Privacy-aware Data Assessment of Online Social Network Registration Processes

Christine Schuppler*
AIT Austrian Institute of Technology
Center for Digital Safety & Security
Vienna, Austria

Maria Leitner†
AIT Austrian Institute of Technology
Center for Digital Safety & Security
Vienna, Austria
maria.leitner@ait.ac.at

Stefanie Rinderle-Ma
University of Vienna
Faculty of Computer Science
Vienna, Austria
stefanie.rinderle-ma@univie.ac.at

ABSTRACT

Privacy and security research has been very active concerning online social networks (OSN) as a vast amount of personal information is used and published (by users) within OSNs. However, most people do not pay attention on what personal information they provide during registration. Depending on what information is provided in (public) OSN profiles, that data might be misused by attackers e.g., for cross-site profile cloning. This paper assesses data provided by the user during the registration of OSNs. Therefore, it is investigated how OSN registration processes are typically modeled, which information is needed to create a profile in OSNs and which attack scenarios can occur based on the provided data. The results contribute to the understanding of OSN registration process design as well as requested data and to replicate and reuse processes for further privacy and security investigations.

CCS CONCEPTS

• **Security and privacy** → **Social network security and privacy**; *Social aspects of security and privacy*; • **Applied computing** → Business process modeling;

KEYWORDS

privacy, online social networks, data assessment, business process

ACM Reference Format:
Christine Schuppler, Maria Leitner, and Stefanie Rinderle-Ma. 2018. Privacy-aware Data Assessment of Online Social Network Registration Processes. In *CODASPY '18: Eighth ACM Conference on Data and Application Security and Privacy, March 19–21, 2018, Tempe, AZ, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3176258.3176950

1 INTRODUCTION

Online social networks (OSN) are widely used by millions of people in everyday lives. OSNs are defined by Boyd and Ellison [4] as *"a web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system"*. Privacy and security are major concerns when it comes to OSNs as a vast amount of personal information is used and published within OSNs. However, most people do not pay attention on what personal information they provide during the registration process [8]. Personal information is extremely valuable and can be sold on black markets (e.g., Dark web). For this reason the personal data is not only interesting for OSNs and third-party domains but also for malicious actors (e.g., hackers or thieves for selling data or identity theft). Already several security attacks were specified for OSNs ins [2, 7, 10]. For example, cloning an existing profile and sending friend requests to friends of the impersonated profile to steal personal data is called same-site and cross-site profile cloning attacks. While same-site profile cloning copies and inserts the profile within the same network, cross-site profile cloning is about copying the profile from OSN A into another network (OSN B). Copying a profile to a new OSN raises less suspicion and therefore is harder to detect. Further examples of attacks are phishing attacks where personal information from OSNs is reused for phishing. For example in [9], it is shown that phishing attacks with information from OSN profiles are four times more effective. Further security attacks include spamming, creation of an digital dossier of personal data, sybil attacks, malware attacks, information leak or de-anonymizing (see [2, 7, 10]). These attacks can have a significant impact on the use of personal information. In this paper, we start our investigation at the beginning - the **OSN user registration**. It is the first process where users "share" personal information with registering in the OSN and further with other OSN users (e.g., friends).

In particular, we want to investigate the following research questions:

(1) How are registration processes of OSNs typically modeled?
(2) Which information is generally needed to create a profile in an OSN?
(3) Which attack scenarios can occur based on the data attributes in OSN registration processes?

The first question (1) will investigate how registration processes are structured in OSNs and if there are differences between OSNs of different domains (e.g., business, leisure, research). This will further support the understanding of processes in OSNs and how online registration is typically designed. With the second question (2), we investigate which minimum data requirements exist to create OSN profiles. This information can be valuable for prospective users having privacy and security concerns for deciding whether or not

[*] Contribution during work at AIT Austrian Institute of Technology.
[†] Corresponding Author.

CODASPY '18, March 19–21, 2018, Tempe, AZ, USA

ACM ISBN 978-1-4503-5632-9/18/03…$15.00
https://doi.org/10.1145/3176258.3176950

to register in OSNs. The last question (3) analyzes which threats or attacks may occur due to the sharing of personal data.

The methodology is outlined as follows. First, we investigate the processes and data use in OSN registration processes. Therefore, we select eleven popular OSNs. In order to identify the structure of registration processes of OSNs, we create and register an avatar and further investigate the process and the information that is required during the registration process. In addition, we examine how existing OSN profiles can be used during the registration process in another OSN. Further we investigate which data attributes are interesting for malicious actors and which attack can occur in which OSN due to declaration of data in the examined registration processes. Based on the results from the analysis, we derive two aggregated registration processes. Due to page limitations we only focus on the reference process model based on all data fields that must be specified or are an optional field of at least fifty percent of all registrations. Business process modeling can be used to analyze security measures (cmp. [14]).

The analysis of the eleven OSN registration processes was conducted between November 17th and 21st, 2015. OSN and their registration processes might have changed since then the analysis results provide a still relevant picture for the current state and future development of OSNs.

The results of this paper can contribute to foster and deepen the understanding of the design of registration processes in OSNs, provide an overview on which mandatory or optional data can be entered by users who want to register a profile in an OSN, simulate OSN registration processes using the two derived processes, and replicate or reuse processes for further privacy and security investigations.

2 METHODOLOGY

As outlined in Figure 1, the methodology consisted of three steps. The first step *Selection of OSNs* focused on identification and selection of OSNs according the defined characteristics in the study. The second step, *Avatar Creation*, established an avatar and all data attributes required (e.g., name, username, email). Furthermore, the step *Investigation of OSN Registration Processes* focused on the privacy-aware data assessment of selected OSNs using the avatar. In the last step *Investigating Potential Threats and Attacks* potential attack scenarios (derived by literature) are investigated based on the attributes found in the previous step. The attacker model in this paper is based on information that is provided (sometimes carelessly) by the users in OSNs and later retrieved by attackers.

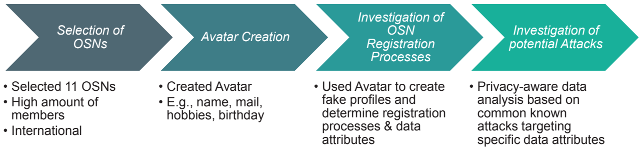
Selection of OSNs
Avatar Creation
Investigation of OSN Registration Processes
Investigation of potential Attacks
• Selected 11 OSNs
• High amount of members
• International
• Created Avatar
• E.g., name, mail, hobbies, birthday
• Used Avatar to create fake profiles and determine registration processes & data attributes
• Privacy-aware data analysis based on common known attacks targeting specific data attributes

Figure 1: Methodology

3 BACKGROUND

In recent years, OSNs have gained importance and the number of users having an OSN account has increased dramatically as shown in [4, 5, 15].

During the registration process of an OSN, user have to declare personal information (further known as data attributes on e.g., name, birthday, city, profession). This can establish a social identity that requires management. For example in [16], the management of the identities itself is discussed focusing also on e.g., the control of data of an identity or ways to get in contact or restriction of profile views. In this paper, we focus only on the process when an identity is created and first data attributes specified for the social identity.

The use and sharing of personal information in OSNs has been investigated in literature. For example [17] discovered that there are design conflicts between the security and privacy goals and the traditional goals of OSNs such as usability and sociability. Gross and Acquisti [8] show that most users of online social networks do not worry about their privacy and provide their personal information carelessly. For instance, exceedingly few users change their default privacy settings, therefore the standard-visibility of the profile is selected which purpose is to maximize the visibility. This creates physical and cyber risks such as stalking social security numbers, stealing identities (identity theft) or creating digital dossiers of the behavior of the user [12].

Moreover [11] and [13] have found that users are not aware of third-party advertisers, data aggregators, external applications and users on the OSN which are not friends having access to private information. External actions while logged into an OSN are tracked and can be used for marketing purposes and more. For example, [11] discovers the role of third party domains in aggregating user related data. Privacy protection for future OSNs has been identified e.g. the minimum and maximum of information which have to be specified for a particular set of interactions must be declared.

4 RESULTS

This paper focused on the investigation of registration processes of OSNs. We examined which information is required (or optional) to create a profile, how profiles can be reused and how data attributes can be misused for attacks. Based on the research questions in Section 1 we came to the following principle findings.

How are registration processes of OSNs typically modeled? We constructed two registration processes illustrating a *reference process model* and aggregated process model for OSN registration. The *reference process model* represents the common behavior of the 11 individual registration processes (Facebook, Twitter, Google+, Instagram, Habbo, Hi5, Twoo, Xing, LinkedIn, Academia.edu and ResearchGate). If we consider the smallest common denominator of the individual OSN registration processes with respect to the data attributes (which would contain the data which has to be specified or is an optional field in all processes), this would comprise only the *password* and *e-mail* attributes. The resulting reference process model would consist of only one step, i.e., a process activity *input personal information*. Thus the definition of considered data attributes is extended to all data attributes that are mandatory or optional in the registration processes of at least **fifty percent** of all analyzed OSNs. As a result, the *first name*, *last name*, *e-mail*

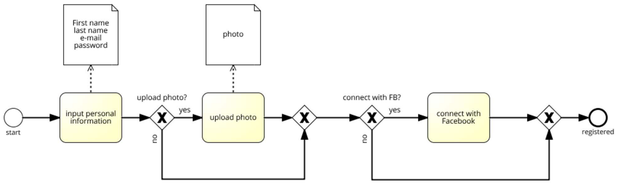

Figure 2: Reference OSN Registration Process Model for Mandatory and Optional Data Attributes in at least 50 percent of the Individual Registration Processes (in BPMN using Signavio)

Table 1: Assessing Data Attributes that can be misused for Attacks (Excerpt)

	Phishing [2, 6, 7, 10]	Profile Cloning [2, 6, 7]	Fake Profiles [6, 8]	Face Recognition [2, 6, 8]
Public profile	✓	✓	✓	
Account Username	✓			
E-Mail	✓			
Birthday	✓			
Photographs				✓
Passwords	✓			
Friendslist		✓	✓	
Credit Card Data	✓			
PIN	✓			
TAN	✓			

✓... utilized data for threat

Public profile: name, location and contact information, educational and employment history, personal preferences, interests and profile photo

and *password* are mandatory fields and are shown in Figure 2. In addition, *connect the profile with Facebook* and *upload a photo* were also optional in most selected OSNs.

Which information is generally necessary to create a profile in an OSN?. We investigated which personal data has to be declared during the registration processes OSNs. The only data which has to be specified in every registration process is the password and e-mail. The first name and last name have to be declared in more than fifty percent of the OSNs, the birthday and gender in exactly half of the selected networks. Uploading a photo or connect the profile with an e-mail account is mostly optional.

Which attack scenarios can occur based on the data attributes in OSN registration processes? Based on a literature review, we analyzed several attacks, i.e. phishing, profile cloning, fake profiles, face recognition (see Table 1) as potential attack scenarios (based on e.g., [2, 6–8, 10]). Luckily, we found that most personal information is optional to specify within the OSN registration processes. However, this does not signify that most users will or will not specify this information during registration. Further analysis would be required. It can be seen for example that many data attributes in OSNs are mandatory or optional. Hence, more publicly specified information can be misused by others and can lead to potential attacks (e.g., fake profiles).

Interpretation of Findings. The most investigated OSNs demand rights of personal information and data of the user during the registration process. The current state of registrations shows that there are possible improvements for the security of the personal information of users. Future OSNs should focus on user privacy and security. For instance, [3] show an approach of an OSN where the user defines the policy over access to private data instead of the OSN by using attribute-based encryption. We are currently not aware of any commercial OSNs using such approaches. One reason could be that personal information is extremely valuable for marketing purposes. For this reason it could be crucial to minimize sharing personal information in the internet to protect the data against third party applications (or use adequate OSN settings).

Finally, it will always depend on the user and which information he/she will share with others in an OSN. The behavior of this changes over time such as shown in [1]. As this paper is mostly about sharing and handling personal information, we feel it is our duty to refer to recommendations on how to handle personal information carefully. We found that recommendations given in [6] help to minimize the usage of personal information and to prevent misuse in general.

REFERENCES

[1] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2015. Privacy and human behavior in the age of information. *Science* 347, 6221 (Jan. 2015), 509–514. https://doi.org/10.1126/science.aaa1465

[2] Abdullah Al Hasib. 2009. Threats of online social networks. *IJCSNS International Journal of Computer Science and Network Security* 9, 11 (2009), 288–93.

[3] Randy Baden, Adam Bender, Neil Spring, Bobby Bhattacharjee, and Daniel Starin. 2009. Persona: an online social network with user-defined privacy. In *ACM SIGCOMM Computer Communication Review*, Vol. 39. ACM, 135–146.

[4] Danah M. Boyd and Nicole B. Ellison. 2007. Social Network Sites: Definition, History, and Scholarship. *Journal of computer-mediated communication* 13, 1 (2007), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

[5] Maeve Duggan, Nicole B Ellison, Cliff Lampe, Amanda Lenhart, and Mary Madden. 2015. Social media update 2014. *Pew Research Center* 9 (2015).

[6] Michael Fire, Roy Goldschmidt, and Yuval Elovici. 2014. Online social networks: threats and solutions. *IEEE Comm. Surveys Tutorials* 16, 4 (2014), 2019–2036.

[7] Hongyu Gao, Jun Hu, Tuo Huang, Jingnan Wang, and Yan Chen. 2011. Security issues in online social networks. *IEEE Internet Computing* 15, 4 (2011), 56–63.

[8] Ralph Gross and Alessandro Acquisti. 2005. Information Revelation and Privacy in Online Social Networks. In *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society (WPES '05).* ACM, New York, NY, USA, 71–80.

[9] Tom N Jagatic, Nathaniel A Johnson, Markus Jakobsson, and Filippo Menczer. 2007. Social phishing. *Commun. ACM* 50, 10 (2007), 94–100.

[10] Prateek Joshi and C-C Jay Kuo. 2011. Security and privacy in online social networks: A survey. In *2011 IEEE Int. Conf.e on Multimedia and Expo.* IEEE, 1–6.

[11] Balachander Krishnamurthy and Craig E. Wills. 2008. Characterizing Privacy in Online Social Networks. In *Proceedings of the First Workshop on Online Social Networks (WOSN '08).* ACM, New York, NY, USA, 37–42.

[12] Katharina Krombholz, Heidelinde Hobel, Markus Huber, and Edgar Weippl. 2015. Advanced social engineering attacks. *Journal of Information Security and Applications* 22 (2015), 113 – 122.

[13] Katharina Krombholz, Dieter Merkl, and Edgar Weippl. 2012. Fake identities in social media: A case study on the sustainability of the facebook business model. *Journal of Service Science Research* 4, 2 (2012), 175–212.

[14] Maria Leitner and Stefanie Rinderle-Ma. 2014. A systematic review on security in Process-Aware Information Systems – Constitution, challenges, and future directions. *Information and Software Technology* 56, 3 (March 2014), 273–293.

[15] Alan Mislove, Hema Swetha Koppula, Krishna P Gummadi, Peter Druschel, and Bobby Bhattacharjee. 2008. Growth of the flickr social network. In *Proceedings of the first workshop on Online social networks.* ACM, 25–30.

[16] Moritz Riesner, Michael Netter, and Günther Pernul. 2013. Analyzing settings for social identity management on Social Networking Sites: Classification, current state, and proposed developments. *Inf. Sec. Technical Report* 17, 4 (May 2013), 185–198.

[17] Chi Zhang, Jinyuan Sun, Xiaoyan Zhu, and Yuguang Fang. 2010. Privacy and security for online social networks: challenges and opportunities. *IEEE Network* 24, 4 (2010), 13–18.