PRIMO

PRIvate data MOnetization: a public good approach using cooperative game theory

Michela CHESSA

Host: EURECOM, Networking and Security department (sup.: Patrick Loiseau)

1. Context and motivation

Personal data, i.e., data generated by or about an individual has intrinsic economic value. Indeed, many companies derive profit from user data acquired through online tracking: Google and Facebook generate profits from targeted advertising, Amazon via recommendations pointing users towards items likely to be purchased, and so on. The common theme across diverse services and providers is the same; collect data about the end-user, mine it to generate user profiles, and monetize the profiles. In this prevailing ecosystem, users are locked out of extracting any monetary value from their own data. They are also negatively affected by privacy implications, as they implicitly forsake control of their data and how it is used. In recent years, an important debate has emerged about these practices, their drawbacks, and the concerns about protecting users from being exploited.

In this context, **allowing users to gain control over their personal data is a key challenge.** In an optimal ecosystem, companies should still have the possibility of collecting users data. However, users who experience a cost for their loss of privacy should play a central role in deciding whether or not they want to release their data, and at which price. Such **an ecosystem with monetization of users data can benefit all parties**: it allows users to derive monetary value from their data in compensation for their loss of privacy and it allows companies to continue using users data for profit. (We note here that the value for users need not be monetary. It could be, e.g., services, but still needs to be quantified.)

To ensure the success of a system giving users the possibility of monetizing their personal data, **it is crucial to design a good market mechanism for determining the value of this data**. Clearly, the value of personal data has a private component related to how much a particular user values his privacy. However, it also has a public component due to the possibility of inferring data about a user from data about others. Indeed, users are naturally embedded in networks such as social networks. When links between users are known, data released by a user can allow statistical inference on data of other users that are connected to him and therefore affect their privacy. The public component of personal data has an important impact on the optimal privacy market mechanism (intuitively, for example, optimal strategies may involve offering incentives to certain users who have a central position in the underlying network). In this project, **we propose an innovative approach seeing privacy as a public good to design an optimal market mechanism for personal data**.

2. Detailed scientific content: objectives, innovation and expected results

The ultimate goal of this project is to design an optimal pricing mechanism for private data monetization that gives each party the appropriate incentives. To achieve this goal, we will use models based on game theory that capture the objective of each agent of the system. A main innovation of the project is the combination of public goods and social networks together with cooperative game theory solutions. This association is essential to obtain a realistic model and therefore derive relevant results for the pricing mechanism that can later be used in a prototype implementation of personal data market.
We identify several intermediate objectives:

i. *Definition of a main usecase and proposition of a good game-theoretic model associated*. To be realistic, the model will include the social network environment in which users are naturally embedded and the public good component of personal data.
ii. *Analysis of the model using cooperative game theory*. We will look in particular for structural properties of the equilibrium solutions and their link to the network structure.
iii. *Application to the design of an optimal pricing mechanism for private data monetization*.

In Objective i, we will propose a good model guided by a practical usecase. We will start by defining the utility function of each player, depending on his decision on selling data and on the decision of other players connected to him though the graph. The network structure will be represented by a weighted graph with weights representing the "inference power". This model of data inference on the graph is very innovative as it captures the public good nature of personal data. After describing the individual utility functions, we will define the aggregate utility of all the players, depending on who is deciding to sell his data. In every possible configuration (i.e., subset of players who decide to sell their data, whatever the quantity), we define the optimal quantity to be sold by each player in order to maximize the aggregate utility.

The main theoretical contribution is expected in Objective ii. The aggregate utility of each configuration, evaluated when the players sell the optimal quantity will be interpreted as the characteristic function of a cooperative game, where the players are the users. The properties of this game will give the optimal cooperation strategy, i.e., for which group of users it is convenient to sell their data. We will investigate key properties such as superadditivity and convexity, and we will investigate the implementation of some known solutions for cooperative games. The core [5], the Shapley value [12] and in particular the Myerson value [11], will be investigated to solve the problem of sharing the total income between the users who decided to sell their data. The Myerson value is particularly interesting as it is the main solution when dealing with a cooperative game where the connections of the players are represented through a graph. We expect to obtain general results on the structure of solutions (stability, optimality) for a class of cooperative games on graphs and on the their link with the graph structure.

These theoretical results will be the key towards Objective iii: to design the appropriate pricing mechanisms that will control incentives for users to sell their data or not. Through our cooperative game-theoretic analysis, we will demonstrate theoretically the feasibility of designing a market for personal data and establish guidelines for pricing this data. The creation of a market for personal data, which is fair and efficient in the interest of users and application providers, can have a very high importance, given the growing concern about user privacy.

The main challenge of the project will be to be able to analyze the solutions of the innovative realistic cooperative game model including the public good nature of personal data. Solving this challenge will permit to design a pricing mechanism with appropriate incentives that ensures wide adoption.

3. History and related work

Academic concerns about the economic value of privacy date back to the mid-90's, with informal discussions about the value of private data by [8] and [14]. More recently, a few experiments have been conducted to quantify the value that users assign to their private data [1, 6].

The game-theoretic analysis of the private data monetization was pioneered by [7], who proposed fair compensation mechanisms for private data. Their solutions are based on the core and on the Shapley value, but their simplistic game model does not take into account

the real utility function appearing when considering privacy as a public good. Recently, a significant thread of research started on selling differential privacy using auctions [3,4,9]. In [15], the authors propose a mechanism called “transactional privacy” where users can sell access to their data through an unlimited supply auction.

None of the works summarized above consider the possibility of inference of data about a user from other users data. In this project, this important aspect is included via our innovative vision of privacy as a public good. Our work will find inspirations from the recent literature on public goods [13] and on pricing in social networks [2]. These works, however, analyze very different contexts and do not use techniques from cooperative game theory.

4. Local context and interest for the labex UCN@Sophia

The *motto* “the user at the heart of the network” perfectly represents our project. In particular our mechanism will provide a technical solution to reduce the lack of control of the user on its own private data and to increase the user confidence in the Internet. By enforcing usage control of data by its owner, the mechanism will meet the user’s privacy and security requirements in order to prevent unnecessary and undesirable data collection. This proposal naturally pertains to the research theme “Security, Privacy and Network Neutrality”, in particular as a reply to the challenge of designing a solution that protects privacy and allows ownership of published user data.

As a post-doc researcher I will be hosted in the Networking and Security department of EURECOM, partner of the Labex UCN@Sophia. I will work with Prof. Patrick Loiseau. He is an expert in network economics and game theory and he has already used public goods successfully for his work on congestion management [10]. My knowledge on cooperative games together with his qualifications will make possible to model and to develop this innovative approach to solve the question of monetization of private data.

Bibliography

[1] What is privacy worth?, A. Acquisti, L. John, and G. Loewenstein, Working paper, 2009.
[2] Optimal pricing in networks with externalities, O. Candogan, K. Bimpikis, A. Ozdaglar, Operations Research, 60(4):883-905, 2012.
[3] Privacy auctions for recommender systems, P. Dandekar, N. Fawaz, and S. Ioannidis, In Proceedings of WINE, pages 309–322, 2012.
[4] Selling privacy at auction, A. Ghosh and A. Roth, In Proceedings of ACM EC, pages 199–208, 2011.
[5] Solutions to general non-zero-sum games, D. B. Gillies, In A. W. Tucker, R. D. Luce (Eds.) Contributions to the Theory of Games IV, Annals of Mathematics Studies 40, Princeton: Princeton University Press, pages 47-85, 1959.
[6] Valuating privacy, B. A. Huberman, E. Adar, and L. R. Fine, IEEE Security & Privacy, 3(5):22–25, 2005.
[7] On the value of private information, J. Kleinberg, C. H. Papadimitriou, and P. Raghavan, In Proceedings of TARK, pages 249–257, 2001.
[8] Markets and privacy, K. C. Laudon, Communications of the ACM, 39(9):92-104, 1996.
[9] Take it or leave it: Running a survey when privacy comes at a cost, K. Ligett and A. Roth, In Proceedings of WINE, pages 378–391, 2012.
[10] Incentive Mechanisms for Internet Congestion Management: Fixed-Budget Rebate versus Time-of-Day Pricing, P. Loiseau, G. Schwartz, J. Musacchio, S. Amin, and S. Shankar Sastry. IEEE/ACM Transactions on Networking, 2013 (to appear).
[11] Graphs and cooperation in Games, R. Myerson, Mathematics of Operations Research, 2:225-229, 1977.

[12] A value for n-Perosn Games, L. S. Shapley, In H. W. Kuhn, A. W. Tucker (Eds.) Contributions to the Theory of Games, Annals of Mathematics Studies 28, Princeton: Princeton University Press, pages 307-317.
[13] Local Public Good Provisioning in Networks: A Nash Implementation Mechanism, IEEE Journal on selected areas in communications, 30(11):2105-2116, 2012.
[14] Economic aspects of personal privacy, H. R. Varian, In W. H. Lehr and L. M. Pupillo, editors, Internet Policy and Economics, Challenges and Perspectives, pages 101–109. Springer, 2009. Published version of a note from 1996.
[15] For sale : your data: by : you, C. Riederer, V. Erramilli, A. Chaintreau, B. Krishnamurthy, and P. Rodriguez, In Proceedings of HotNets, pages 13:1–13:6, 2011.