12th International Workshop on Data Privacy Management

co-located with

Gamle museet
Dronningensgate 4 Kvadraturen,
0152 Oslo

Welcome message from the Program Chairs

On behalf of the Organizing Committee, we would like to welcome you to the 12th International Workshop on Data Privacy Management (DPM 2017), held in Oslo, Norway, the 14th and 15th of September of 2017, in conjunction with the 22nd European Symposium on Research in Computer Security (ESORICS) 2017. The DPM series started in 2005 when the first workshop took place in Tokyo (Japan). Since then, the event has been held in different venues: Atlanta, USA (2006); Istanbul, Turkey (2007); Saint Malo, France (2009); Athens, Greece (2010); Leuven, Belgium (2011); Pisa, Italy (2012); Egham, U.K. (2013); Wroclaw, Poland (2014); Vienna, Austria (2015); and Crete, Grece (2016).

The aim of DPM is to promote and stimulate the international collaboration and research exchange on areas related to the management of privacy-sensitive information. This is a very critical and important issue for organizations and end-users. It poses several challenging problems, such as translation of high-level business goals into system-level privacy policies, administration of sensitive identifiers, data integration and privacy engineering, among others.

For this workshop edition we received 51 submissions, and each one was evaluated on the basis of significance, novelty, and technical quality. The Program Committee, formed by 41 members, performed an excellent task and with the help of additional 18 referees all submissions went through a careful review process (three or more reviews per submission). In the end, 16 full papers were accepted for presentation at the event. The program is completed with a keynote talk given by Vicenç Torra (University of Skövde, Sweden) on *integral privacy*.

We would like to thank everyone who helped at organizing the event, including all the members of the Organizing Committee of both ESORICS and DPM 2017. Our gratitude goes to Pierangela Samarati, Steering Committee Chair of the ESORICS Symposium, for all her arrangements to make possible the satellite events, Socratis Katsikas, the Workshops Chair of ESORICS 2017. Last but, by no means least, we thank to all to the DPM 2017 Program Committee members, additional reviewers, all the authors who submitted papers, and all the workshop attendees. Finally, we want to acknowledge the support received from the sponsors of the workshop: Institut Mines-Telecom (Telecom SudParis), CNRS Samovar UMR 5157 (R3S team), Universitat Autonoma de Barcelona, UNESCO Chair in Data Privacy, Universitat Rovira i Virgili, and project TIN2014-55243-P from the Spanish MINECO

September 2017

Joaquin Garcia-Alfaro
Guillermo Navarro-Arribas

Keynote Speaker

Dr. Vicenç Torra

Professor, PhD, University of Skövde, Sweden

Title: *Privacy models and disclosure risk: integral privacy*

Abstract

The literature offers a set of privacy models which can be used to specify what kind of inference are we interested to avoid. Reidentification, k-anonymity and differential privacy are well known examples of these privacy models. Most of them can be classified according to the two dimensions: attribute and identity disclosure. In this talk, we will give an overview of the privacy models and describe some of our work on evaluating the worst-case scenario for some masking methods, and discuss the privacy model of integral privacy.

Bio

Vicenç Torra is a Full Professor at the University of Skövde (Sweden) since 2014. Before, he was Assoc. Prof. Research Track (1999-2014) at the Artificial Intelligence Research Institute (IIIA-CSIC) and Assoc. Prof. (1994-1999) at the Universitat Rovira i Virgili, and Visiting researcher at the University of Tsukuba (Japan). He is an IEEE and EurAI Fellow, and ISI Elected member.

His research interests include: data privacy (statistical disclosure control and privacy preserving data mining), approximate reasoning, aggregation operators and data fusion, decision making and game theory. He has published in major journals and conferences and has written six books. Among the latter, he has published "Data Privacy" (Springer, 2017), "Scala: From a functional programming perspective" (Springer, 2016), and "Modeling Decisions" (with Y. Narukawa, Springer, 2007). His research has been funded by national and international projects (e.g. EU ones). He is currently working on a project on Big data privacy (VR 2016-03346, 2017-2020). He is regularly involved in editorial boards and program committees. He is the founder and Editor-in-chief of the Transactions on Data Privacy (www.tdp.cat).

Thursday, September 14, 2017

09:00 – 10:30

Room: Big Time
General Welcome & Keynote Talk
Keynote Title: *Privacy models and disclosure risk: integral privacy*
Speaker: Dr. Vicenç Torra

11:00 – 12:30

Session 1: Privacy, logics, and computational models

Room: Big Time
Chair: Joaquin Garcia-Alfaro (Telecom SudParis and Paris-Saclay University)

A Proof Calculus for Attack Trees in Isabelle

Florian Kammueller (Middlesex University London and TU Berlin)

Attack trees are an important modeling formalism to identify and quantify attacks on security and privacy. They are very useful as a tool to understand step by step the ways through a system graph that lead to the violation of security policies. In this paper, we present how attacks can be refined based on the violation of a policy. To that end we provide a formal definition of attack trees in Isabelle's Higher Order Logic: a proof calculus that defines how to refine sequences of attack steps into a valid attack. We use a notion of Kripke semantics as formal foundation that then allows to express attack goals using branching time temporal logic CTL. We illustrate the use of the mechanized Isabelle framework on the example of a privacy attack to an IoT healthcare system.

Confidentiality of Interactions in Concurrent Object-Oriented Systems

Olaf Owe (University of Oslo), Toktam Ramezanifarkhani (University of Oslo)

We consider a general concurrency model for distributed systems, based on concurrent objects communicating by asynchronous methods. This model is suitable for modeling of service-oriented systems, and gives rise to efficient interaction avoiding active waiting and low-level synchronization primitives such as explicit signaling lock operations. This concurrency model allows us to focus on information flow at a high level of abstraction, Our approach is formalized by a high-level language with a secrecy-type system ensuring noninterference in object interactions. We prove soundness based on an operational semantics incorporating runtime secrecy levels.

Using Oblivious RAM in Genomic Studies

Nikolaos Karvelas (TU Darmstadt), Andreas Peter (University of Twente), Stefan Katzenbeisser (TUD)

Since the development of tree-based Oblivious RAMs by Shi et al. it has become apparent that privacy preserving outsourced storage can be practical. Although most current constructions follow a client-server model, in many applications, such as Genome Wide Association Studies (GWAS), it is desirable that multiple entities can share data, while being able to hide access patterns not only from the server, but also from any other entities that can access parts of the data. Inspired by the efficiency and simplicity of Path-ORAM, in this work, we study an extension of Path-ORAM that allows oblivious sharing of data in a multi-client setting, so that accesses can be hidden from the server and from other clients. We address various challenges that emerge when using Path-ORAM in a multi-client setting, and prove that with adequate changes, Path-ORAM is still secure in a setting, where the clients are semi-honest, do not trust each other, but try to learn the access patterns of each other. We demonstrate our ORAM construction in a GWAS setting. Our experiments show that in databases storing 2^{23} data blocks (corresponding to a database holding 2^{17} blocks per client, capable of storing human genome in the form of SNPs, for 100 clients), the average query time is less than 7 seconds, yielding a secure and practical solution.

14:00 – 15:30

Session 2: Privacy and encrypted search

Room: Big Time

Chair: Guillermo Navarro-Arribas (Universitat Autonoma de Barcelona)

Towards Efficient and Secure Encrypted Databases: Extending Message-Locked Encryption in Three-Party Model

Yuuji Furuta (Osaka University), Naoto Yanai (Osaka University), Masashi Karasaki (Nippon Telegraph and Telephone West Corporation), Katsuhiko Eguchi (Nippon Telegraph and Telephone West Corporation), Yasunori Ishihara (Osaka University), Toru Fujiwara (Osaka University)

In database systems with three parties consisting of a data owner, a database manager and a data analyst, the data owner uploads encrypted data to a database and the data analyst delegated by the data owner analyzes the data by accessing to the database without knowing plaintexts. In this work, towards an efficient and secure scheme whose encryption can be processed in real time, we extend message-locked encryption (Bellare et al., EUROCRYPT 2013), where parts of ciphertexts are generated from their plaintexts deterministically. In particular, we introduce both delegations of relational search between ciphertexts from a data owner to a data analyst, and re-encryption of ciphertexts such that ciphertexts of the message-locked encryption become truly probabilistic against a database manager. We call the scheme \textit{message-locked encryption with re-encryption and relational search}, and formalize the security, which is feasible and practical, in two cases, i.e., any relationship in a general setting and only an equality test in a restricted setting. Both settings are useful from a standpoint of trade-offs between the security and the efficiency. We also propose an instantiation with the equality test between ciphertexts.

Searchable Encrypted Relational Databases: Risks and Countermeasures

Mohamed Ahmed Abdelraheem, Tobias Andersson (SICS), Christian Gehrmann (Lund University)

We point out the risks of providing security to relational databases via searchable symmetric encryption (SSE) schemes by mounting an inference attack exploiting the properties of relational databases and the leakage of searchable encryption schemes. We show that record-injection attacks mounted on relational databases have worse consequences than their file-injection counterparts on unstructured databases. Moreover, we discuss some techniques to reduce the effectiveness of inference attacks exploiting the access pattern leakage existing in SSE schemes.

Private verification of access on medical data: an initial study

Thais Bardini Idalino (University of Ottawa), Dayana Spagnuelo (University of Luxembourg), Jean Everson Martina (Universidade Federal de Santa Catarina)

Patient-centered medical systems promote empowerment of patients, who can decide on the accesses and usage of their personal data. To inspire a sense of trust and encourage the adoption of such systems, it is desired to allow one to verify whether the system has acted in accordance with the patients' preferences. However, it is argued that even audit logs and usage policies, normally used when verifying such property, may already be enough for one to learn sensitive information, e.g., the medical specialists a given patient has visited in the past. This is not only damaging for the patients, but is also against the interests of the medical system, which may lose back the trust earned and gain a bad reputation. Verifiability should not come at the expense of patients' privacy. It is, therefore, imperative that these systems take necessary precautions towards patient's information when providing means for verifiability. In this work we study how to realize that. In particular, we explore how searchable encryption techniques could be applied to allow the verification of systems in a private fashion, providing no information on patient's sensitive data.

16:00 – 17:30
Session 3: Data privacy, data mining, and applications

Room: Big Time

Chair: Florian Kammueller (Middlesex University London and TU Berlin)

Default Privacy Setting Prediction by Grouping User's Attributes and Settings Preferences

Toru Nakamura (KDDI Research, Inc.), Welderufael Berhane Tesfay (Goethe University Frankfurt), Shinsaku Kiyomoto (KDDI Research, Inc.), Jetzabel Serna (Goethe University Frankfurt)

While user-centric privacy settings are important to protect the privacy of users, often users have difficulty changing the default ones. This is partly due to lack of awareness and partly attributed to the tediousness and complexities involved in understanding and changing privacy settings. In previous works, we proposed a mechanism for helping users set their default privacy settings at the time of registration to Internet services, by providing personalised privacy-by-default settings. This paper evolves and evaluates our privacy setting prediction engine, by taking into consideration users' settings preferences and personal attributes (e.g. gender, age, and type of mobile phone). Results show that while models built on users' privacy preferences have improved the accuracy of our scheme; grouping users by attributes does not make an impact in the accuracy. As a result, services potentially using our prediction engine, could minimize the collection of user attributes and based the prediction only on users' privacy preferences.

δ-privacy: Bounding Privacy Leaks in Privacy Preserving Data Mining

Zhizhou Li (The Voleon Group), Ten H. Lai (The Ohio State University)

We propose a new definition for privacy, called δ-privacy, for privacy preserving data mining. The intuition of this work is, after obtaining a result from a data mining method, an adversary has better ability in discovering data providers' privacy; if this improvement is large, the method, which generated the response, is not privacy considerate. *δ*-privacy requires that no adversary could improve more than *δ*. This definition can be used to assess the risk of privacy leak in any data mining methods, in particular, we show its relations to differential privacy and data anonymity, the two major evaluation methods. We also provide a quantitative analysis on the tradeoff between privacy and utility; rigorously prove that the information gains of any *δ*-private methods do not exceed *δ*. Under the framework of *δ*-privacy, it is able to design a pricing mechanism for privacy-utility trading system, which is one of our major future works.

Threshold Single Password Authentication

Devriş Işler (Koç University), Alptekin Küpçü (Koç University)

Passwords are the most widely used form of online user authentication. In a traditional setup, the user, who has a human-memorable low entropy password, wants to authenticate with a login server. Unfortunately, existing solutions in this setting are either nonportable or insecure against many attacks, including phishing, man-in-the-middle, honeypot, and offline dictionary attacks. Three previous studies (Acar et al. 2013, Bicakci et al. 2011, and Jarecki et al. 2016) provide solutions secure against offline dictionary attacks by additionally employing a storage provider (either a cloud storage or a mobile device for portability). These works provide solutions where offline dictionary attacks are impossible as long as the adversary does not corrupt both the login server and the storage provider. For the first time, improving these previous works, we provide a more secure generalized solution employing multiple storage providers, where our solution is proven secure against offline dictionary attacks as long as the adversary does not corrupt the login server and threshold-many storage providers. We dene ideal and real world indistinguishability for threshold single password authentication (Threshold SPA) schemes, and formally prove security of our solution via ideal-real simulation. Our solution provides security against all the above-mentioned attacks, including phishing, man-in-the-middle, honeypot, and offline dictionary attacks, and requires no change on the server side. Thus, our solution can immediately be deployed via a browser extension (or a mobile application) and support from some storage providers. We further argue that our protocol is efficient and scalable, and provide performance numbers where the user and storage load are only a few milliseconds.

Towards A Toolkit for Utility and Privacy-Preserving Transformation of Semi-structured Data Using Data Pseudonymization

Saffija Kasem-Madani (University of Bonn), Michael Meier (University of Bonn), Martin Wehner (University of Bonn)

We present a flexibly configurable toolkit for the automatic pseudonymization of datasets that keep certain utility. We define some possible utility requirements and corresponding utility options a pseudonym can meet. Based on that, we define a policy language that can be used to produce machine-readable utility policies. The utility policies are used to configure the toolkit to produce a pseudonymized dataset that offers the utility options. Here, we follow a confidentiality-by-default principle. I.e., only the data mentioned in the policy is transformed and included in the pseudonymized dataset. All remaining data is kept confidential. This stays in contrast to common pseudonymization techniques that replace only personal or sensitive data of a dataset with pseudonyms, while keeping any other information in plaintext.

Friday, September 15, 2017

09:00 – 10:30

Session 4: User Privacy

Room: Big Time

Chair: Jordi Herrera-Joancomartí (Universitat Autonoma de Barcelona)

Privacy Dashcam - Towards Lawful Use of Dashcams Through Enforcement of External Anonymization

Paul Wagner (Karlsruhe Institute of Technology), Pascal Birnstill (Fraunhofer IOSB), Erik Krempel (Fraunhofer IOSB), Sebastian Bretthauer (Karlsruhe Institute of Technology), Jürgen Beyerer (Fraunhofer IOSB)

Dashcams are small, dashboard mounted camera systems that continuously monitor the area around a vehicle and record video images on a portable storage device. According to many data protection authorities, dashcams constitute surveillance systems that are oper- ated by private individuals in public places. By continuously acquiring personal data they interfere disproportionately with the right of informational self-determination. One approach to make dashcams compliant to data protection law is to automatically identify personal information – at least pedestrian's faces and license plates – in the captured video image and subsequently disguise them. Even though appropriate anonymization methods exist, high computational costs prevent their use in portable dashcams. This article presents a new approach that enforces the anonymization of encrypted dashcam videos on a dedicated computer system, before the user gets access to the videos. To accomplish this, classified images are safeguarded by usage control techniques on the way from the camera to the anonymization component. By ap- plying the developed system, any existing dashcam can ultimately be enhanced by privacy protection capabilities.

DLoc: Distributed Auditing for Data Location Compliance in Cloud

Mojtaba Eskandari (University of Trento), Bruno Crispo (University of Trento), Anderson Santana De Oliveira (SAP)

The prevalence of mobile devices and their capability to access high speed Internet has transformed them into a portable pocket cloud interface. In order to protect user's privacy, the European Union Data Protection regulations restrict the transfer of European users' personal data within the geographical boundaries of the European Union itself. The matter of concern, however, is the enforcement of such regulations. Since cloud service provision is independent of physical location and data can travel to various servers, it is a challenging task to determine the location of data and enforce jurisdiction policies. In this paper we introduce a framework,

named DLoc, which enables the end-users to track the location of their data after being transferred to the cloud. DLoc does not require a network of monitoring servers (landmarks) and does not need to reside and run within the target server. It uses a proof of data possession technique to guarantee that the cloud storage service possess the particular file and estimates its location(s) in a distributed manner without requiring the collaboration of the data controller or cloud provider. Empirical evaluations demonstrate that DLoc provides a better accuracy than its rival approaches in real world scenarios.

Inonymous: Anonymous Invitation-Based System

Sanaz Taheri Boshrooyeh (Koç University), Alptekin Küpçü (Koç University)

In invitation-based systems, a user is allowed to join upon receipt of a certain number of invitations from the existing members. The system administrator approves the new membership if he authenticates the inviters and the invitations, knowing who is invited by whom. However, the inviter-invitee relationship is privacy-sensitive information and can be exploited for inference attacks: The invitee's profile (e.g., political view or location) might leak through the inviters' profiles. To cope with this problem, we propose Inonymous, an anonymous invitation-based system where the administrator and the existing members do not know who is invited by whom. We formally define and prove the inviter anonymity and the unforgeability of invitations. Inonymous is efficiently scalable in the sense that once a user joins the system, he can immediately act as an inviter, without re-keying and imposing overhead on the existing members. We also present InonymouX , an anonymous cross-network invitation-based system where users join one network (e.g., Twitter) using invitations of members of another network (e.g., Facebook).

11:00 – 12:40

Session 5: Applied cryptography and privacy

Room: Big Time

Chair: Cristina Pérez-Solà (Universitat Autonoma de Barcelona)

PCS, a privacy-preserving certification scheme

Nesrine Kaaniche, Maryline Laurent, Pierre-Olivier Rocher, Christophe Kiennert (Telecom SudParis), Joaquin Garcia-Alfaro (Telecom SudParis)

We present *PCS*, a privacy-preserving certification mechanism that allows users to conduct anonymous and unlinkable actions. The mechanism is built over an attribute-based signature construction. The proposal is proved secure against forgery and anonymity attacks. A use case on the integration of PCS to enhance the privacy of learners of an e-assessment environment is presented. The goal is to allow the learners of an e-assessment platform to reveal only required information to certificate authority providers.

Order-Preserving Encryption Using Approximate Integer Common Divisors

James Dyer (University of Manchester), Martin Dyer and Jie Xu (University of Leeds)

We present a new, but simple, randomized order-preserving encryption (OPE) scheme based on the general approximate common divisor problem (GACDP). This scheme requires only *O(1)* arithmetic operations for encryption and decryption. We show that the scheme has optimal information leakage under the assumption of uniformly distributed plaintexts, and we indicate that this property extends to some non-uniform distributions. We report on an extensive evaluation of our algorithms. The results clearly demonstrate highly favorable execution times in comparison with existing OPE schemes.

Privacy-Preserving Deterministic Automata Evaluation with Encrypted Data Blocks

Giovanni Di Crescenzo (Vencore Labs), Brian Coan (Vencore Labs), Jonathan Kirsch (Vencore Labs)

Secure computation (i.e., performing computation while keeping inputs private) is a fundamental problem in cryptography. In this paper, we present an efficient and secure 2-party computation protocol for deterministic automata evaluation, a problem of large practical relevance. Our result is secure under standard assumptions and bypasses roadblocks in previous general solutions, like Yao's garbled circuits and Gentry's lattice-based fully homomorphic encryption, by performing secure computations over data blocks (instead of bits) and using typical-size (instead of impractically large) cryptographic keys and ciphertexts. An important efficiency property achieved is that the number of both asymmetric and symmetric cryptographic operations in the protocol is *sublinear* in the size of the circuit representing the computed function (specifically, improving linear-complexity protocols by a multiplicative factor equal to a block size). All previous protocols for deterministic automata evaluation required a linear number of asymmetric cryptographic operations. Even though not as general as in the two mentioned techniques, our result is applicable to the class of all constant-space computations.