

Handbook of Mobile Data Privacy

Aris Gkoulalas-Divanis • Claudio Bettini
Editors

Handbook of Mobile Data Privacy

Editors

Aris Gkoulalas-Divanis
IBM Watson Health Headquarters
Cambridge, MA, USA

Claudio Bettini
Dipartimento di Informatica
University of Milan
Milan, Italy

ISBN 978-3-319-98160-4 ISBN 978-3-319-98161-1 (eBook)

<https://doi.org/10.1007/978-3-319-98161-1>

Library of Congress Control Number: 2018958707

© Springer Nature Switzerland AG 2018

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.

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

*Dedicated to my parents, Aspa and Dimitris,
and to my lovely wife Elena for all their
support.*

–Aris Gkoulalas-Divanis

*Dedicated to my parents, Anna Maria and
Giovanni, and to my son Matteo.*

–Claudio Bettini

Preface

The advances in mobile devices and positioning technologies, together with the progress in spatiotemporal database research, have made possible the tracking of mobile devices (and their human companions) at a very high accuracy while supporting the efficient storage of mobility data in data warehouses. This has provided the means to collect, store, and process mobility data of an unprecedented quantity, quality, and timeliness. As ubiquitous computing pervades our society, user mobility data represents a very useful, but also extremely sensitive, source of information. On the one hand, the movement traces that are left behind by the mobile devices of the users can be very useful in a wide spectrum of applications such as urban planning, traffic engineering, and environmental pollution management. On the other hand, the disclosure of mobility data to third parties may severely jeopardize the privacy of the users whose movement is recorded, leading to abuse scenarios such as user tailing and profiling.

A significant body of research work has been conducted in the last 15 years in the area of mobility data privacy, along a number of important research directions such as privacy-preserving mobility data management, privacy in location sensing technologies and location-based services, privacy in vehicular communication networks, privacy in location-based social networks, and privacy in participatory sensing systems. This work has identified important privacy gaps in the use of human mobility data and has resulted to the adoption of international laws for location privacy protection (e.g., in EU, the USA, Canada, Australia, New Zealand, Japan, Singapore), as well as to a large number of interesting technologies for privacy-protecting mobility data, some of which have been made available through open-source systems and featured in real-world applications.

The overarching aim of this book is to survey the field of mobility data privacy and to present the fundamental principles and theory, as well as the state-of-the-art research, systems, and applications, to a wide audience including non-experts. Emphasis in the book is given toward coverage of the most recent directions in mobility data privacy. The structure of the book closely follows the main categories of research works that have been recently undertaken to protect user privacy in the context of mobility data and applications. After the first introductory chapters,

which cover the fundamentals around the offering of privacy in mobility data, each subsequent chapter of the book surveys an important research problem related to mobility data privacy and discusses the corresponding privacy threats that have been identified and the solutions that have been proposed. The last part of the book is devoted to state-of-the-art systems for mobility data privacy, as well as to real-world applications where privacy-protection techniques have been applied.

We would like to note that this book is primarily addressed to computer science and statistics researchers and educators, who are interested in topics related to mobility privacy. We expect that the book will be also valuable to industry developers, as it covers the state-of-the-art algorithms for offering privacy. To ease understanding by nonexperts, the chapters contain a lot of background material, as well as many examples and citations to related literature. By discussing a wide range of privacy techniques, providing in-depth coverage of the most important ones, and highlighting promising avenues for future research, this book also aims at attracting computer science and statistics students to this fascinating field of research.

Boston, MA, USA
Milan, Italy
June 2018

Aris Gkoulalas-Divanis
Claudio Bettini

Acknowledgments

We would like to thank all authors who contributed chapters to this handbook, for their valuable contributions. This work would not have been possible without their efforts. A total of 31 authors who hold positions in leading academic institutions and industry, in Italy, Hungary, Austria, Switzerland, Greece, France, Germany, the USA, and the UK, have contributed 15 chapters to this book. We sincerely thank them for their hard work and the time they devoted to this effort.

In addition, we would like to express our deep gratitude to all the expert reviewers of the chapters for their constructive comments, which significantly helped toward improving the organization, readability, and overall quality of the book.

Last but not least, we are indebted to Susan Lagerstrom-Fife and Caroline Flanagan from Springer, for their great guidance and advice in the preparation and completion of this handbook, as well as to the typesetting and publication team at Springer for their valuable assistance in the editing process.

Contents

1	Introduction to Mobility Data Privacy	1
	Aris Gkoulalas-Divanis and Claudio Bettini	
Part I Fundamentals for Privacy in Mobility Data		
2	Modeling and Understanding Intrinsic Characteristics of Human Mobility	13
	Jameson L. Toole, Yves-Alexandre de Montjoye, Marta C. González, and Alex (Sandy) Pentland	
3	Privacy in Location-Sensing Technologies	35
	Andreas Solti, Sushant Agarwal, and Sarah Spiekermann-Hoff	
Part II Main Research Directions in Mobility Data Privacy		
4	Privacy Protection in Location-Based Services: A Survey	73
	Claudio Bettini	
5	Analyzing Your Location Data with Provable Privacy Guarantees.....	97
	Ashwin Machanavajjhala and Xi He	
6	Opportunities and Risks of Delegating Sensing Tasks to the Crowd.....	129
	Delphine Reinhardt and Frank Dür	
7	Location Privacy in Spatial Crowdsourcing	167
	Hien To and Cyrus Shahabi	
8	Privacy in Geospatial Applications and Location-Based Social Networks	195
	Igor Bilogrevic	

9	Privacy of Connected Vehicles	229
	Jonathan Petit, Stefan Dietzel, and Frank Kargl	
10	Privacy by Design for Mobility Data Analytics	253
	Francesca Pratesi, Anna Monreale, and Dino Pedreschi	
Part III Usability, Systems and Applications		
11	Systems for Privacy-Preserving Mobility Data Management	281
	Despina Kopanaki, Nikos Pelekis, and Yannis Theodoridis	
12	Privacy-Preserving Release of Spatio-Temporal Density	307
	Gergely Acs, Gergely Biczók, and Claude Castelluccia	
13	Context-Adaptive Privacy Mechanisms	337
	Florian Schaub	
14	Location Privacy-Preserving Applications and Services	373
	Ioannis Boutsis and Vana Kalogeraki	
	Index	399

About the Authors



Gergely Acs CrySyS Lab, Budapest University of Technology and Economics, Budapest, Hungary (acs@crysys.hu).

Gergely Acs is an Assistant Professor at the Budapest University of Technology and Economics (BUTE), Hungary, and a member of the Laboratory of Cryptography and System Security (CrySyS). Before joining BUTE, Gergely was a research scholar and engineer at INRIA, France. His research focuses on different aspects of data privacy and security, including privacy-preserving machine learning, data anonymization, and privacy risk analysis. He received his MS and PhD degrees from BUTE.



Sushant Agarwal Institute for Management Information Systems, Vienna University of Economics and Business, Vienna, Austria (sushant.agarwal@wu.ac.at).

Sushant works on the European Union project SERAMIS as a PhD researcher at WU, where he explores questions related to privacy in pervasive computing. He received an MTech degree in Aerospace Engineering from the Indian Institute of Technology (IIT), Bombay, in 2014. During his studies, he interned at the Institute for Manufacturing (IfM) in the University of Cambridge and was involved in projects dealing with data quality and failure diagnosis for Boeing and Hitachi, respectively.



Claudio Bettini EveryWare Lab, Department of Computer Science, University of Milan, Italy (claudio.bettini@unimi.it).

Claudio Bettini is Professor of Computer Science at the University of Milan, where he leads the EveryWare laboratory at the Computer Science Department. Claudio received his PhD in Computer Science from the University of Milan in 1993. He has been a post-doc at IBM Research, NY, and, for more than a decade, an Affiliate Research Professor at the Center for Secure Information Systems at George Mason University, VA. His research interests cover the areas of mobile and pervasive computing, data privacy, temporal and spatial data management, and intelligent systems.

In 2009, he co-edited the Springer book *Privacy in Location-Based Applications*, offering a first overview of the main research results in the field. Among his many organizational activities, he acted as General Chair of IEEE PerCom 2017 and IEEE MDM 2013 and as TPC Chair of IEEE PerCom 2013. He is a member of the steering committee of IEEE PerCom and Associate Editor of the *Pervasive and Mobile Computing* journal and previously Editor of *The VLDB Journal* and of the *IEEE Transactions on Knowledge and Data Engineering*. In 2011, he founded EveryWare Technologies, a startup developing innovative mobile apps and services for privacy and assistive technologies. He is a senior member of the IEEE Computer Society.



Gergely Biczók CrySyS Lab, Budapest University of Technology and Economics, Budapest, Hungary ([bic-zok@crysys.hu](mailto:biczok@crysys.hu)).

Gergely Biczók is an Assistant Professor at the CrySyS Lab at the Budapest University of Technology and Economics (BME). He received his PhD (2010) and MSc (2003) degrees in Computer Science from BME. He was a Postdoctoral Fellow at the Norwegian University of Science and Technology (2011–2014) and at the Future Internet Research Group of the Hungarian Academy of Sciences (2014–2016). Previously, he was a Fulbright Visiting Researcher to Northwestern University (2007–2008) and held a researcher position at Ericsson Research (2003–2007). His research interests center around the economics of networked systems including privacy, incentives, and game theory.



Igor Bilogrevic Senior Research Scientist, Google Inc., Zurich, Switzerland (ibilogrevic@google.com).

Igor Bilogrevic is a Senior Research Scientist at Google, where he works on machine learning and data mining techniques in order to build novel privacy and security features in products. He obtained his PhD on the privacy of context-aware mobile networks from EPFL (Switzerland), where he investigated privacy issues at different layers of the network stack and proposed privacy-enhancing mechanisms for information sharing that work across different layers. He has previously worked in collaboration with the Nokia Research Center and PARC (a Xerox Company) on privacy challenges in pervasive mobile networks, encompassing data privacy, social community privacy, location privacy, information sharing, and private data analytics. His main areas of interest include applications of machine learning for privacy, private data analytics, contextual intelligence, applied cryptography, and user experience.



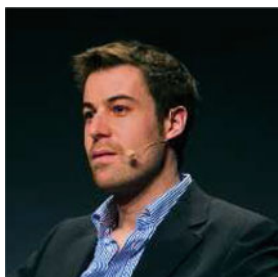
Ioannis Boutsis Athens University of Economics and Business, Athens, Greece (mpoutsis@aueb.gr).

Ioannis Boutsis received BSc (2009) and MSc degrees in Computer Science (2011) and PhD degree in Distributed Systems (2017), all from the Athens University of Economics and Business (AUEB). During his PhD, Ioannis worked in several EU projects including Insight-FP7 and NGHCS-ERC. Since 2017, he works for Amazon Video at the Amazon Development Centre, London, as a Software Development Engineer.



Claude Castelluccia Saint Ismier Cedex, France (claude.castelluccia@inria.fr).

Claude Castelluccia is a Research Director at Inria, in France, where he leads the PRIVATICS team (Privacy Models, Architectures and Tools for the Information Society). Claude has held visiting research positions at UC Irvine and Stanford University, USA. His past research interests include networking, Internet protocols, network security, and applied cryptography. His current research is on Internet privacy and security with a focus on anonymized analytics, data anonymization, data transparency, and Internet-based surveillance analysis. He is also interested in the economical and legal aspects of data privacy. He has chaired and participated in many Program Committees (ACM CCS, PETS, WiSec, etc.), co-founded the ACM WiSec conference, and advised more than 10 PhD students.



Yves-Alexandre de Montjoye Department of Computing, Imperial College, London, UK (demon-tjoye@imperial.ac.uk).

Yves-Alexandre de Montjoye is an Assistant Professor (Lecturer) at Imperial College London where he leads the Computational Privacy Group, and a Special Adviser to EC Commissioner Vestager. He is affiliated with the Data Science Institute and Department of Computing. Previously, he was Postdoctoral Researcher at Harvard University, working with Latanya Sweeney and Gary King, and he received his PhD from MIT, under the supervision of Alex “Sandy” Pentland. His research aims at understanding how the unicity of human behavior impacts the privacy of individuals in large-scale metadata datasets, and his work has been covered in *The New York Times*, BBC News, CNN, *Wall Street Journal*, *Harvard Business Review*, *Le Monde*, *Die Spiegel*, *Die Zeit*, and *El Pais*, and in reports of the World Economic Forum, United Nations, OECD, FTC, and the European Commission, as well as in talks at TEDxLLN and TEDxULg. Yves-Alexandre was recently named an Innovator under 35 for Belgium (TR35). He is a Fellow of the ID³ Foundation and the B.A.E.F. Foundation and a Research Associate at Data-Pop.



Stefan Dietzel Humboldt-Universität zu Berlin, Department of Computer Science, 10099 Berlin, Germany (stefan.dietzel@hu-berlin.de).

Dr. Stefan Dietzel is a Postdoctoral Researcher with the Chair of Computer Engineering at Humboldt-Universität zu Berlin. Stefan's research interests are in security, scalability, and privacy aspects of wireless distributed computing. In several projects, he is investigating how to create communication protocols that are scalable and efficient while, at the same time, protecting the users' privacy and being secure against insider and outsider attackers.

Stefan holds a doctoral degree in Computer Science from the University of Twente, The Netherlands. During his time as a doctoral candidate, Stefan worked at the Institute of Media Informatics and the Institute of Distributed Systems at Ulm University and at the Distributed and Embedded Systems group at the University of Twente; before, he studied media informatics at Ulm University in Germany.



Frank Dürr Institute of Parallel and Distributed Systems, University of Stuttgart, Stuttgart, Germany (frank.duerr@ipvs.uni-stuttgart.de).

Frank Dürr is a Senior Researcher and Lecturer at the Institute of Parallel and Distributed Systems (IPVS) at the University of Stuttgart in Germany. He received both his doctoral degree and diploma in Computer Science from the University of Stuttgart. Frank Dürr is currently leading the mobile and context-aware systems and software-defined networking (SDN) groups of the Department of Distributed Systems at IPVS. He was technical coordinator of the Collaborative Research Center (SFB) 627 "Spatial World Models for Mobile Context-Aware Applications (Nexus)," funded by the German Research Foundation (DFG). His research interests include mobile and pervasive computing—in particular, location privacy and mobile sensing—as well as software-defined networking (SDN) and time-sensitive networking (TSN).



Aris Gkoulalas-Divanis IBM Watson Health Headquarters, Cambridge, MA 02142-1123, USA (gkoulala@us.ibm.com).

Aris Gkoulalas-Divanis is the Technical Lead on Data Protection and Privacy for IBM Watson Health. He received his PhD in Computer Science from the University of Thessaly. His PhD dissertation was awarded the Certificate of Recognition and Honorable Mention in the 2009 ACM SIGKDD Dissertation Awards. Aris has been a Postdoctoral Research Fellow in the Department of Biomedical Informatics of Vanderbilt University (2009–2010) and a Research Scientist in IBM Research-Zurich (2010–2012) and in IBM Research-Ireland (2012–2016). His research interests are in the areas of privacy-preserving data mining, privacy in trajectories and location-based services, privacy in medical data, and knowledge hiding. In these areas, he has published more than 80 research works, including four Springer books, and he has applied for or being granted more than 25 patents. He is an Associate Editor of the *Knowledge and Information Systems* (KAIS) journal, the *IEEE Transactions on Information Forensics and Security* (T-IFS), the *International Journal of Research & Development Innovation Strategy* (IJRDIS), and the *International Journal of Knowledge-Based Organizations* (IJKBO). Since 2014 he is an Area Editor for *ACM Computing Reviews* (CR). Aris is a senior member of IEEE; a professional member of ACM, SIAM, and AAAS.



Marta C. González University of California, Berkeley, CA, USA (martag@berkeley.edu).

Marta C. González is an Associate Professor of City and Regional Planning at the University of California, Berkeley, and a Physics Research faculty in the Energy Technology Area (ETA) at the Lawrence Berkeley National Laboratory (Berkeley Lab). With the support of several companies, cities, and foundations, her research team develops computer models to analyze digital traces of information mediated by devices. They process this information to manage the demand in urban infrastructures in relation to energy and mobility. Her recent research uses billions of mobile phone records to understand the appearance of traffic jams and the integration of electric vehicles into the grid, smart meter data records to compare the policy of solar energy adoption, and card transactions to identify habits in spending behavior. Prior to joining Berkeley, Marta worked as an Associate Professor of Civil and Environmental Engineering at MIT, and a member of the Operations Research Center and the Center for Advanced Urbanism. She is a member of the scientific council of technology companies such as Gran Data, PTV, and the Pecan Street Project consortium.



Xi He Duke University, Durham, NC, USA (hexi88@cs.duke.edu).

Xi He is a PhD student at the Computer Science Department of Duke University. Her research interests lie in privacy-preserving data analysis and security. She has an MS degree from Duke University and a double degree in Applied Mathematics and Computer Science from the University of Singapore. Xi has been working with Prof. Machanavajjhala on privacy since 2012. She has published in SIGMOD, VLDB, and CCS and has given tutorials on privacy at VLDB 2016 and SIGMOD 2017. She received best demo award on differential privacy at VLDB 2016 and was awarded a 2017 Google PhD Fellowship in Privacy and Security.



Vana Kalogeraki Athens University of Economics and Business, Athens, Greece (vana@aueb.gr).

Vana Kalogeraki is an Associate Professor at Athens University of Economics and Business leading the Distributed and Real-Time Systems research. Previously, she has held positions as an Associate and Assistant Professor at the Department of Computer Science at the University of California, Riverside, and as a Research Scientist at Hewlett-Packard Labs in Palo Alto, CA. She received her PhD from the University of California, Santa Barbara, in 2000. Prof. Vana Kalogeraki has been working in the field of distributed and real-time systems, participatory sensing systems, mobility, and crowdsourcing for over 20 years and has co-authored over 170 papers in journals and conferences proceedings, including co-authoring the OMG CORBA Dynamic Scheduling Standard. She was awarded a Marie Curie Fellowship; three best paper awards at ACM DEBS 2017, IEEE IPDPS 2009, SAINT 2008; two Best Student Paper Awards at PETRA 2016 and SAINT 2011; a UC Regents Fellowship Award, UC Academic Senate Research Awards; and a research award from HP Labs. Her research has been supported by an ERC Starting Independent Researcher Grant, the European Union, joint EU/Greek “Aristeia” grants, joint EU/Greek “Thalis” grant, NSF, and gifts from SUN and Nokia.



Frank Kargl Ulm University, 89069 Ulm, Germany (frank.kargl@uni-ulm.de).

Prof. Dr. Frank Kargl is the Director of the Institute of Distributed Systems at Ulm University, where a main focus of his research is on security and privacy protection in distributed systems in domains like automotive and industrial control systems. He holds a doctoral degree in Computer Science from the Ulm University and previously held a Professor position at the University of Twente.

Prof. Dr. Frank Kargl was involved in a number of national and international projects related to automotive security and privacy like SeVeCom, PRECIOSA, CONVERGE, and PRESERVE, that all worked toward a holistic security and privacy solution for connected cars and Intelligent Transport Systems (ITS). He has co-authored more than 100 peer-reviewed publications

in this area and actively contributed to standardization of security and privacy in ITS and beyond, through participation in bodies like C2C-CC or ETSI, and as co-chair of major conferences and workshops like ACM WiSec, ACM Vanet, IEEE WiVeC, and IEEE VNC.



Despina Kopanaki Department of Informatics, University of Piraeus, Athens, Greece (dkopanak@unipi.gr).

Despina Kopanaki is a PhD candidate in Informatics at the Department of Informatics, University of Piraeus, Greece, under the supervision of Professor Yannis Theodoridis. She got her bachelor degree from the Department of Statistics, Athens University of Economics and Business in Greece (2000–2004). She holds an MSc in Applied Economics and Finance from the Department of Economics, Athens University of Economics and Business (2004–2006). Her research interests include privacy preservation and data mining in moving object databases.



Ashwin Machanavajjhala Duke University, Durham, NC, USA (ashwin@cs.duke.edu).

Ashwin Machanavajjhala is an Assistant Professor in the Department of Computer Science, Duke University, and an Associate Director at the Information Initiative at Duke (iiD). Previously, he was a Senior Research Scientist in the Knowledge Management group at Yahoo! Research. His primary research interests lie in algorithms for ensuring privacy in statistical databases and augmented reality applications. His work on ℓ -Diversity that appeared in IEEE ICDE 2006 received the Influential Paper Award in 2017. He is also a recipient of the National Science Foundation Faculty Early CAREER Award in 2013, and the 2008 ACM SIGMOD Jim Gray Dissertation Award Honorable Mention. Ashwin graduated with a PhD from the Department of Computer Science, Cornell University, and a BTech in Computer Science and Engineering from the Indian Institute of Technology, Madras.



Anna Monreale Computer Science Department of University of Pisa, Pisa, Italy (annam@di.unipi.it).

Anna Monreale is a post-doc at the Computer Science Department of the Pisa University and a member of the KDD-LAB, a joint research group with the Information Science and Technology Institute of the National Research Council in Pisa. She received her MS and PhD degrees in Computer Science from the University of Pisa in 2007 and 2011, respectively. Her research is in privacy-aware data mining and data publishing, and in privacy-preserving outsourcing of analytical tasks.



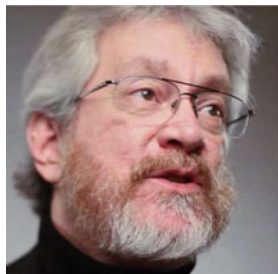
Dino Pedreschi Computer Science Department of University of Pisa, Pisa, Italy (pedre@di.unipi.it).

Dino Pedreschi is a Full Professor of Computer Science at the University of Pisa. His research interests are in data mining and in privacy-preserving data mining. He is a member of the program committee of the main international conferences on data mining and knowledge discovery. He has been granted a Google Research Award (2009) for his research on privacy-preserving data mining and anonymity-preserving data publishing.



Nikos Pelekis Department of Statistics & Insurance Science, University of Piraeus, Athens, Greece (pedre@di.unipi.it).

Nikos Pelekis is an Assistant Professor at the Department of Statistics and Insurance Science, University of Piraeus, Greece. His research interests include all topics of data science. He has been particularly working for almost 20 years in the field of Mobility Data Management and Mining. Nikos has co-authored one monograph and more than 60 refereed articles in scientific journals and conferences, receiving more than 700 citations, while he has received three best paper awards, won the SemEval'17 competition, and was ranked third in the ACM SIGSPATIAL'16 data challenge. He has offered several invited lectures in Greece and abroad (including PhD/MSc/summer courses at Rhodes, Milano, KAUST, Aalborg, Trento, Ghent, JRC Ispra) on mobility data management and data mining topics. He has been actively involved in more than ten European and National R&D projects. Among them he is or was principal researcher in GeoPKDD, MODAP, MOVE, DATASIM, SEEK, DART and datAcron, and Track & Know.



Alex (Sandy) Pentland MIT Media Lab, Cambridge, MA 02139, USA (sandy@media.mit.edu).

Professor Alex “Sandy” Pentland directs the MIT Connection Science and Human Dynamics labs and previously helped create and direct the MIT Media Lab and the Media Lab Asia in India. He is one of the most-cited scientists in the world, and *Forbes* recently declared him one of the “seven most powerful data scientists in the world,” along with Google founders and the Chief Technical Officer of the United States. He has received numerous awards and prizes such as the McKinsey Award from *Harvard Business Review*, the 40th Anniversary of the Internet from DARPA, and the Brandeis Award for work in privacy.



Jonathan Petit OnBoard Security, Wilmington, MA 01887, USA (jpetit@onboardsecurity.com).

Dr. Petit is the Senior Director of Research for Security Innovation, Inc. He is in charge of leading projects in security and privacy of automated and connected vehicles. He has conducted extensive research in detecting security vulnerabilities in automotive systems. He published the first work on potential cyber attacks on automated vehicles and remote attacks on automated vehicle sensors. He has supported communications security design and automotive cybersecurity analysis through OEM and NHTSA-sponsored projects. Previously, he was a Research Fellow in the Computer Security Group at University College Cork, Ireland. From 2011 to 2014, he was a Postdoctoral Researcher at the University of Twente, The Netherlands, where he co-coordinated the European-Union funded FP7 PRESERVE project. He received his PhD in 2011 from Paul Sabatier University, Toulouse, France.



Francesca Pratesi Information Science and Technology Institute of the National Research Council, Pisa, Italy (francesca.pratesi@isti.cnr.it).

Francesca Pratesi is a Postdoctoral Researcher at the Information Science and Technology Institute of the Italian National Research Council, and a member of the Knowledge Discovery and Data Mining Lab, a joint research group with the Italian National Research Council in Pisa. She received her MS degree and PhD degree in Computer Science from the University of Pisa.

in 2013 and 2017, respectively. Her research interests include data mining, data privacy, and spatiotemporal data analysis.



Delphine Reinhardt University of Göttingen, Göttingen, Germany (reinhardt@cs.uni-goettingen.de).

Delphine Reinhardt is a full professor for Computer Security and Privacy at University of Göttingen, Germany. Before 2018, she was an assistant professor at the University of Bonn and also associated to the Fraunhofer Institute for Communication, Information Processing and Ergonomics (FKIE). She completed her doctoral degree in computer science with distinction in 2013 at Technische Universität Darmstadt, Germany. Her doctoral thesis received awards by the Communication and Distributed Systems Group (KuVS) of the German Informatics Society (GI) and Information Technology Society of the Association for Electrical, Electronic & Information Technologies (VDE-ITG) as well as the “Vereinigung von Freunden der Technische Universität zu Darmstadt e.V.” for outstanding academic achievements. Since 2009, she holds a double-degree in electrical engineering from TU Darmstadt and Ecole Nationale Supérieure de l’Electronique et ses Applications (ENSEA), France. She has served as program committee member and reviewer for more than 60 international conferences and journals. Her research interests include privacy and security, trust and reputation, and usability in ubiquitous computing and beyond.



Florian Schaub University of Michigan, Ann Arbor, MI, USA (fschaub@umich.edu).

Florian Schaub is an Assistant Professor in the School of Information at the University of Michigan. His research focuses on empowering users to effectively manage their privacy in complex socio-technological systems. His research interests span privacy, human-computer interaction, mobile and ubiquitous computing, and the Internet of Things. Before joining UMSI, he was a Postdoctoral Fellow in the School of Computer Science at Carnegie Mellon University. He holds a doctoral degree and Diploma in Computer Science from the University of Ulm, Germany, and a Bachelor in Information Technology from Deakin University, Australia.



Cyrus Shahabi University of Southern California, Los Angeles, CA, USA (shahabi@usc.edu).

Cyrus Shahabi is a Professor of Computer Science, Electrical Engineering and Spatial Sciences at the University of Southern California (USC). He is also the Director of the NSF's Integrated Media Systems Center (IMSC), the Informatics Program and the Information Laboratory (InfoLAB) at USC's Viterbi School of Engineering. He received his BS in Computer Engineering from Sharif University in 1989 and then his MS and PhD degrees in Computer Science from USC in May 1993 and August 1996, respectively. He authored two books and more than 200 research papers in the areas of databases, GIS, and multimedia. Dr. Shahabi is a fellow of IEEE and a recipient of the ACM Distinguished Scientist Award in 2009, the 2003 US Presidential Early Career Awards for Scientists and Engineers (PECASE), and the NSF CAREER Award in 2002.



Andreas Solti Institute of Information System, Vienna University of Economics and Business, Vienna, Austria (solti@ai.wu.ac.at).

Andreas Solti is a Postdoctoral Researcher at the Institute for Information Business at the WU Vienna, Austria. He completed his PhD studies entitled "Probabilistic Estimation of Unobserved Process Events" in the business process technology group of the Hasso Plattner Institute of the University of Potsdam. Currently, he is working on the European Union FP7 project SERAMIS on topics related to sensor data analysis. Andreas has published more than 25 international research papers in journals and highly competitive conferences, including Information Systems, BPM, CAiSE, ICSOC, EDOC. He has served in several program committees of conferences, including BPI, BPIC, and EDOC.



Sarah Spiekermann-Hoff Institute for Management Information Systems, Vienna University of Economics and Business, Vienna, Austria (spiek@wu.ac.at).

Sarah chairs the Institute for Management Information Systems at the Vienna University of Economics and Business (WU Vienna). She has published more than 80 articles on the social and ethical implications of computer systems, and given more than a hundred presentations and talks about her work throughout the world. Her main expertise is electronic privacy, disclosure behavior, and ethical computing. Sarah has co-authored US/EU privacy regulation for RFID technology and has regularly worked as an expert and advisor to companies and governmental institutions, including the EU Commission and the OECD. She was on the advisory board of the Foundation of Data Protection of the German Parliament. She maintains a blog on *The Ethical Machine* at Austria's leading daily newspaper *Standard.at* and is on the board of the Austrian art and science think-tank GlobArt. Before being tenured in Vienna in 2009, Sarah was Assistant Professor at the Institute of Information Systems at Humboldt University Berlin (Germany), where she headed the Berlin Research Centre on Internet Economics (2003–2009).



Yannis Theodoridis Department of Informatics, University of Piraeus, Athens, Greece (ytheod@unipi.gr).

Professor Yannis Theodoridis is faculty member with the Department of Informatics and Director of the Data Science Lab, University of Piraeus, Greece. He serves or has served as member of the editorial boards of *ACM Computing Surveys* (2016–) and the *International Journal of Data Warehousing and Mining* (2005–), general co-chair for SSTD'03 and ECML/PKDD'11, PC vice-chair for IEEE ICDM'08, and PC member for numerous conferences in the fields of data management and data mining. His research interests include data science (big data management and analytics) for mobility information. He has co-authored three monographs and more than 100 refereed articles in scientific journals and conferences, with over 10,000 citations according to Google Scholar. He holds a Diploma (1990) and PhD (1996) in Computer Engineering, both from the National Technical University of Athens, Greece.



Hien To University of Southern California, Los Angeles, USA (hto@usc.edu).

Hien To is a Software Engineer at Amazon Mechanical Turk. He is a PhD candidate in Computer Science, University of Southern California since 2017. He received his Bachelor of Engineering degree in Computer Science from Hanoi University of Science and Technology in 2010. He worked as a Software Engineer at the sVNG Corporation from 2010 to 2011. Hien's research interest is on data privacy and crowdsourcing. He developed new privacy-preserving techniques for efficient task assignment in spatial crowdsourcing. He contributed to various projects funded by NSF, Google, Microsoft, Oracle, and Northrop Grumman. He authored more than ten research papers in top conferences and journals. Hien serves as the publicity Chair of GeoRich 2017 and local Chair of SIGSPATIAL 2018. He is a member of ACM and IEEE.



Jameson L. Toole Fritz Inc., Boston, MA. (jltoole@mit.edu).

Jameson Toole is the co-founder and CEO of Fritz, a company helping developers to optimize, deploy, and manage machine learning models on every mobile device and platform, and of Warehouse Inc., which aims to build efficient, robust, and useful technologies for mobile network operators. Previously, he ran the Data Engineering team at Jana Mobile and spent summers as a Software Engineering Intern at GoogleX, working with Project Wing to change the way we transport things with drones. Dr. Toole holds a PhD degree in Engineering Systems from the Massachusetts Institute of Technology, where he used data collected from mobile phones to understand urban mobility, social behavior, and economic outcomes. In addition to the PhD, he holds an MS and undergraduate degrees in physics, economics, and mathematics from the University of Michigan. He is a leading expert on using big data to understand transportation, human behavior, and mobility. Dr. Toole has authored multiple peer-reviewed papers and built teams of data scientists for companies like Google and Jana Mobile.

About the Editors



Aris Gkoulalas-Divanis IBM Watson Health Headquarters, Cambridge, MA 02142-1123, United States (gkoulala@us.ibm.com).

Aris Gkoulalas-Divanis is the Technical Lead on Data Protection and Privacy for IBM Watson Health. In this role, he is responsible for the design and integration of novel privacy-protection methods to the IBM Watson Platform for Health and the IBM Cloud, to support IBM customers' privacy and utility requirements. Aris received his PhD in Computer Science from the University of Thessaly in 2009. His PhD dissertation was awarded the Certificate of Recognition and Honorable Mention in the 2009 ACM SIGKDD Dissertation Awards. Aris has been a Postdoctoral Research Fellow in the Department of Biomedical Informatics of Vanderbilt University (2009–2010) and a Research Scientist in IBM Research-Zurich (2010–2012) and IBM Research-Ireland (2012–2016). His research interests are in the areas of databases, data mining, privacy-preserving data mining, privacy in trajectories and location-based services, privacy in medical data, and knowledge hiding. In these areas, he has given many seminars and two tutorials (ECML/PKDD 2011, SDM 2012), he has published more than 80 research works—including four Springer books, and he has applied for or being granted more than 25 patents (received three IBM Invention Achievement Awards). Aris is a regular reviewer in top-quality journals, as well as participates in the program committee of many prestigious conferences. He serves as an associate editor of the *Knowledge and Information Systems* (KAIS) journal, the *IEEE Trans-*

actions on Information Forensics and Security (T-IFS), the *International Journal of Research & Development Innovation Strategy (IJRDIS)*, and the *International Journal of Knowledge-Based Organizations (IJKBO)*. Since 2014 he is an area editor for *ACM Computing Reviews (CR)* covering the Information Systems category. Aris is a senior member of IEEE; a professional member of ACM, SIAM, and AAAS; and an at-large member of UPE and Sigma Xi.



Claudio Bettini EveryWare Lab, Department of Computer Science, University of Milan, Italy (claudio.bettini@unimi.it).

Claudio Bettini is Professor of Computer Science at the University of Milan, where he leads the EveryWare laboratory at the Computer Science Department. Claudio received his PhD in Computer Science from the University of Milan in 1993. He has been post-doc at IBM Research, NY, and, for more than a decade, an Affiliate Research Professor at the Center for Secure Information Systems at George Mason University, VA. His research interests cover the areas of mobile and pervasive computing, data privacy, temporal and spatial data management, and intelligent systems. He is the author of over 150 peer-reviewed publications on these topics. He acted as co-PI of several international projects on data privacy and contributed as project reviewer in the EU Horizon H2020 ERC and FET initiatives. Among his many organizational activities, he acted as General Chair of IEEE PerCom 2017 and IEEE MDM 2013 and as TPC Chair of IEEE PerCom 2013. He is a member of the steering committee of IEEE PerCom and Associate Editor of the *Pervasive and Mobile Computing* journal and previously Editor of *The VLDB Journal* and of the *IEEE Transactions on Knowledge and Data Engineering*. In 2011, he founded EveryWare Technologies, a startup developing innovative mobile apps and services for privacy and assistive technologies. He is a senior member of the IEEE Computer Society.

List of Figures

Fig. 2.1	Mobile phones are increasingly being used to collect high-resolution mobility data. This figure from de Montjoye et al. [15] depicts (a) a sequence of calling events made by a user at different locations. (b) These events are localized to the area served by the closest mobile phone tower to the use and (c) can be aggregated into individual specific neighborhoods where a user is likely to be found at different times of the day or week	16
Fig. 2.2	(a) Individual mobility trajectories are passively collected from mobile devices [23]. (b) Measuring the distribution of radius of gyrations, r_g within a population of 100,000 users in a European country reveals considerable heterogeneity in typical travel distance of individuals. Moreover, this distribution cannot be explained by modeling each individual’s movement as realizations of a single Levy flight process [23]. (c and d) Show the slower than linear growth in new locations visited over time $S(t)$ and that the probability a location is visited next is inversely proportional to the frequency it has been visited in the past [54]. (e) This preferential return contributes to strikingly high predictability $R(t)$ over time while (f) the number of unique locations visited in any given hour is highly periodic and corresponds to the sleep-wake cycles of individuals [55]	19

Fig. 2.3	(a) Removing geographic coordinates from locations and only focusing on a set of unique places and the directed travel between them, mobility motifs reveal that the daily routines of people are remarkably similar. Despite over 1 million unique ways to travel between 6 or fewer points, just 17 motifs are used by 90% of the population. Moreover, the frequency of their appearance in CDR data matches very closely with more traditional survey methods [52]. (b) Despite this similarity and predictability, our movement displays a high degree of unicity. Just four spatiotemporal points is enough to differentiate a user from 95% of all others individuals [15] 23	23
Fig. 2.4	(a) The radiation model accounts for intervening opportunities, producing more accurate estimates of flows between two places than more traditional gravity models [53]. (b) Routing millions of trips measured from CDR data to real road networks makes it possible to measure the importance of a road based on how many different locations contribute traffic to it, K_{road} . Understanding how transportation systems perform under different loads presents new opportunities to solve problems related to congestion and make infrastructure more efficient [63] 26	26
Fig. 2.5	(a) Global air travel has dramatically increased the speed at which diseases can spread from city to city and continent to continent [44]. (b) Mobility also adds context to social networks. When two individuals visit the same locations can suggest the nature of a social relationship [60]. (c) Mobility and the access it provides has strong correlations with economic outcomes. Children have dramatically different chances at upward economic mobility in certain places of the United States than others [10] 29	29
Fig. 3.1	Illustration for trilateration 41	41
Fig. 3.2	Illustration for 2-step fingerprinting. (a) and (b) depict the first step for training and (c) shows the second step for positioning 41	41
Fig. 3.3	An illustration of a shopping trip with RFID readers at a point of sales and an exit gate 49	49
Fig. 3.4	An illustration of a shopping trip with RFID readers at an interaction point, a POS and an exit gate 51	51
Fig. 3.5	An illustration of a shopping trip with RFID RTLS, a POS and an exit gate 52	52
Fig. 3.6	An illustration of a shopping trip with WiFi locating system 54	54

Fig. 3.7	An illustration of a shopping trip with RFID RTLS, Wi-Fi locating system, a POS and an exit gate	56
Fig. 3.8	Balancing act for legitimate interests as defined in the GDPR [81]	60
Fig. 5.1	DPT framework overview	111
Fig. 6.1	Crowdsensing architecture	131
Fig. 7.1	Threat models in spatial crowdsourcing. W and R denote workers and requesters, respectively. The dotted circles surrounding them denote that they are protected from an untrusted entity shown in the first column. After tasking and reporting, the assignment and reporting links between W and R represent the established connections during each phase. The dashed links indicate connections that are oblivious to the corresponding malicious entity. (a) Push mode. (b) Pull mode	172
Fig. 7.2	Screenshots of TaskRabbit web application from worker Bob. (a) Task locations. (b) Task price. (c) Task status. (d) Performed tasks	174
Fig. 7.3	A framework for privacy protection during tasking and reporting in the pull mode. Dashed entities are malicious, while others are trusted	177
Fig. 7.4	Examples of range dependency and all-inclusivity. (a) Query formation. (b) Range dependency leak. (c) All-inclusivity leak	181
Fig. 7.5	Distance estimation methods. (a) Centroid-point method. (b) Expected probabilistic method	183
Fig. 7.6	Differentially private framework for spatial crowdsourcing. (a) System architecture. (b) Worker PSD using adaptive grid	186
Fig. 8.1	A thematic map (top) and a reference map (bottom). The thematic maps shows the poverty rate by county in the U.S. in 2014 [123], whereas the reference map shows the U.S. territory by State, together with topographic, transportation and demographic information (images: (top) https://www.census.gov/did/www/saipe/data/statecounty/maps/iy2014/Tot_Pct_Poor2014.pdf , (bottom) https://upload.wikimedia.org/wikipedia/commons/7/7d/United_states_wall_2002_us.jpg , last retrieved Dec. 6, 2016)	205

Fig. 8.2	Online crime map for the region of Berkeley, California, during a 1-week period between Oct. 18–24th, 2016. The map shows the different types of reports, such as thefts, burglaries, assaults, vandalisms, and the place where they were reported by the Berkeley Police (image shown with permission from crimemapping.com , http://www.crimemapping.com/ , last retrieved Dec. 6, 2016)	206
Fig. 8.3	System architecture of a generic Location-Based Social Network. The links represent possible ways by which location data can be attached to the content posted by either people or organizations (adapted from [129]). The solid and dashed lines correspond to the actions that people and organizations can perform, respectively	211
Fig. 9.1	The connected car ecosystem	234
Fig. 9.2	Abstract pseudonym lifecycle	240
Fig. 9.3	Pay-as-You-Drive insurance models, once with the classical, privacy-invasive model (left) and once with the PriPAYD model that ensures no user data is inadvertently leaked (right)	244
Fig. 9.4	POPCORN protocol for privacy-preserving charging of electric vehicles	246
Fig. 10.1	(a) Milan GPS Trajectories, (b) characteristic points, (c) spatial clusters, (d) tessellation of the territory, and (e) generalized trajectories	260
Fig. 10.2	10 largest clusters of the original trajectories (top) and of the anonymized trajectories (down)	262
Fig. 10.3	F-measure for comparison of the clusterings of the anonymized dataset versus the clustering of the original trajectories	263
Fig. 10.4	The places taxonomy	265
Fig. 10.5	The empirical disclosure probability on Milano dataset	266
Fig. 10.6	(a) Number of patterns extracted from Milan data and (b) coverage of the patterns varying the support threshold	267
Fig. 10.7	CCDFs of <i>Flow per Link</i> (Left); CCDFs of <i>Flow per Zone</i> (Right)	272
Fig. 10.8	Visualization of <i>Flow per Link</i> (A-B) and <i>Flow per Zone</i> (C-D)	273
Fig. 11.1	A big picture of the system architecture [8]	285
Fig. 11.2	The architecture of Hermes++ [22]	287
Fig. 11.3	Generating a fake trajectory over a set of line segments [22]	288
Fig. 11.4	Sensitive location tracking and sequential tracking attacks to user privacy [22]	292
Fig. 11.5	Protecting sensitive locations of user trajectories [22]	293

Fig. 11.6	Selecting segments from real trajectories [22]	294
Fig. 11.7	Prohibiting sequential tracking: (a) case I, (b) case II [22]	296
Fig. 11.8	The result of a range query in its (a) original vs. (b) NWA anonymized version [23]	299
Fig. 11.9	T-Optics applied on (a) the original vs. (b) the anonymized dataset [23]	300
Fig. 11.10	Private-Hermes architecture [23]	301
Fig. 11.11	HipStream architecture [31]	303
Fig. 12.1	Never-Walk-Alone anonymization. Original dataset (city of Oldenburg in Germany) with 1000 trajectories (left) and its anonymized version (NWA from [2]) with $k = 3$ where the distance between any points of two trajectories within the same cluster is at most 2000 m (right) (image courtesy of Gábor György Gulyás)	320
Fig. 12.2	IRIS cells of Paris (left) and Voronoi-tesselation of tower cells (right)	324
Fig. 12.3	Algorithm 12.1 before improvements ($\varepsilon = 0.3$, $\delta = 2 \cdot 10^{-6}$, $\ell = 30$). (a) Large counts. (b) Small counts around local minimas	328
Fig. 12.4	Algorithm 12.1 after improvements ($\varepsilon = 0.3$, $\delta = 2 \cdot 10^{-6}$, $\ell = 30$). (a) Scaling. (b) Smoothing	329
Fig. 12.5	Mean relative error and Pearson correlation of each IRIS cell ($\varepsilon = 0.3$, $\delta = 2 \cdot 10^{-6}$, $\ell = 30$). (a) Naive Gaussian Perturbation (Avg. MRE: 1.01, PC: 0.47). (b) Algorithm 12.1 (Avg. MRE: 0.17, PC: 0.96)	330
Fig. 13.1	Conceptual view of a context-aware system based on the layer models by Baldauf et al. [13] and Knappmeyer et al. [74]	341
Fig. 13.2	Context-adaptive privacy mechanisms align context-aware system capabilities with the cognitive privacy regulation process in order to support privacy decision making and regulation	349
Fig. 14.1	The PCube app. (a) Login. (b) Add friends. (c) Set proximity	382
Fig. 14.2	The Locaccino app. (a) Map. (b) Privacy settings. (c) History	383
Fig. 14.3	The CrowdAlert app. (a) Login. (b) Report. (c) Settings	392

List of Tables

Table 3.1	Comparison of different technologies for location tracking [43]	44
Table 4.1	A classification of location-based services (SP = Service Provider)	78
Table 5.1	Table of notation	100
Table 7.1	Attacks on SC users	169
Table 7.2	Three tasks requested by requester Alice	175
Table 7.3	Overview of problem focuses (Re: reporting, Ta: tasking); privacy techniques used (Ps: pseudonym, Cl: cloaking, Pt: perturbation, Ex: exchange-based, En: encryption-based); threats (W: worker, T: requester, S: server); trusted third party (TTP); optimization type (ST: single task, MT: multiple tasks). x and (x) represent primary and secondary aspects, respectively	176
Table 8.1	Properties of different methods for geographic content generation	201
Table 8.2	Categorization of different research works according to the adversary type (In: internal, E: external), adversarial model (M: malicious, S: semi-honest, T: trusted third-party), the goal of the adversary (L: location, Ab: Absence, C: co-location, Id: identity, Ac: activity) and the proposed or suggested privacy protection mechanism (spatial/temporal cloaking, elimination, fake data, cryptography)	216

Table 12.1 Examples for k - and k^m -anonymity, where each row represents a record, public and sensitive attributes are not distinguished, and temporal information is omitted for simplicity 313

Table 13.1 Overview of context-adaptive privacy mechanisms 355

Table 14.1 Location-based applications, data sharing and privacy mechanisms 385

Table 14.2 University initiative location based app applications, data sharing and privacy mechanisms 386