Marco Gaboardi and Chris J. Skinner

Special issue on the theory and practice of differential privacy

Article (Published version) (Refereed)

Journal of Privacy and Confidentiality (2016) **7**, Number 2, 1–2

Special Issue on the Theory and Practice of Differential Privacy

Marco Gaboardi* and Chris J. Skinner,† *Guest Editors*

This special issue presents papers based on contributions to the first international workshop on the "Theory and Practice of Differential Privacy" (TPDP) held in London, UK, 18 April 2015, as part of the European joint conference on Theory And Practice of Software (ETAPS).

Differential privacy is a mathematically rigorous definition of the privacy protection provided by a data release mechanism: it offers a strong guaranteed bound on what can be learned about a user as a result of participating in a differentially private data analysis. Researchers in differential privacy come from several areas of computer science, including algorithms, programming languages, security, databases and machine learning, as well as from several areas of statistics and data analysis. The workshop was intended to be an occasion for researchers from these different research areas to discuss the recent developments in the theory and practice of differential privacy.

The program of the workshop included 10 contributed talks, 1 invited speaker and 1 joint invited speaker with the workshop "Hot Issues in Security Principles and Trust" (HotSpot 2016). Participants at the workshop were invited to submit papers to this special issue. Six papers were accepted, most of which directly reflect talks presented at the workshop.

In this special issue

Thomas Steinke and Jonathan Ullman study lower bounds on the sample complexity of differentially private algorithms for answering one-way marginal queries. The authors study two problems: the gap between the best possible error under pure and approximate differential privacy, and the gap between the best possible average and best possible worst-case error (over queries). By using known techniques in a novel way, this work fills important gaps in the literature such as showing that the difference between worst-case and average error in approximate differential privacy is sub-logarithmic. Moreover, the results of Steinke and Ullman also imply similar results for other kind of analysis such as private empirical risk minimization and private principal component analysis.

Fragkiskos Koufogiannis, Shuo Han, and George J. Pappas study the problem of releasing private data under differential privacy when the privacy parameter ϵ can be increased over time. The scenario that they consider is the one where a data curator privately releases some data with a privacy level ϵ, and then decide to increase this level to ϵ'. Their result shows that the data curator can release an answer that is

*University at Buffalo, SUNY gaboardi@buffalo.edu.
†London School of Economics C.J.Skinner@lse.ac.uk.

ϵ'-differentially private and exactly as accurate as if the data curator had decided to achieve ϵ'-differential privacy to start with. That is, there is no accuracy loss incurred by the data curator by having changed her mind after the fact. This is the first paper addressing this interesting problem and the results by Fragkiskos Koufogiannis, Shuo Han, and George J. Pappas are encouraging and can open new scenarios in differential privacy.

Marco Gaboardi, Emilio Jesús Gallego Arias, Justin Hsu, Aaron Roth, and Zhiwei Steven Wu propose an algorithm for releasing answers to collections of queries over high dimensional relational data under differential privacy. While this algorithm runs in worst-case exponential time, it has the benefit that the computationally expensive task can be formulated as a combinatorial optimization problem that does not need to be solved privately and can be handled with an off-the-shelf solver. The authors prove rigorous privacy guarantees and show experimentally that the algorithm's runtime scales favorably with the number of queries and the dimension of the data. This paper combines in a compelling way an interesting theoretical idea with an experimental analysis.

Michael Backes, Aniket Kate, Praveen Manoharan, Sebastian Meiser, and Esfandiar Mohammadi present a formal framework to analyze anonymity guarantees for anonymous communication (AC) protocols. This framework is based on an abstract property defined by combining (in)distinguishability games that are typical in computational security, with a multiplicative bound inspired by differential privacy. This framework allows a great amount of flexibility which is useful to analyze different notions of anonymity like sender anonymity, receiver anonymity, sender unlikability, and relationship anonymity. This paper shows how the bound described by differential privacy can have applications that go beyond the traditional ones.

Mohammad Alaggan, Sébastien Gambs and Anne-Marie Kermarrec consider an extension of differential privacy that allows different users to have different levels of privacy. They name this privacy definition "heterogeneous differential privacy". Similar definitions have been considered before in the literature showing the interest for a more personalized notion of differential privacy. The main contribution of their work is the "stretching mechanism", inspired to the Laplace mechanism, which is used to guarantee heterogeneous differential privacy for numeric valued outputs. The authors provide a privacy and an utility analysis of this new mechanism, and some experiments showing its usefulness.

Hamid Ebadi and David Sands conclude the special issue with a paper which is atypical for the Journal of Privacy and Confidentiality. They contribute to the area of programming languages for differential privacy. More specifically this contribution presents a formal model for the language PINQ, a library for differentially private data manipulation. The authors model a simplified version of PINQ, "Featherweight PINQ" using a technique from formal logic named probabilistic transition system. This formal model permits formal guarantees about the differential privacy of PINQ programs that the original language PINQ did not provide. This work shows how differential privacy can benefit from work in related areas of computer science.