Recommender Systems for Privacy Management: A Framework

Curtis Rasmussen
School of Computer Science
University of Guelph,
Guelph, Canada
crasmuss@uoguelph.ca

Rozita Dara, Member, IEEE
School of Computer Science
University of Guelph,
Guelph, Canada
drozita@uoguelph.ca

***Abstract*— Social media and online service providers are increasingly collecting personal information. In order for users to make decisions about their online privacy, they will have to read through a dense and hard-to-understand privacy policy. We developed a recommender system to help users make more pertinent decisions with regards to their privacy by providing them with recommendations and warnings based on their privacy preferences. Our ultimate goal is to build intelligent recommender systems that can process any combination of user data and privacy policies to provide recommendations for privacy management to the user.**

Keywords: privacy; privacy statement; ontology; knowledge base; recommender systems; decision making

I. INTRODUCTION

As the use of social media websites increases, as well as the widespread movement to include more personalization options in online services, the amount of information about a person that is not online in some form is dwindling. Almost all free web services exist to collect personal data, and this collection of information has turned into a lucrative business, largely in the form of targeted advertisements. This trend in data collection has raised serious privacy concerns among users and customers of such services as they no longer are only the subjects, but also are the producers of such data. Collection and usage of personal data by the service provider is one issue, the business model for misuse of personal data, sharing data with secondary and third parties is another growing trend which is deeply concerning.

Many studies and surveys have suggested that privacy policies are complex, vague, hard to read, and usually do not touch on issues that users care about [1,2]. It is also hard for an average user to understand technical or legal terminologies in the privacy statements. There have been attempts to create a tool to help users understand privacy policies and make decisions based on their personal preferences ([3] – [5]).

Recommender systems are tools that predict users' preferences for an item based on user behavioral data [9]. We use a similar concept in which we utilize user behavioral data and preferences for privacy managements, privacy policies, privacy statements, and several machine-readable knowledge bases [10] to recommend best action to the users or warn them about the consequences. More specifically, as a proof-of-concept, we have focused on social media and have implemented a recommender tool for such applications.

II. PROPOSED APPROACH

Our Privacy Recommender system uses ontology for analyzing the privacy policies of social media websites. By using ontology to model the terms and concepts found in human-readable privacy policies, our goal has been to develop a framework that is generalizable across different applications and will be able to empower users by producing meaningful recommendations. In addition, using ontology enables us to reuse the various constructs we develop for analyzing any privacy policy even in applications that do not require recommendations.

A. Privacy Management Recommender Systems Framework

Our proposed framework is depicted in Figure 1. This framework is generic and can be applied for different applications and domains:

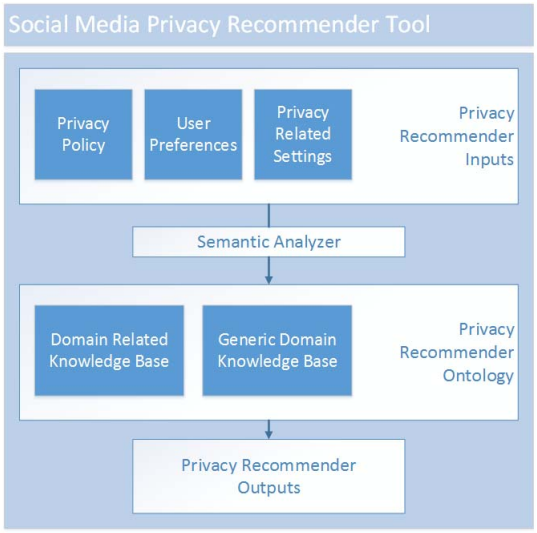

Figure 1. The generic framework

- System Input: input to the system could be collected from the user, tools' privacy policies and settings, as well as the historical data collected in the system databases.
- Semantic Analyzer: Textual inputs are then inputted to the system's semantic analyzer, which is a system capable of transforming them into terminologies that are understandable by the knowledge bases.

978-1-4799-3466-9/14 $31.00 © 2014 IEEE
DOI 10.1109/HASE.2014.43

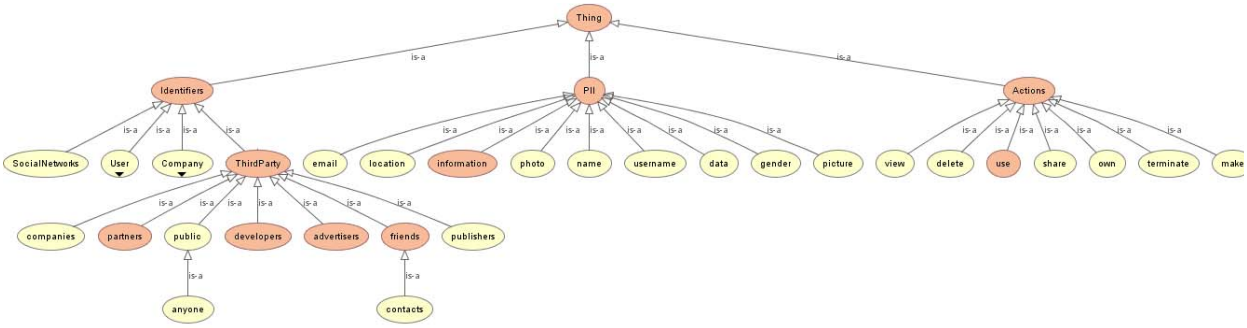

Figure 3. Sample ontology tree for privacy statement extraction

- Ontology Engine: The ontology engine consists of several knowledge bases: the Domain Related Knowledge Base (DRKB), and the Generic Domain Knowledge Base (GDKB). The GDKB holds generalized information regarding collection, use, and disclosure of personal data. On the other hand, the DRKB contains knowledge about specific application/domain (i.e. in our case social media).
- System Output: the output of the system may vary depending on the option that users select. User may wish to only extract relevant statements from long privacy policies or may decide to receive recommendations and warnings, along with the corresponding privacy policy statement.

B. Sample Results

This recommender system has been implemented using the Java programming language. We have also used OWL tool Protégé to construct the ontologies [23].

The user can either type in his request as free-text or can select one from a list of predefined options we have provided. Users requested will be parsed by the semantic analyzer and inputted to the ontology engine. The GDKB processes the terminologies provided by the semantic analyzer and extracts relevant statements from privacy policies or settings pages.

Take a user who is concerned about the third parties his information will be shared with. Ideally, this information would be presented to him plainly: a simple list of whom his information will be shared with. Our tool is able to give the user that kind of information. Given this statement from Facebook's privacy policy: "We use the information we receive about you in connection with the services and features we provide to you and other users like your friends, our partners, the advertisers that purchase ads on the site, and the developers that build the games, applications, and websites you use."

In Figure 3, these terms, as well as their parents, are highlighted in red. This information is then presented to the user as follows: "Your information may be shared with these third-parties: friends, partners, advertisers, developers." This user can now clearly see with whom his information will be shared, without having it obfuscated. By examining Figure 3 we can see how this statement is constructed: the system sees that some *personal information* is being used by third parties. It can deduce that this is what the user wants to know, given he is interested in who will have access to his information.

III. CONCLUSION

In this paper, we presented an implementation of a framework for building privacy management recommender systems. The framework operates based on three categories of data that are readily available to the user: privacy policies, user preferences, and privacy setting instructions. The recommender system is capable of warning and providing feedback to end-users. Although we are in early stages of this project, we were able to show that our proposed framework and ontology rules are able to provide context-aware recommendations. Our goal is that, with automatic comprehension of human-readable privacy policies combined with other data attributes, we can provide end-users with a practical and user-friendly tool to manage privacy.

REFERENCES

[1] J. Mullin, "New privacy bill requires apps to disclose how they share personal data," Ars Technica, 2013, http://arstechnica.com/tech-policy/2013/05/new-privacy-bill-requires-apps-to-disclose-how-they-share-personal-data/ (Accessed 22 July 2013).

[2] C. Jensen and C. Potts, "Privacy policies as decision-making tools: an evaluation of online privacy notices," in *Proc. of the Int. Conf. on Human Factors in Computing Systems*, NY, USA, 2004, pp. 471–478.

[3] L. Fang, H. Kim, K. LeFevre, and A. Tami, "A privacy recommendation wizard for users of social networking sites," in *Proc. of the 17th Conf. on Computer and communications security*, USA, 2010, pp. 630–632.

[4] K. Bernsmed, I. A. Tondel, and A. A. Nyre, "Design and Implementation of a CBR-based privacy agent," in *Proc. 7th Int. Conf. on Availability, Reliability and Security (ARES)*, 2012, pp. 317–326.

[5] K. Ghazinour, S. Matwin, and M. Sokolova, "Monitoring and recommending privacy settings in social networks," in *Proc. of the Joint EDBT/ICDT Workshops*, NY, USA, 2013, pp. 164–168.

[6] J. Debattista, S. Scerri, I. Rivera, and S. Handschuh, "Ontology-based rules for recommender systems," in SeRSy, 2012, pp. 49-60.

[7] M. Uschold, M. Gruninger, "Ontologies: principles, methods and applications," *Knowledge engineering review*, vol. 11, no. 2, pp. 93–136, 1996.

[8] The Protégé ontology editor and knowledge acquisition system, *Protégé*, 2013, http://protege.stanford.edu (Accessed 22 July 2013).