Privacy Protection of User Profiles in Personalized Information Systems

Author: Javier Parra-Arnau, Thesis Advisors: J. Forné and D. Rebollo-Monedero
contact email: xparnau@gmail.com

I. Introduction

Recent years have witnessed the accelerated growth of a rich variety of personalized information systems (PISs) of unprecedented sophistication, which have been integrating seamlessly into our daily lives. Examples of these systems comprise personalized Web search and news, resource tagging in the semantic Web and multimedia recommendation systems. The key enabling technology of such systems is *personalization*, a research area that has received great attention lately and whose aim is to tailor information-exchange functionality to the specific interests of their users. To accomplish this functionality, most personalized information systems capitalize on, or lend themselves to, the construction of *profiles*, either directly declared by a user, or inferred from past activity, not only of the user in question, but also from the profiles of users with whom social relationships are known to the information system.

Personalized services therefore allow users to deal with the overwhelming overabundance of information, but inevitably at the expense of *privacy*, especially when profiling is conducted across several information systems. Besides, the enrichment of these services with data from social networks creates additional opportunities with respect to information sharing but, at the same time, increases the user privacy risks. Figure 1 shows an example of user profile modeled as a list of categories of interest.

II. Measuring the Privacy of User Profiles

A variety of privacy-enhancing technologies (PETs) have been proposed to enable the provision of new services and functionalities aimed at mitigating those privacy threats. Unfortunately, these technologies have not yet gained wide adoption. This is because it remains unclear whether their overall benefits outweigh their typically costly deployment and/or integration, as well as the operational cost that arises due to the fact that PETs typically come with penalties in terms of utility and performance, when compared to more privacy-invasive alternatives [1]. Assessing the privacy provided by a PET is, therefore, crucial to both determine its overall benefit and compare its effectiveness with other technologies. In other words, privacy metrics, accompanied with utility metrics, provide a quantitative means of contrasting the suitability of two or more privacy-enhancing mechanisms.

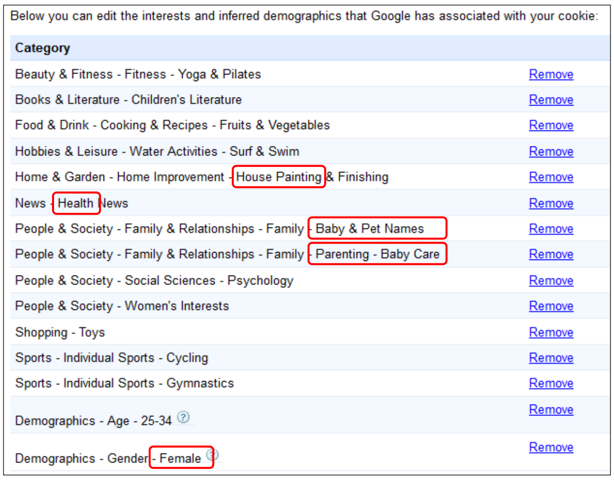
Building upon well-established principles of information theory and statistics, we make a first contribution in this direction by proposing Kullback-Leibler (KL) divergence as a criterion for quantifying the privacy of user profiles. Our metric, which encompasses Shannon's entropy as a special case, is examined, on the one hand, under the beautiful perspective of the method of types and large deviation theory, and on the other, under Jaynes' rationale behind entropy-maximization methods. The proposed privacy measure contemplates a user profile modeled as a normalized histogram of user data, e.g., tags, ratings or queries, across a predefined set of categories of interest. In addition, we consider two distinct adversary models–an attacker aimed at targeting users who deviate from the average profile of interests; and another attacker whose objective is to classify a given user into a group of users.

III. Privacy-Enhancing Technologies in Personalized Information Systems

Equipped with a quantitative measure of privacy and utility, we investigate PETs providing *hard privacy*. By hard privacy, the privacy research literature refers to the case in which users mistrust communicating entities, e.g., the personalized information provider or the network operator, and thus strive to reveal as little private information as possible. This is in contrast to those privacy-preserving systems that build upon the assumptions of *soft privacy*, what means that users entrust their private data to these systems, which are therefore responsible for the protection of their data.

Under the assumptions of hard privacy, this thesis contemplates two conceptually-simple strategies that capitalize on the principle of data perturbation. First, we consider the *suppression* of tags in the scenario of the semantic Web, and secondly, the combination of the *forgery* and suppression of ratings in personalized recommendation systems. Figure 2 provides a depiction of one of these approaches. Specifically, we illustrate the case of tag suppression, whereby users may wish to refrain from tagging certain resources. In doing so, the actual user profile q, that is, the profile capturing the user genuine interests, is observed from the outside as a perturbed profile; we refer to this profile as the apparent user profile s. Consequently, the adoption of our approach enables users to avoid being accurately profiled by the service provider, or in general, by any attacker capable of collecting the tags posted by users.

Our second strategy contemplates the submission of false information, together with the aforementioned suppression technique, but in the scenario of



Below you can edit the interests and inferred demographics that Google has associated with your cookie:

Category	
Beauty & Fitness - Fitness - Yoga & Pilates	Remove
Books & Literature - Children's Literature	Remove
Food & Drink - Cooking & Recipes - Fruits & Vegetables	Remove
Hobbies & Leisure - Water Activities - Surf & Swim	Remove
Home & Garden - Home Improvement - House Painting & Finishing	Remove
News - Health News	Remove
People & Society - Family & Relationships - Family - Baby & Pet Names	Remove
People & Society - Family & Relationships - Family - Parenting - Baby Care	Remove
People & Society - Social Sciences - Psychology	Remove
People & Society - Women's Interests	Remove
Shopping - Toys	Remove
Sports - Individual Sports - Cycling	Remove
Sports - Individual Sports - Gymnastics	Remove
Demographics - Age - 25-34 ?	Remove
Demographics - Gender - Female ?	Remove

Figure 1. Example of user profile, as shown by Google [2]. The interest of this user in the categories highlighted in red might reveal she is pregnant or planning to get pregnant. If this information ended up in the hands of her employer, her job might be at risk.

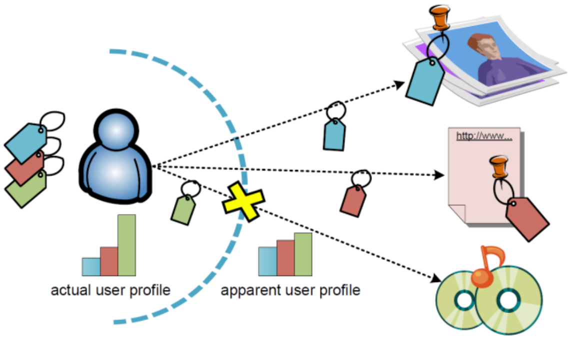


Figure 2. Tag suppression in the semantic Web.

recommendation systems. More precisely, in our approach, users rate items, e.g., movies, music or books, as they normally do. However, when their privacy is being compromise, users may want to submit some ratings to items that do not reflect their actual interests.

IV. On the Trade-Off between Privacy and Utility

By adopting our strategies, users enhance their privacy to a certain extent, without having to trust an external entity or the network operator. Nevertheless, this is inevitably at the expense of a loss in data utility. For example, in the case of tag suppression, privacy comes at the cost of a degradation in the semantic functionality of the Web, since tags has the purpose of associating meaning with resources. On the other hand, the forgery and suppression of ratings in recommendation systems come with penalties in terms of the accuracy of the prediction generated by the recommender. In a nutshell, data-perturbative mechanisms pose an inherent *trade-off* between privacy and utility.

One of the objectives of this thesis is precisely to investigate the trade-off posed by such PETs. For this purpose, first we formulate mathematically the compromise between these two contrasting aspects; and secondly we tackle the issue in a systematic fashion by applying the methodology of multiobjective optimization. Our extensive theoretical analysis includes a close-form solution to the mathematical problem of tag suppression on the one hand, and to the problem of the forgery and suppression of ratings on the other. In addition, we characterize the optimal trade-off between the aspects of privacy and utility. Figure 3 illustrates the trade-off between privacy, measured as the Shannon entropy of the apparent user profile $\mathrm{H}(s)$, and the tag suppression rate $\sigma \in [0,1)$, i.e., the proportion of tags a user is willing to eliminate. Figure 4 shows the contours of the function modeling the trade-off among privacy risk, forgery rate ρ and suppression rate σ.

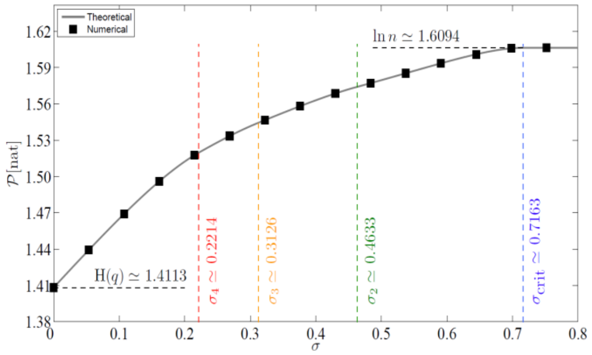


Figure 3. Privacy-utility trade-off in tag suppression.

V. Experimental Evaluation of our Privacy-Enhancing Technologies

Having investigated the privacy-utility trade-off posed by such PETs, we study the impact of those mechanisms on a real world application scenario. In particular, we assess the level of privacy attained by those users suppressing tags, and also how this mechanism may affect a parental control filter that enforces blocking conditions on resources (e.g., Web pages, videos or pictures), on the basis of the tags associated with them. More accurately, we contemplate an enhanced collaborative tagging system that consists of a “traditional” bookmarking service, such as Delicious (http://delicious.com), and two main additional services built on top of it. Such services address two main issues. The former allows users to specify certain policies to control the access to the browsed data, and the latter features our tag suppression mechanism.

Our experimental evaluation shows how our PET allows users to enhance their privacy to a certain extent. In addition, we assess the impact that suppression has on utility, by considering the percentage of tags that each bookmark loses as a result of the elimination of tags. Lastly, we quantitatively evaluate the degradation in the classification of Web content, in terms of false negatives, false positives, precision and recall. Our results indicate that our technique does not have a significant impact on the accuracy of a parental control filter.

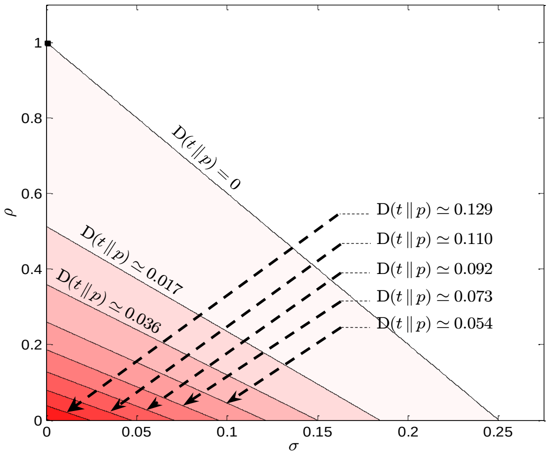


Figure 4. We measure privacy *risk* $\mathcal{R}$ as the KL divergence between the apparent user profile t, resulting from the addition of false ratings and the suppression of genuine ratings, and the population’s distribution of ratings p, that is, $\mathcal{R} = \mathrm{D}(t\|p)$. This figure plots the contours of the privacy risk function for different values of forgery rate ρ and suppression rate σ.

VI. Acknowledgments

This work was partly supported by the Spanish Government through projects Consolider Ingenio 2010 CSD2007-00004 “ARES", TEC2010-20572-C02-02 “Consequence".

VII. References

[1] J. Borking, “Why adopting privacy enhancing technologies takes so much time, in: S. Gutwirth, Y. Poullet, P. Hert, R. Leenes (Eds.), Proc. Comput. Priv., Data Prot. (CPD), Springer-Verlag, 2011, pp. 309-341.

[2] Google Ads Preferences. Available at http://www.google.com/ads/preferences.