Fine-Grained Privacy Setting Prediction Using a Privacy Attitude Questionnaire and Machine Learning

Frederic Raber, Felix Kosmalla, Antonio Krueger

Fine-grained Privacy Setting Prediction using a Privacy Attitude Questionnaire and Machine Learning

Frederic Raber, Felix Kosmalla and Antonio Krueger

DFKI, Saarland Informatics Campus

Abstract. This paper proposes to recommend privacy settings to users of social networks (SNs) depending on the topic of the post. Based on the answers to a specifically designed questionnaire, machine learning is utilized to inform a user privacy model. The model then provides, for each post, an individual recommendation to which groups of other SN users the post in question should be disclosed. We conducted a pre-study to find out which friend groups typically exist and which topics are discussed. We explain the concept of the machine learning approach, and demonstrate in a validation study that the generated privacy recommendations are precise and perceived as highly plausible by SN users.

1 Introduction

The tradeoff between privacy and utility in a social network (SN) has been a research problem from the beginning, since SNs are largely used in public. Still, there is no acceptable solution that provides an optimal tradeoff between privacy and utility while keeping the user burden at a minimum. Social network providers tried to tackle this problem by introducing *friend* lists or *circles*. Users create one or more lists containing a subset of their online friends, and publish a new post exactly to the people inside these lists. Still, the SN users have the burden of manually setting the appropriate privacy setting for each of these groups in order to achieve a perfect privacy setting. Recent studies have shown that only 17% of all posted content is shared using friend lists [5].

We argue that every single post needs its own privacy setting, and should only be disclosed to a specific list of users, depending on the *topic* of the post. To decrease the user burden, the privacy settings should be derived automatically, for example by using a machine learning approach. Although most social networks like Facebook or Google+ only allow a binary decision on the privacy settings (e.g. to disclose or not), we think that a user decision on privacy is a decision that is not ultimately binary. A SN user does not only think "I do not at all want my drinking buddies to know that I am a ballet dancer as a hobby" or "I would really like my co-dancers to see the pictures of that ballet contest". There are also some groups of people, like university friends, where a user would say "It is OK if they see it. I do not want to completely cut them off from that information, but I also do not want to draw too much attention to it". In this

case, the user would take some middle road, for example by sharing the post and the pictures with the university friends, but hiding them from their timelines.

2 Related work

Several publications in the past have offered questionnaires to capture privacy attitudes. Starting with Westin scales [3] as a very general form of questionnaire, newer questionnaires like the IUIPC [4] provide a very specific privacy attitude regarding privacy towards online companies. Wisniewski et al. [10] created a privacy scale to observe how social connectedness corresponds with a user's privacy desires on a social network, which we also included in our questionnaire.

There are also other systems that use machine learning for the prediction of privacy settings, for example by labeling some of the friends with privacy permissions and using a supervised learning approach [6,7,2]. Other approaches additionally take the post content into account, by using latent Dirichlet allocation (LDA) and maximum entropy to predict settings for a new post based on the privacy settings chosen in earlier posts[8]. Although the idea seems promising, research has shown that privacy behavior in online social networks does not correspond to actual privacy desires; this is known as the privacy paradox [1]. We therefore decided to capture the privacy attitude using a distinct privacy questionnaire rather than trying to extract it from the user's SN behavior. Furthermore, all approaches so far rely on a binary decision (disclose/undisclose) for a privacy setting, whereas our approach offers five distinct privacy levels.

3 Approach

In a final implementation of our approach, the post topic is extracted and shown on the left side in Figure 1, while the proposed privacy settings for a selection of friend groups are displayed on the right side. As stated in the introduction, the proposed privacy settings are not only disclose/undisclose, but five different privacy levels as follows: On level 1, everything is disclosed and shown on the wall. Level 2 means the content does not appear on the recipients' news wall, whereas level 3 completely hides comments and graphical content. Level 4 hides the entire post, and level 5 also hides it from the recipient's direct friends, so it cannot be propagated to him by word of mouth. What exactly is hidden, is denoted by the small pictograms next to each friend group.

For suggesting the permissions, we use a machine learning technique called ridge regression. As input features, we use the measures calculated from the answers to the aforementioned two questionnaires [4,10] and the topic of the post, or only the questionnaire answers (called "generic" in Table 1). As an output, we receive for every friend group a privacy level between 1 and 5.

We performed three user studies to first find out which topics are most frequently discussed in people's social activities (online and offline) and which friend groups exist; second, to gather training data for the machine learning algorithm and to validate its precision; and third, to validate the approach in a scenario as

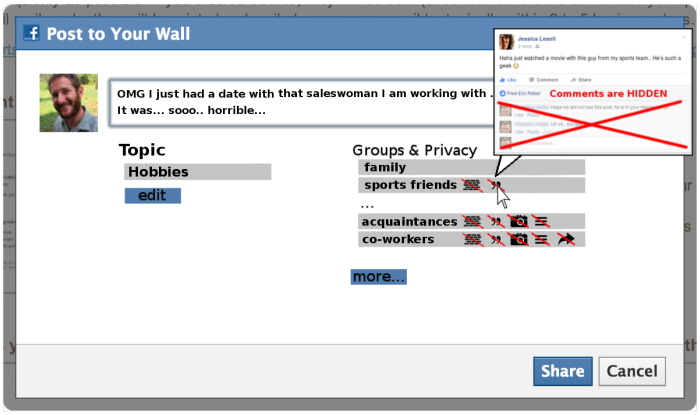

Fig. 1. Envisioned user interface concept of a privacy setting prediction system.

realistic as possible, introducing the proposed settings of our machine learning prediction to Facebook users. All studies were performed using online questionnaires; participants were recruited using *prolific academic*, an online recruiting portal similar to Amazon Mechanical Turk.

For the first, we asked 15 participants to list their friend groups and most frequently discussed topics in their social life in a free-text form. We merged the answers using an axial coding approach [9]. The most frequent topics were (in descending order) family affairs, events, movies, politics, food, work, hobbies, travel, music and sports. The friend groups that were mentioned most frequently were extended and immediate family, work friends, close friends, acquaintances and school/university as well as online and sports friends.

In the second study ("main study"), we let 100 participants first answer the two aforementioned questionnaires, followed by a matrix where they had to enter a privacy level for each topic/friend group pair. We trained and validated the regression with a ten-fold cross validation. The mean squared error (MSE) between the prediction and the actual result can be found in Table 1.

For the third study, called the "validation study", we again let 31 persons fill out the two privacy questionnaires in the first part. But this time, we let them copy and paste ten of their own Facebook posts that match our list of topics, and enter the topic of the post into the questionnaire. The website then proposed a privacy setting, using the ridge regression trained with the data of the former study. The participants were asked to adapt the settings if needed, and answer on a five-point Likert scale whether they would use the system on Facebook. Again we calculated the mean squared error between the adapted and the proposed settings. 67% of the participants stated that they would likely or very likely use our system, supporting the design of our approach. The results in Table 1 show that the trends are similar for both studies: For almost all topics, we can achieve a mean squared error less than one. Hobbies, travel and family are predicted best, whereas sports and politics are hardest to predict; maybe

	Main study	Validation study	
topic	*mean squared error*	*mean squared error*	*# posts*
family	0.87	0.93	38
events	0.91	0.85	24
movies	0.91	0.26	23
politics	1.05	0.91	14
food	0.88	0.46	28
work	1.00	1.17	18
hobbies	0.83	0.86	29
travel	0.88	0.64	17
sports	1.31	0.6	22
generic	0.96	0.78	230

Table 1. Amount of posts and mean squared error for the selected topics with machine learning in the main and the validation study.

because of the diverse nature of sports, where the exact sport affects whether it is likely to be shared or not. Posts about football are more common and socially accepted than posts about ballet, for example. Politics and work are also hard to predict by privacy attitude; this could be caused by the fact that here, the political interest or the job itself affects whether you want to share your thoughts, rather than a pure privacy attitude. A professor is more likely to share his work with a community than a cleaner would be.

4 Lessons learned and future work

We did background research to find friend groups and topics that are prsent in people's online and offline social life, and conducted two studies to find out whether it is possible to propose fine-grained privacy settings based on privacy attitude and the topic of the post. We tested and evaluated in two different scenarios: In the main study, users had no proposed setting, and had to enter their desired setting without support. In contrast to this, they had a proposed setting they had to adapt in the validation study. In both cases, we achieved an acceptable precision for most of the topics. Nevertheless, there are some topics like work and politics that seem not to depend on the privacy attitude, but rather on the actual occupation or political interest of the person.

Instead of a binary decision, our approach supports five privacy levels of disclosure, which offer to show only parts of the post to some friend groups, such as only the textual content without images or comments. For this study, we used an example implementation of the privacy levels. In future work, we would like to conduct further studies to determine which parts of the post users would hide depending on the post's sensitivity, and how an optimal implementation of the levels looks. Finally, we would like to offer a prototype of the proposed interface as a Facebook plugin, to be able to check whether the achieved prediction precision is sufficient for everyday use, and whether the tool is accepted by users.

References

1. Barnes, S.B.: A privacy paradox: Social networking in the united states. First Monday 11(9) (2006), http://firstmonday.org/ojs/index.php/fm/article/view/1394
2. Fang, L., LeFevre, K.: Privacy wizards for social networking sites. In: Proceedings of the 19th International Conference on World Wide Web. pp. 351–360. WWW '10, ACM, New York, NY, USA (2010), http://doi.acm.org/10.1145/1772690.1772727
3. Kumaraguru, P., Cranor, L.F.: Privacy Indexes: A Survey of Westin's Studies. Tech. Rep. CMU-ISRI-5-138, Institute for Software Research International, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA (December 2005)
4. Malhotra, N.K., Kim, S.S., Agarwal, J.: Internet users' information privacy concerns (iuipc): The construct, the scale, and a causal model. Info. Sys. Research 15(4), 336–355 (Dec 2004), http://dx.doi.org/10.1287/isre.1040.0032
5. Mondal, M., Liu, Y., Viswanath, B., Gummadi, K.P., Mislove, A.: Understanding and specifying social access control lists. In: Symposium On Usable Privacy and Security (SOUPS 2014). pp. 271–283. USENIX Association, Menlo Park, CA (Jul 2014), https://www.usenix.org/conference/soups2014/proceedings/presentation/mondal
6. Shehab, M., Cheek, G., Touati, H., Squicciarini, A., Cheng, P.C.: User centric policy management in online social networks. In: Policies for Distributed Systems and Networks (POLICY), 2010 IEEE International Symposium on. pp. 9–13 (July 2010)
7. Shehab, M., Touati, H.: Semi-supervised policy recommendation for online social networks. In: Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012). pp. 360–367. ASONAM '12, IEEE Computer Society, Washington, DC, USA (2012), http://dx.doi.org/10.1109/ASONAM.2012.66
8. Sinha, A., Li, Y., Bauer, L.: What you want is not what you get: Predicting sharing policies for text-based content on facebook. In: Proceedings of the 2013 ACM Workshop on Artificial Intelligence and Security. pp. 13–24. AISec '13, ACM, New York, NY, USA (2013), http://doi.acm.org/10.1145/2517312.2517317
9. Strauss, A., Corbin, J.M.: Basics of Qualitative Research : Techniques and Procedures for Developing Grounded Theory. SAGE Publications (September 1998), http://www.amazon.co.uk/exec/obidos/ASIN/0803959400/citeulike-21
10. Wisniewski, P., Islam, A.N., Knijnenburg, B.P., Patil, S.: Give social network users the privacy they want. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work #38; Social Computing. pp. 1427–1441. CSCW '15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2675133.2675256