On Protecting Private Information in Social Networks: A Proposal

Bo Luo [#], Dongwon Lee [*]

[#] *Department of EECS, The University of Kansas, Lawrence, KS, USA*
[*] *College of IST, The Pennsylvania State University, University Park, PA, USA*
[#] `bluo@ku.edu` [*] `dongwon@psu.edu`

***Abstract*— As online social networks get more popular, it becomes increasingly critical to preserve user privacy in such networks. In this paper, we propose our preliminary results on defining and tackling information aggregation attacks over online social networks. We first introduce three major threats towards private information in online social networks. We conceptually model private information into multilevel and discretionary models. Then, we articulate information aggregation attacks under discretionary model. Finally, we present our preliminary design of "privacy monitor," a framework that allows users to define their own privacy scheme, and track their actual privacy disclosure to check for any unwanted leakage.**

I. Introduction

Recently, online social network has emerged as a promising area with many products and a huge number of users. With the development of information retrieval and search engine techniques, it becomes very convenient to extract users' personal information that is readily available in various social networks. Malicious or curious users take advantage of these techniques to collect others' private information. Therefore, it is critical to enable users to control their information disclosure and effectively maintain privacy over online social networks.

While many people work on extracting information or learning knowledge from social networks, little research has been devoted to security and privacy in such networks. Existing research on web security and privacy falls short since most of them prevent people from giving out information. They are not suitable for social networks, where users intend to share information and socialize. In online social networks, in particular, user privacy is under three types of threats.

Threat 1: Out-of-context information disclosure.

Users give out information to social network community, assuming that community is trusted at some level. For instance, people trust *LinkedIn* as a professional/business network, so that they build profiles with educational background and employment history. They assume that their information would stay in the network. Unfortunately, this assumption is not always valid. In many cases, the information could be easily accessed from outside of the context due to wrong configuration, mal-functioning code, or user's misunderstanding. There are many real world examples: messages sent to an email-based social network may be archived at a repository and accessible to the open public; stalkers follow people through social networks [6], gadgets and add-ons may access users' profiles [12], etc.

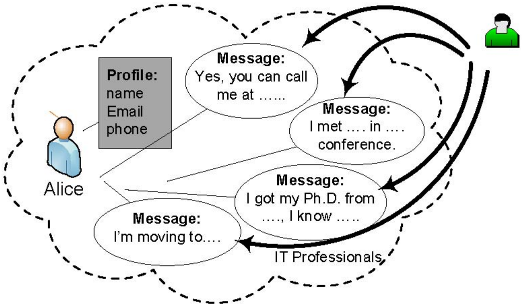

Fig. 1. In-network information aggregation attack.

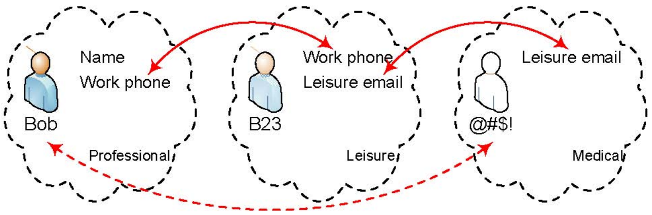

Fig. 2. Re-identification through cross-network information aggregation attack.

Threat 2: In-network information aggregation.

Suppose Alice is a user of an online social network for IT professionals. She has carefully made her profile hidden and not used her real identity in the forum. However, as she socializes in the network, she gives out little pieces of private information. As shown in Figure 1, each time she posts the messages, she thinks it's ok since the amount of revealed privacy is limited. Unfortunately, she does not realize that an attacker can easily track all her messages. As shown in the figure, through information aggregation over all her messages, a significant amount of privacy is lost.

Threat 3: Cross-network information aggregation.

A study in 2007 showed significant amount of member overlap in popular social networks [14]. However, users of multiple social networks may not want information from different contexts to mix up with each other. For instance, people in a professional network may not talk about their hobbies; or people do not want to reveal their education background in a leisure community. Figure 2 demonstrates cross-network information aggregation attack. In this example, Bob puts his name, work phone, work email, etc on his homepage. In a leisure community, he uses his work phone and a leisure email in his profile. He also occasionally participates in a medical network for health issues. To keep it private, he only uses his leisure email. However, as shown in Figure 2, his identities in different types of social networks are connected by the information he provides. With the developments of information retrieval techniques, an attacker could easily discover the bridges, and then aggregate information from connected sources to acquire Bob's privacy.

In all the cases, private information is voluntarily released. Users are unaware of the potential risks, or have made wrong assumptions. Considering the increasing use of social networks and development of information retrieval techniques, user privacy becomes more vulnerable without any powerful protection tools. Meanwhile, conservative users, who are aware of the risk of private information disclosure, choose not to give out any information and stay away from the benefits of social networks.

II. Social networks

A. Openness level and access equivalency

In some social networks, information is accessible to the general public, while others require registration, contribution to the network, consent of information owner, or other authentication. We use the term "*openness level*" (OL) to measure how information in a social network could be accessed. For instance, "*OL = public*" means accessible by everyone including crawlers. Social networks in this category are best in visibility but worst in preserving privacy. "*OL = registration-required*" brings better privacy. However, a customized niche search engine could still crawl these networks.

An *access-equivalent (AE)* group contains all the sites that have the same openness level. For instance, all open social networks are in the same AE group. All networks requiring simple registration are in the same group. However, *OL=Stanford-alumni* and *OL=MIT-alumni* indicates two groups.

An online social network does not necessarily belong to only one group. Many of them sit across multiple levels, e.g. partially open to public and partially restricted. Some networks allow users to choose the openness level of their information, e.g., Live Spaces let users choose from: open to public, to friends (and friends' friends), or to a specified list.

B. Honest but curious observer

Honest but curious model (semi-honest model) is widely used in computer security research (e.g. in two-party computation). In this model, all parties are assumed to follow protocols properly (honest); but they keep all inputs and intermediate results to infer knowledge (curious). We define honest but curious observers in a similar way: an honest but curious observer tries to obtain as much information about targeted user(s) as possible, and manipulates information to seek for other's privacy. However, s/he does not break any law or conduct any illegal activity, e.g. no phishing emails or hacking into any web site (s/he may register for any public forums though). Therefore, s/he remains perfectly legal (honest), but is very aggressive in seeking information (curious).

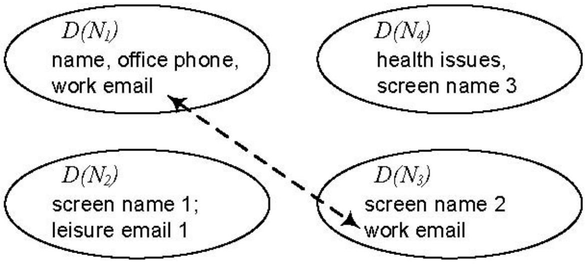

Fig. 3. Discretionary privacy model.

Conventional honest but curious parties may keep all intermediate results or previous inputs. However, this assumption is not valid in our settings, since an observer could not have the excessive resource to keep a complete history of the entire web. In our assumptions, an observer only sees what is currently available, including those in search engine caches.

III. Private information models

There are different types of private information in social networks: (1) user profiles usually contain email, phone numbers, location, and more; it also includes information submitted for the "theme" of the social network, e.g. education and employment history for LinkedIn. (2) User generated content such as blogs, messages posted to forums, etc. We have defined multilevel and discretionary models to manage private information.

In multilevel model, private information is managed in hierarchically organized categories. Information objects in higher levels are considered more private than objects in lower levels. The basic rule in this model is that information only flows from lower level to higher level and not vice versa.

In the discretionary model, private information objects are organized into sets. Objects in the same set are regarded as the same privacy level, and they could be released to one place.

Definition 1 (Privacy disclosure set)**.** In discretionary privacy model, set $D(N)$ represents the collection of private information objects that had been given or would be given to a social network N.

Figure 3 gives an example of discretionary privacy model. In this example, user defines four sets of privacy objects. S/he trusts professional network N_1 with $D(N_1) = \{\texttt{name, office phone, work email}\}$, two leisure networks with $D(N_2)$ and $D(N_3)$, and a health care community N_4 with $D(N_4) = \{\texttt{health issues, screen name 3}\}$. One item may belong to multiple sets as they could be included in different combinations (e.g., $D(N_1) \cap D(N_3) = \{\texttt{work email}\}$). In this paper, we only focus on discretionary model due to space constrains. However, please note that all our analysis and solutions are applicable to multi-level model with minor modification.

IV. Private information disclosure

In Section I, we have presented three privacy vulnerabilities in social networks. Now we formally model them. When user discloses private information $D(N)$ to a social network, an honest-but-curious observer observes $D'(N)$. When $D'(N) = D(N)$, observer sees exactly what user intends to reveal, and user privacy is preserved. However, when $D'(N) - D(N) \neq \phi$, we consider there is undesired private information disclose – user does not want observer to see $D'(N) - D(N)$.

Proposition 1 (Privacy disclosure)**.** *Private information disclosed to social network N_i is considered visible to all social networks in its access-equivalent group.*

This proposition says that when user discloses information to a social network, any observer who can access its access-equivalent group is able to see the information. Users always take it for granted that the "context" of submitted information is the targeted social network. However, the true context is the whole AE group.

Proposition 2 (Maximum privacy disclosure)**.** *Given an online social network N_k and user's privacy disclosure $D(N_k)$; user's actual private information disclosure to site N_k is:* $D'(N_k) = \bigcup D(N_{i,n})$, *where* $D(N_{i,n}) \cap D(N_{i,n-1}) \neq \phi, ..., D(N_{i,1}) \cap D(N_k) \neq \phi, \forall D(N_{i,j})$.

This proposition is defined with recursion. $D(N_{i,n})$ denotes all privacy disclosure sets that can be connected with original $D(N_k)$ through n steps. Therefore, $D'(N_k)$ is the union of all the sets ($D(N_{i,1})$) that has a (nonempty) intersection with $D(N_k)$, and all the sets ($D(N_{i,2})$) that has an intersection with $D(N_{i,1})$, and so on. If we define relation xRy as "sets x and y have intersection," then we are looking at the transitive closure of $D(N_k)$ and R.

This proposition describes information aggregation attacks. When a curious observer gets information of targeted user from one social network, s/he seeks for all the private information sets that could be connected – directly ($D(N_{i,1}) \cap D(N_k) \neq \phi$) or indirectly ($D(N_{i,2}) \cap D(N_{i,1}) \neq \phi$, and so on). For instance, when we look back to the example in Figure 2, Bob gives $D(N_M)$ to the medical network, $D'(N_L)$ to leisure network, and $D'(N_P)$ to professional network. However, due to the intersections between these sets, we can connect them together, so that $D'(N_M) = D(N_M) \cup D(N_L) \cup D(N_P)$. Note that we are assuming exact matching here. There are further questions such as approximate matching and disambiguation. Their technical details are out of scope of this paper.

To defeat against information aggregation attacks, we need to minimize $\bigcup D(N_{i,n})$, i.e. to cut off intersections between private information sets. Ideally, $D(N_k)$ should be independent from all other $D(N_i)$ – which means information disclosed to N_k cannot be linked with information disclosed to any other social network.

V. Solution overview

With the private information disclosure model described above, we are able to tackle privacy vulnerabilities introduced in Section I. We propose a *privacy monitor* for social networks users. The system structure is shown in Figure 4. Here we briefly describe major components of this system, as well as examples of challenges.

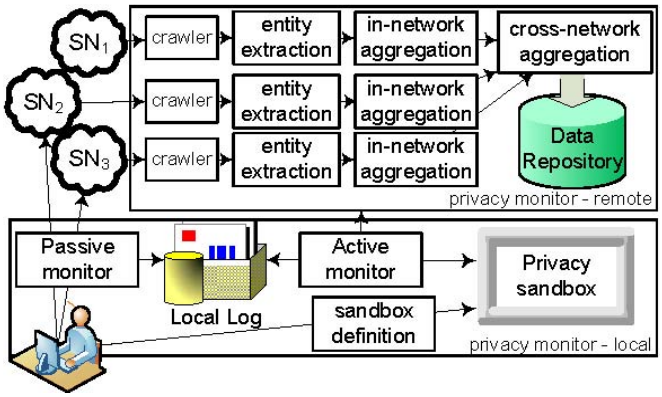

Fig. 4. System structure of privacy monitor.

1. Information extraction. First, we simulate the activities of curious observers to construct an information repository. We make use of information retrieval techniques to get data from various social networks. We also use "registered crawlers" to monitor information in semi-private networks. Practically, these crawlers may violate social networks' terms of use; this issue requires further investigation. In this phase, one major challenge is to extract entities from web pages. Due to the nature of social networks, there are plenty of structural information and patterns to be utilized. Moreover, private information in social networks is more than profiles and entities. The other challenge is to identify/extract valuable private information items.

2. Information aggregation. For the private information objects extracted in component 1, we apply in-network and cross-network aggregations, according to Proposition 2, to obtain $D'(N)$ for all users. The processed data are then stored in the data repository. As shown in the figure, components 1 and 2 consist the remote portion of privacy monitor.

In this step, information aggregation is especially challenging. For instance, we have the problem of *information object linkage*: giving two pieces of information (I_1, I_2), which are linked to any profile, the profile context "Alice", and the probability of I_1 belonging to Alice is: $P(Alice, I_1) = p_1$, determine $P(Alice, I_2|I_1)$. E.g. we found a phone number that may belong to Alice, and another email that appears with the phone number, the problem is that how we can determine if the email also belongs to Alice. Globally, we have the problem of *cross-network profile linkage*: giving two profiles from two different networks, how do we determine if they belong to the same person? Giving large volume of profiles from two social networks, how can we efficiently identify matches?

3. Privacy sandbox definition. The local portion of privacy monitor maintains user's privacy plan, called "*privacy sandbox*." In the sandbox, users define private information objects as well as *privacy disclosure sets*. Users also define the desired $D(N)$ for social networks with different OLs. In this component, a well-structured sandbox and a well-designed, easy-to-use user interface is the major challenge.

4. Passive monitor. In communicating with the remote components, the local privacy monitor maintains a list of social networks and their properties. When users are about to submit information to a social network, provides them with the OL and the *AE group* so that user knows who will be able to access the information. It also keeps a history of collections of information that users have released to any network and its *AE group*,so that user could make the judgement whether there's risk for information aggregation attack.

5. Active monitor. More importantly, privacy monitor actively tracks private information disclosed to the internet. Local tracking checks logs of information submitted to various social networks to get $D'_L(N)$. While remote tracking uses $D(N)$ as seed to check our private information repository (constructed in steps 1 and 3) to get $D'_R(N)$. Moreover, it uses $D(N)$ as seed to check with search engines to obtain $D'_S(N)$. Finally we have $D'(N) = D'_L(N) \cup D'_R(N) \cup D'_S(N)$. Whenever we discover differences between $D(N)$ and $D'(N)$, an alert is issued to the user. User could either change the settings in privacy sandbox, or correct the problem by revoking some private information.

6. Privacy report. We periodically present $D'(N)$ to the user for review. In many cases $D'(N)$ may be too large or messy to read, therefore, knowledge extraction is applied to estimate what valuable information would be learn from $D'(N)$.

There are also system level challenges. For instance, the system needs to be trusted and well secured. Because of space constrains, we do not go into details here.

VI. Related work

In recent years, web privacy has drawn significant attention from both the industry and the research community. Most of existing research efforts fall in the following categories:

(1) Privacy-preserving data processing. This includes privacy-preserving collaboration [3], preserving privacy data mining and publishing [7]. Notably, [10] infers user attributes from friends' attributes using a Bayesian network. [2] studies attacks on anonymized social network data using graph information. [16] protects social network users from neighborhood attacks, in which anonymous users are re-identified through 1-neighborhood subgraph. [13] introduces *k-degree anonymous* for social network graphs, and propose algorithms for efficient k-degree anonymization. Finally, [9] models three types of adversary knowledge that could be used to re-identify vertexes from anonymized social network graphs, and tackle the problem through graph generalization.

(2) Privacy-preserving web browsing. Most works focus on anonymity – hiding user identities in web browsing or other applications [5] [15], so that they could not be tracked down to their origin or distinguished from a set of users. [1] continuously collects information submitted to the internet and visualizes to the user, in order to stimulate self-awareness. On the other hand, there's also research on user behavior study, user education, or policy/legal issues [11], [8] [12], [4].

Our work is significantly different from existing research. Instead of using published (and anonymized) data, we work on information that is available on the web. We don't limit user from providing information – we recognize that users need to socialize, which implies information sharing. More over, existing work mostly focus on graph structure while we focus on nodes/attributes as well as structural information.

VII. Conclusion

In this position paper, we share our preliminary results in defining information aggregation attacks, modeling private information, and articulating private information disclosure under the model. We also present the overall design of our solution, namely privacy monitor. Our framework handles private information over the Internet, where social network is the largest source of such information. However, we are still able to monitor other sources such as personal webpages.

References

[1] K. Abdullah, G. Conti, and E. Sobiesk. Self-monitoring of web-based information disclosure. In *WPES '07: Proceedings of the 2007 ACM workshop on Privacy in electronic society*, pages 56–59. ACM, 2007.

[2] L. Backstrom, C. Dwork, and J. Kleinberg. Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography. In *Proceedings of ACM international conference on World Wide Web*, pages 181–190, 2007.

[3] G. Blosser and J. Zhan. Privacy Preserving Collaborative Social Network. In *International Conference on Information Security and Assurance (ISA)*, pages 543 – 548, April 2008.

[4] J. Caverlee and S. Webb. A large-scale study of myspace: Observations and implications for online social networks. In *Proceedings of the International Conference on Weblogs and Social Media*, 2008.

[5] R. Dingledine, N. Mathewson, and P. Syverson. Tor: the second-generation onion router. In *USENIX Security Symposium*, 2004.

[6] B. Dubow. Confessions of 'Facebook stalkers'. USA Today, March 2007.

[7] K. B. Frikken and P. Golle. Private social network analysis: how to assemble pieces of a graph privately. In *Proceedings of ACM workshop on Privacy in electronic society*, pages 89–98, 2006.

[8] R. Gross, A. Acquisti, and I. H. John Heinz. Information revelation and privacy in online social networks (the facebook case). In *Proceedings of ACM workshop on Privacy in the electronic society*, 2005.

[9] M. Hay, G. Miklau, D. Jensen, D. Towsley, and P. Weis. Resisting structural re-identification in anonymized social networks. *Proc. VLDB Endow.*, 1(1):102–114, 2008.

[10] J. He, W. W. Chu, and Z. Liu. Inferring privacy information from social networks. In *IEEE International Conference on Intelligence and Security Informatics*, pages 154–165, 2006.

[11] B. A. Huberman, E. Adar, and L. R. Fine. Valuating privacy. *IEEE Security and Privacy*, 3(5):22–25, 2005.

[12] M. Irvine. Social network users overlook privacy pitfalls. USA Today, April 2008.

[13] K. Liu and E. Terzi. Towards identity anonymization on graphs. In *SIGMOD '08: Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 93–106, 2008.

[14] A. Patriquin. Connecting the social graph: Member overlap at opensocial and facebook. Compete.com Blog, available at: http://blog.compete.com/2007/11/12/connecting-the-social-graph-member-overlap-at-opensocial-and-facebook/, 2007. Accessed on December 10, 2008.

[15] F. Saint-Jean, A. Johnson, D. Boneh, and J. Feigenbaum. Private web search. In *WPES '07: Proceedings of the 2007 ACM workshop on Privacy in electronic society*, pages 84–90. ACM, 2007.

[16] B. Zhou and J. Pei. Preserving privacy in social networks against neighborhood attacks. In *Proceedings of the 24th International Conference on Data Engineering (ICDE)*, April 2008.