vulnerabilities. The risk assessment step was divided into four major processes: likelihood determinations, impact analysis, risk determination and control recommendations. It represents the probability that a potential vulnerability could be exercised by a given threat source in qualitative way (high, medium, low) and determines the adverse impact resulting from a successful threat exercise of vulnerability and finally it represents, in a qualitative way, the risk levels and control recommendations to reduce this risk in CC



system. In the risk mitigation step, the cloud provider develops risk treatment plans (RTP) to mitigate vulnerability and threat. Finally, the cloud provider should monitor a risk treatment plan.

### 3 Cloud Computing System and Cyber Security Challenge in IaaS

Cloud computing provides cloud individuals and organizations with a more efficient, flexible and cost reduction services like computing resources. However, this environment is prone to several security threats risk [3, 20, 21, 25, 28, 29] and especially the infrastructure as a service layer due to the use of new technologies like virtualization that create more hard threats. In addition, IaaS consists of several components that have been developed to be used together in a shared and outsourced environment which carries several security concerns.

We detailed in this section the security issue of each component and discuss the proposed solutions and recommendations. Then, we present threats that threaten this cloud layer.

#### 3.1 IaaS Components

IaaS delivery model consists of several components that are connected together to run system's operations. However, employing those components together in a shared environment carries multiple security issues. Security issues in IaaS components are described as follows [12, 19, 20, 28–32].

**Service Level Agreement.** Cloud Computing emerges a set of technologies, and using SLA in cloud is the solution to guarantee acceptable level of quality of service. SLA is the contract between the service provider and the client to make the trust for quality of services and guaranty uptime. In the other hand, monitoring and enforcing the SLA in a dynamic environment is a big concern for example in Cloud systems, it is necessary to monitor this quality attributes continuously [16]. The solution is the Web Service Level Agreement (WSLA) framework [17] which is developed for SLA monitoring and enforcement in Service Oriented Architecture (SOA). WSLA manages SLA in Cloud Computing environment by allowing the third parties innovation with task of maintaining the SLA provisions in cloud computing [30].

**Utility Computing.** Utility computing plays a curtail role in Cloud Computing where users can only pay for usage service time. It offers two of the main features of the CC namely scalability, and pay-as-you-go usage service. However, several issues can threaten the usage of this technology as it can be attractive targets of attackers. Attacker may aim to access services without paying or can go further to drive specific client bill to unmanageable level. In such cases, keeping system healthy and well-functioning is a responsibility of provider [12, 19].

**Cloud Software.** Cloud software is that key which joins the cloud components together that they can act as single component [18]. Cloud system uses several software that can be either open source or commercially close source which are vulnerable to several attacks and bugs. For instance, cloud users use the protocol SOAP to exchange structured information in the implementation of web services in computer networks. It uses XML for its message format negotiation and transmission [19, 29].

WS-Security, a standard extension for security in SOAP, addresses the security for web services. It determines SOAP header that carried WS-Security extensions and determines how existing XML security standards like XML signature and XML encryption are applied to SOAP messages. However, attackers can attack against the XML services security standards and attack against the web services that can lead to break the communication of services. One solution to avoid these types of attacks is the XML encryption where data is encrypted and then decrypted to get the original [19, 31].

**Networks and Virtualization.** Internet connectivity plays an important role in delivering a service over the internet. There are some issues in networks and internet connectivity likes: Man in the middle attack where a hacker manipulates the network connectivity by creating middle man addressing, from where the attacker can access all the confidential data and permissions. Another type of attack is flooding attack, in which an unauthorized user sends a huge amount of request that lead more chances to attack on that demand. The possible solutions are traffic encryption by using point to point protocols that encrypt the connectivity to avoid the externals attacks [19, 22, 32]. Another countermeasure is, by proper network monitoring on services that whether all networking parameters are working properly or not. Externals attacks can also avoided by implementing firewalls, use of Intrusion Detection System and Intrusion and Prevention System that not only keep external threats out, but it can also alert you to more subtle problems by intercepting outgoing data as well [19, 22, 32].

**Computer Hardware.** IaaS offers an interface to a set of distributed physical resources (like network components, CPUs, and storage devices) to deliver a shared services to serve multiple consumers. Virtualization offers a secure share of the computer resources and a controlled communication on hardware and network level. Even though the private organizations used to move the hardware components into locked rooms accessible only by the authorized and trusted persons to protect the resources, a study showed that over 70% of all attacks on organizations' sensitive data and resources occurred internally (i.e., from inside the organization itself) [20].

IaaS providers play an essential role in protecting clients' data. Whatever the level of the data security, it can be part of retired or replaced storage devices. Usually, companies don't have restrictive policy to manage the retired devices that could be accidentally devolved to untrusted people. Each organization is supposed to assure clients' data security along its life cycle (It may use encryption technique) [28].

**Security Threats.** As a basic technology for Cloud Computing services, virtualization facilitates aggregation of multiple separate systems into a single hardware platform by virtualizing the computing resources. Virtualization plays a crucial role in sharing computer resources in IaaS of cloud system however it addresses plenty of security

problems. The other factor is the reliability of data stored within cloud provider's hardware [21]. In addition, IaaS layer consists of several components that have integrated to deliver computing resources as a service to their cloud subscribers.

In this section, we discuss virtualization risks and vulnerabilities that affect particularly IaaS components in addition to the proposed security solutions and recommendation.

**Data Leakage Protection and Usage Monitoring.** Data stored in the cloud must be kept confidential. The problem for these stored data is to know who is accessing the information, how it is accessed, location from where it is accessed and what happened to accessed information later. These concerns can be resolved by up to date right data managing services by restricting the data usages. The usage should be monitored continuously.

**Logging and Reporting.** To ensure effective and suitable deployment of IaaS proper and useful logging and reporting techniques must be applied. Robust logging and reporting solutions are crucial for service management and optimization. They help to keep track of where the information is, who accesses it, which machines are handling it and which storage arrays are responsible for it.

**Authentication and Authorization.** It is the most crucial countermeasure a system has to maintain. For every application, a user name and password is not most secure authentication mechanism. Therefore the cloud provider must use a multifactor authentication.

**End to End Encryption.** IaaS needs to use encryption techniques. For example whole disk encryption allows data encryption including user files on the disk which prevents offline attacks.

## 4 Risk Management Model for Cloud Computing System: Infrastructure as a Service

The IaaS layer for Cloud Computing system is prone to several security risks as it is discussed in previous section. We propose in this section a risk management approach for cloud computing system in order to reduce security threats risks in the IaaS layer.

The developed approach is based on the NIST risk management framework for information systems [10]. We use this approach to propose a generic model for security risk management for Cloud Computing system. We group the NIST methodology in three main stages as shown in Fig. 1:

- Step 1: Security risk assessment: in this step we identify threats and vulnerabilities of the information system.
- Step 2: Risk mitigation: in this step we security risk assessment approach to evaluate security risk.
- Step 3: Security risk control: The step recommends the security controls to be used.



**Fig. 1.** The NIST security management process.

#### 4.1 Threat and Vulnerability Identification

Cloud Computing offers all the advantages of a public utility system, in terms of economy of scale, flexibility and convenience but it raises substantial issues such as loss of control and loss of security [1, 11, 13, 23, 25, 27]. Security is the most challenges in cloud environment. Indeed, by trusting critical data to a service provider (externalization of service), a user (whether an individual or an organization) takes risks with the availability. The sections below highlight security threats related to Cloud Computing technology and their potential risks.

Virtualization technology causes major security risks for CC systems [22, 23, 25]. In fact, a Cloud Computing system is threatened by many types of attacks including security threats between the subscriber and the datacenter, the hypervisor and the VMs and also between the VMs themselves as mentioned in [11, 24–26]. Each of these types of attacks is examined in more detail below as it was defined in [11, 24–26]:

- Monitoring Virtual Machines from host
- Virtual machine modification
- Threats on communications between virtual machines and host
- Placement of malicious VM images on physical systems
- Flooding attacks
- Denial of service (DoS)
- Data loss or leakage
- Malicious insiders
- Account, service and traffic hijacking
- Abuse and nefarious use of Cloud Computing
- Insecure application programming interfaces
- Monitoring VMs from other VMs
- Virtual machine mobility
- Threats on communications between virtual machines

For Cloud Computing systems security risks affect mainly the following security requirements: availability, confidentiality and integrity [25, 27].

#### 4.2 Risk Analysis

From the risk assessment literature, a number of metrics has evolved to measure security risks. In fact, we have two types of metrics: qualitative and quantitative models. There are few qualitative models that estimate the security risks like GRAMM,

SecAgreement, and QUIRC [1–3, 9, 11, 14] while the other works are based on qualitative models like SLE, ALE, MTTF, MTBF, MFC and the M<sup>2</sup>FC [4, 11, 23–26]. Qualitative models are based on numerical values to estimate the security risk. However, qualitative approaches can be taken by defining risk in more subjective terms. It does not use variable values to estimate risks but rather evaluate qualitatively the influence of each variable on the risk.

Therefore, we use a quantitative model to assess security risk in Cloud Computing environment. We used the Multi-dimensional threat classification model (M<sup>2</sup>FC), presented in [26], to analyze security risk. This method assesses the cost of the failure of an information system security with regards to system's stakeholders, threats dimensions (architectural components, deployment sites...), and security threats.

The M<sup>2</sup>FC model:

- Varies by stakeholder, and takes into account the variance of the stakes that a stakeholder has in meeting each security requirement.
- Takes into account threats perspectives to reduce security risk to each system.
- Considers changes in systems like changes in the deployment, components and changes in user access policies.
- Takes into account threats dimensions and perspectives aspect and allows identifying critical dimensions that cause the biggest costs.

**Definition of Multi-dimensional Mean Failure Cost.** The Multi-dimensional Mean Failure Cost (M<sup>2</sup>FC) assessment model, proposed in [26], takes into account the stakeholders assessment of the cost related to their requirements with regard to the elements of two dimensions. So, we can say that the model is stakeholder based. That is why, in the following model, the set H of stakeholders and the set R of their requirements are distinguished from the set of the leading dimension and the set of the other considered dimensions.

- The Model

Let S be the set of elements in the leading dimension, D be the set of elements of the other considered dimension, H be the set of stakeholders, R is a set of requirements, and T be a set of threats. For every element  $s \in S$ , we define the Multi-dimensional Mean Failure Costs  $M(s; D)$  of elements as follows [26]:

$$M(s; D) = V_s \circ PFR_s \circ C_s \circ P_s \quad (1)$$

Where:

- We denote  $\circ$  by the matrix multiplication operation.
- $V_s$  is a matrix that each entry  $(i; j)$  represents the value of the stake that stakeholder  $H_i$  has in meeting requirement  $R_j$ .
- $PFR_s$  is a matrix that each entry  $(i; j)$  represents the probability of failing requirement  $R_i$  due to a failure originating from element  $d_j \in D$ .
- $C_s$  is a matrix that each entry  $(i; j)$  represents the probability that an element  $d_i \in D$  fails once the threat  $t_j$  has materialized.

- $P_s$  is a column vector that each entry  $i$  represents the probability that threats materialize during unitary period of operation.

If we consider  $S$  is a set of site on which the system is deployed, and  $D$  is the set of components located in the site  $s \in S$ , then  $M(s; D)$  gives the Multi-dimensional Mean Failure Cost per site and its components.

**Illustrative Example.** We illustrate by a real example the use of the  $M^2FC$  model on a simple case study adapted from similar cases studies found in [3, 4, 6] to estimate the security of a system in terms of loss incurred by each stakeholder. We consider the previous case study adopted in our previous work in [26].

In this case study, we considered the architectural perspective in which we consider as dimensions: the deployment sites dimension and the architectural components. Our assessment varies according to the stakes that each stakeholder has in meeting each security requirement per system site. We take the deployment or sites dimension as the leading dimension. For each site, we have the lists of stakeholders, security requirements, components, and threats.

We assume that the considered system is deployed on two sites. We have  $S = \{Site1, Site2\}$  and the component dimension  $D$  has the following elements which is equals to  $D = \{Browser (Br), Web server (WS), Proxy server (PS)\}$ . We recognize two stakeholders for this example namely: *local user* and *external user*. We consider as well the set  $R$  of requirements to include the three facets of security  $R = \{Availability, Integrity, Confidentiality\}$ . We consider that we dealing with the set  $T = \{Virus (Vr), Data Bases Attacks (DBA), Denial of service (DoS)\}$  of threats.

In this case, we assumed that both locations have the same stakeholders and the same security requirements. Table 1 gives the values of the matrix of stakes  $Vs$  in thousands of dollars per hour (\$K/H). Tables 2, 3 and 4 represent respectively the values of matrices  $PFRs$  and  $Cs$  and the vector  $Ps$ . Finally we obtained the vector of Multidimension Mean Failure Costs  $M(Site, Component)$  as it is shown in Table 5 for the components of each site.

**Table 1.** Matrix of stakes: cost of failing a security requirement in \$K/H [26].

	Security requirement		
	Availability	Confidentiality	Integrity
Stakeholders			
Local user	0,5	0,9	0,9
External user	0,03	0,04	0,9

**Risk Mitigation.** Mitigation involves fixing the flaw or providing some type of compensatory control to reduce the likelihood or impact associated with the flaw [10, 19, 20]. The objective of this step is to mitigate the risks by introducing appropriate countermeasures. This aims is thus to develop a risk mitigation plans to support the organization’s missions and priorities.

The hierarchical linear structure of the  $M^2FC$  calculation formula, produced from three matrices and a vector, allows diversity in the control classes of the latter. We can

**Table 2.** Probabilities of failure requirements matrix for sites [26].

	Components						
	Site 1				Site 2		
	Br	WS	PS	No_Failure (NoF)	Br	WS	No_Failure (NoF)
Security requirements							
Availability	0,2	0,1	0,3	0,4	0,2	0,3	0,5
Confidentiality	0,1	0,1	0,2	0,6	0,3	0,3	0,7
Integrity	0,1	0,2	0,4	0,3	0,2	0,2	0,6

**Table 3.** Probabilities of failure components matrix for sites [26].

	Threats						
	Site 1				Site 2		
	Vr	DBS	No_Threat (NoT)	VR	DBA	DoS	No_Threat (NoT)
Security requirements							
Br	0,2	0,5	0,3	0,2	0,2	0,1	0,5
WS	0,5	0,4	0,1	0,2	0,35	0,1	0,35
PS	0,3	0,6	0,1	–	–	–	–
No_Failure (NoF)	0,1	0,3	0,6	0,11	0,2	0,09	0,6

**Table 4.** Threats probabilities occurrence vector for sites [26].

Threats	Probability	
	Site 1	Site 2
Vr	0,03	0,04
DBA	0,02	0,015
DoS	–	0,05
No_Threat (NoT)	0,95	0,840

**Table 5.** Stakeholders multi-dimensional mean failure cost for sites considering the dimension components [26].

Stakeholders	M <sup>2</sup> FC (\$K/h)	
	Site 1	Site 2
Local user	0,7914	1,1611
External user	0,2749	0,4503

control the multi dimensional mean failure costs by controlling any one of these factors. For the sake of argument, we classify security measures according to which factor they involve. We briefly discuss this classification, below:

- Mitigation measures: controlling the matrix of stakes (Vs). Countermeasures in this family are designated to reduce the impact of failures on costs incurred by systems 'stakeholders.
- Failure tolerance measures: Controlling the Probabilities of failure requirements per threat dimension matrix (PFRs). This kind of measures aims to minimize the impact of threats dimension (like a component, period of season...) failures on system failures by enhancing the failure tolerance of the system.
- Fault tolerance measures: Controlling the Probabilities of failure threats dimensions matrix (Cs). Countermeasures are designates to minimize the incidence of threats dimension failures by eliminating or mitigating component vulnerabilities.
- Evasive measures: Controlling the threats probabilities occurrence vector (Ps). This category of measures aims to conceal component vulnerabilities, or otherwise making it harder to exploit them.

As virtualization is a fundamental technology for Cloud Computing systems (virtual machines), that allows aggregation of multiple unconnected systems into a single hardware platform by virtualizing the computing resources (like network, CPUs, memory, and storage) in the IaaS layer. Virtualization lets reducing management complexity of physical computing resources. Virtualization provides multi tenancy and scalability, and these are two significant characteristics of Cloud system [21, 22].

As has been shown in the previous section that the infrastructure as a service layer for Cloud Computing environments risky by several types of attacks and vulnerabilities, we will focus on virtualization security effect on this layer.

In this paper we will focus on hypervisor security. To ensure a good hypervisor security, the organization can adopt the solution of hypervisor duplication.

- Duplicate hypervisor

The hypervisor or the virtual machine manager is the brain of virtualization technology. It is a software, firmware, or hardware that creates and runs virtual machines (VMs). The hypervisor serves as a virtual operating platform that executes the guest operating system for an application. Host servers are designed to run multiple VMs cloud user to run more operating system at the same time over on a single physical system, also provide hardware abstraction to run end user OS and multiplexes underlying hardware resources.

Virtualization also provides several advantages for cloud computing environments, including resource sharing, VM isolation, and load balancing. In a cloud environment, these capabilities enable scalability, high utilization of pooled resources, rapid provisioning, workload isolation, and increased uptime [11, 22].

Renting double hypervisor is creating duplication in the Cloud Computing system, and as service, is putting the same components (VMs, Operation system...) and configuration of the system in distinct hypervisor. This solution reduces the probability that the components fail to the half. In our case we compute the  $M^2FC$  using (matrix of stakes, matrix of probabilities of failure requirements, matrix of probabilities of failure threats dimension) matrix and the probabilities of threat vector. If we adopt this solution each probability in probabilities of failure threat dimension matrix will be as follow:



$$M(s; D) = V_s \circ PFR_s \circ C_s \circ \frac{1}{2} P_s \tag{2}$$

Using the 3 and the new probabilities of threats vector (Ps), the new M<sup>2</sup>FC values are shown in Table 6.

**Table 6.** New values of stakeholders multi-dimensional mean failure cost for sites considering the dimension components.

Stakeholders	M <sup>2</sup> FC (\$K/h)	
	Site 1	Site 2
Local user	0,3957	0,5805
External user	0,1347	0,2251

The new values of M<sup>2</sup>FC appear more interesting but we can't decide if this solution is better than others without computing the return on investments. In fact, M<sup>2</sup>FC values are lower than the first ones.

## 5 Conclusion

With the advantages of flexibility, storage, sharing and easy accessibility, Cloud Computing is a trending technology driven world. This new technology offers services at different layers, simulating the functions performed by applications, operating systems, or physical hardware. Cloud Computing services are given as three layers of services which provide infrastructure resources, application platform and software as services to the cloud subscribers.

While the cloud offers these advantages, until some of the risks are lack of security. For this purpose our paper focus on security risks on infrastructure as a service layer for Cloud Computing systems because it includes the main components and computing resources to run all cloud services.

We present in this paper main components of the infrastructure as a service (IaaS) layer and security challenges of them. We propose as well an information risk management approach for better understanding critical areas of interest in cloud computing environment and to identifying security threats and vulnerabilities. This approach is covering all of cloud infrastructure models and it is based on a quantitative model to evaluate security risks namely the multi dimensional mean failure cost (M<sup>2</sup>FC).

## References

1. Barry, B., LiGuo, H.: Value-based software engineering: a case study. *IEEE Comput.* **36**, 33–41 (2003)

2. Saripalli, P., Walters, B.: QUIRC: a quantitative impact and risk assessment framework for cloud security. In: *The Proceedings of the IEEE 3rd International Conference on Cloud Computing*, pp. 280–288 (2009)
3. An, M., Chen, Y., Baker, C.J.: A fuzzy reasoning and fuzzy-analytical hierarchy process based approach to the process of railway risk information: a railway risk management system. *Inf. Sci.* **181**, 3946–3966 (2011)
4. Wang, J.A., Xia, M., Zhang, F.: Metrics for information security vulnerabilities. In: *Proceedings of Intellect base International Consortium*, vol. 1, pp. 284–294 (2009)
5. Yuri, D., Leon, G., Cees, L.: *Web services and grid security vulnerabilities and threats analysis and model*, Bas Oudenaarde, Advanced Internet Research Group, University of Amsterdam, Kruislaan 403, NL-1098 SJ Amsterdam, The Netherlands (2000)
6. ISO/IEC 27005: *Information Technology—Security Techniques—Information Security Risk Management*, International Organization for Standardization (2007)
7. Emam, A.H.M.: Additional authentication and authorization using registered email-ID for cloud computing. *Int. J. Soft Comput. Eng.* **3**, 110–113 (2013)
8. ISO. BS ISO 31000: *Risk management. Principles and guidelines* (2009)
9. ISO. BS ISO/IEC 27005: *Information technology. Security techniques. Information security risk management* (2011)
10. Gary, S., Alice, G., Alexis, F.: *Risk Management Guide for Information Technology Systems Recommendations of the National Institute of Standards and Technology, NIST Special Publication 800-30* (2002)
11. Xuan, Z., Nattapong, W., Hao, L., Xuejie, Z.: Information security risk management framework for the cloud computing environments. In: *10th IEEE International Conference on Computer and Information Technology (CIT 2010)* (2010)
12. Padhy, R.P., Patra, M.R., Satapathy, S.C.: Cloud computing: security issues and research challenges. *Int. J. Comput. Sci. Inf. Technol. Secur.* **1**(2), 136–146 (2011)
13. Sangroya, A., Kumar, S., Dhok, J., Varma, V.: Towards analyzing data security risks in cloud computing environments. In: Prasad, S.K., Vin, H.M., Sahni, S., Jaiswal, M.P., Thipakorn, B. (eds.) *ICISTM 2010. CCIS*, vol. 54, pp. 255–265. Springer, Heidelberg (2010). [https://doi.org/10.1007/978-3-642-12035-0\\_25](https://doi.org/10.1007/978-3-642-12035-0_25)
14. Alberts, C., Dorofee, A., Stevens, J., Woody, C.: *Introduction to the OCTAVE approach*. Software Engineering Institute (2003)
15. Kevin, S.: *Virtualisation as a Blackhat Tool*. In: *Network Security*, pp. 4–7. Elsevier, New York (2007)
16. SLA Management Team: *SLA Management Handbook*, 4th edn. Enterprise Perspective (2004)
17. Frankova, G.: Service level agreements: web services and security. In: Baresi, L., Fraternali, P., Houben, G.-J. (eds.) *ICWE 2007. LNCS*, vol. 4607, pp. 556–562. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-73597-7\\_54](https://doi.org/10.1007/978-3-540-73597-7_54)
18. Patel, P., Ranabahu, A., Sheth, A.: Service level agreement in cloud computing. In: *Cloud Workshops at OOPSLA 2009* (2009)
19. Bineet, K.J., Mohit, K.S., Bansidhar, J.: Security threats and their mitigation in infrastructure as a service. *Perspect. Sci.* **8**, 462–464 (2016)
20. Wesam, D., Ibrahim, T.: *Infrastructure as a service security: challenges and solutions* (2008)
21. Subashini, S., Kavitha, V.: A survey on security issues in service delivery models of cloud computing. *J. Netw. Comput. Appl.* **34**, 1–11 (2011)
22. Intel IT Center: *Planning Guide: Virtualization and Cloud Computing*, White paper (2013)
23. Mouna, J., Latifa, B.R.: A multi-dimensional mean failure cost model to enhance security of cloud computing systems. *Int. J. Embed. Real Time Commun. Syst. (IJERTCS)* **7**(2), 1–14 (2016)

24. Mouna, J., Latifa, B.R.: Mean failure cost extension model towards a security threats assessment: a cloud computing case study. *J. Comput.* **10**, 184–194 (2015)
25. Mouna, J., Latifa, B.R., Anis, B.A., Ali, M.: A cyber security model in cloud computing environments. *J. King Saud Univ. Comput. Inf. Sci.* **25**, 63–75 (2013)
26. Mouna, J., Latifa, B.R., Ridha, K.: A multidimensional approach towards a quantitative assessment of security threats. In: *ANT/SEIT 2015*, vol. 52, pp. 507–514 (2015)
27. Mouna, J., Latifa, B.R.: Surveying and analyzing security problems in cloud computing environments. In: *CIS 2014*, pp. 689–693 (2014)
28. Lawal, B.O., Ogude, C., Abdullah, K.K.A.: Security management of infrastructure as a service in cloud computing. *Afr. J. Comput. ICT Ref. Format* **6**, 137–146 (2013)
29. Ibrahim, A.S., Hamlyn-Harris, J., Grundy, J.: Emerging security challenges of cloud virtual infrastructure. In: *The Asia Pacific Software Engineering Conference 2010 Cloud Workshop* (2010)
30. Jaiswal, P.R., Rohankar, A.W.: Infrastructure as a service: security issues in cloud computing. *IJCSMC* **3**, 707–711 (2014)
31. Jenson, M., Schwenk, J., Gruschka, N., Lo Iacono, L.: Ontechnical security issues in cloud computing. *IEEE* (2009)
32. Krutz, R.L., Vines, R.D.: *Cloud Security*. Wiley Publication, Indianapolis (2014)

## Author Index

- Alepis, Efthimios 93  
Ali, Sameer 131
- Ba, Haihe 516  
Bai, Jin 256
- Cao, Buqing 308  
Cao, Xuefei 56  
Cao, Yu 31  
Chang, Rui 381  
Cheng, Hanni 17  
Cheng, Xiawei 241
- Dang, Lanjun 56  
Dang, Tran Khanh 501, 561  
de Oliveira, Italo R. 584  
Deng, Hua 420  
Dou, Wanchun 280  
Duan, Guihua 42
- Elahi, Haroon 169
- Fan, Kai 56  
Feng, Yaokai 461  
Fregnani, Jose A. 584  
Fu, Chong 432  
Fu, Yulong 56
- Gao, Bei-li 432  
Gao, Wuqiang 201, 332  
Gong, Weiwei 298  
Guo, Guibing 256  
Guo, Ying 31
- Haque, Reazul 131  
Hei, Xiaojun 17  
Hori, Yoshiaki 461  
Hu, Xiao-hua 368  
Huang, Xianfei 121
- Ji, Sai 395  
Jiang, Shengyi 229
- Jiang, Wenjun 420  
Jouini, Mouna 594
- Kaspin, Ir. Rizaludin 131  
Kwang, Lee Ching 131  
Kwuimi, Raoul 571
- La Marra, Antonio 545  
Leite, Alessandro F. 584  
Leng, Bo 201  
Li, Jian-bin 1, 201  
Li, Jiantao 420  
Li, Jin 487  
Li, Ping 487  
Li, Shuai 446  
Li, Simin 241  
Li, Xu 169  
Li, Yanmei 87  
Li, Yuanlong 108  
Li, Yunshi 516  
Liang, Jianwu 31  
Liu, Akang 241  
Liu, Dengzhi 395  
Liu, Jianxun 308  
Liu, Meilin 446  
Liu, Pin 186  
Liu, Qi 395  
Liu, Qin 186  
Liu, Yuan 256  
Liu, Zhenlan 1  
Long, Min 368  
Long, Zhenyue 409  
Luo, Entao 42
- Ma, Zhaohui 474  
Man, Yujia 381  
Martinelli, Fabio 530, 545  
Massacci, Fabio 501  
Miao, Tiantian 395  
Mori, Paolo 545
- Ouyang, Liubo 269

- Patsakis, Constantinos 93  
 Peng, Fei 368  
  
 Qi, Fang 87, 108  
 Qi, Lianyong 280  
  
 Rabai, Latifa Ben Arfa 594  
 Ren, Jiangchun 516  
 Ren, Ju 332  
 Ren, Lu 381  
 Ren, Milin 353  
 Ren, Xueqi 474  
 Ribeiro, Vitor F. 584  
 Rizos, Athanasios 545  
  
 Sakurai, Kouichi 461  
 Saracino, Andrea 545  
 She, Liang 201  
 Sheikhalishahi, Mina 530  
 Shen, Guiquan 409  
 Shen, Jian 395  
 Shou, Zhaoyu 241  
 Son, Ha Xuan 501, 561  
 Song, Qiang 1  
 Song, Yunkui 409  
 Sun, Jing 432  
  
 Tan, Saw Chin 131  
 Tan, Zhenhua 256  
 Tang, Hua 474  
 Tang, Hui 269  
 Tang, Mingdong 308  
 Tang, Yankun 139  
 Tang, Zhe 87, 108  
 Thampi, Sabu M. 212  
 Thi, Que Nguyet Tran 561  
 Tian, Yuan 269  
  
 Usha, Athira 212  
  
 Van, Huy Luong 561  
  
 Wang, Changji 229  
 Wang, Chen 395  
 Wang, Gaocai 121  
 Wang, Guojun 42, 169, 186, 332  
 Wang, Jialin 17  
 Wang, Nao 121  
 Wang, Qifan 42  
  
 Wang, Xiaodong 70, 78  
 Wang, Xingwei 256  
 Wang, Xinming 474  
 Wang, Xue 432  
 Wang, Yongjun 516  
 Wang, Zhiying 516  
 Wei, Lihao 409  
 Wei, Songjie 353, 446  
 Weigang, Li 584  
 Wu, Faguo 321  
 Wu, Jiechao 474  
 Wu, Qianqian 353  
  
 Xu, Cheng 153  
 Xu, Guangwu 139  
 Xu, Hongyun 153  
 Xu, Mengzhen 153  
 Xu, Yang 201, 332  
 Xue, Minhui 17  
  
 Yan, Zheng 139  
 Yao, Wang 321  
 Ye, Heng 487  
 Yin, Hongjian 344  
 Yin, Qing 381  
 Yu, Shui 280  
 Yuan, Yuan 229  
 Yusoff, Zulfadzli 131  
  
 Zeng, Jijun 409  
 Zeng, Quanrun 201, 332  
 Zeng, Sumeng 308  
 Zhang, Boyun 269  
 Zhang, Gao-yuan 432  
 Zhang, Gen 298  
 Zhang, Leyou 344  
 Zhang, Peng 139  
 Zhang, Xiao 321  
 Zhang, Xuyun 280  
 Zhang, Yaoxue 332  
 Zhao, Gansen 474  
 Zheng, Zhiming 321  
 Zhou, Huaizhe 516  
 Zhou, Xu 298  
 Zhou, Zhibin 186  
 Zhu, Daxin 70, 78  
 Ziri, Salvatore Renato 131  
 Zuva, Tranos 571