Accepted Manuscript

Privacy Leakage Analysis in Online Social Networks

Yan Li , Yingjiu Li , Qiang Yan , Robert H. Deng

PII: S0167-4048(14)00158-8

DOI: 10.1016/j.cose.2014.10.012

Reference: COSE 847

To appear in: *Computers & Security*

Received Date: 16 June 2014

Revised Date: 16 September 2014

Accepted Date: 19 October 2014

Please cite this article as: Li Y, Li Y, Yan Q, Deng RH, Privacy Leakage Analysis in Online Social Networks, *Computers & Security* (2014), doi: 10.1016/j.cose.2014.10.012.

Privacy Leakage Analysis in Online Social Networks

Yan Li[a,*], Yingjiu Li[a], Qiang Yan[a], Robert H. Deng[a]

[a]*School of Information Systems, Singapore Management University*

Abstract

Online Social Networks (OSNs) have become one of the major platforms for social interactions, such as building up relationship, sharing personal experiences, and providing other services. The wide adoption of OSNs raises privacy concerns due to personal data shared online. Privacy control mechanisms have been deployed in popular OSNs for users to determine who can view their personal information. However, user's sensitive information could still be leaked even when privacy rules are properly configured. We investigate the effectiveness of privacy control mechanisms against privacy leakage from the perspective of information flow. Our analysis reveals that the existing privacy control mechanisms do not protect the flow of personal information effectively. By examining representative OSNs including Facebook, Google+, and Twitter, we discover a series of privacy exploits. We find that most of these exploits are inherent due to the conflicts between privacy control and OSN functionalities. The conflicts reveal that the effectiveness of privacy control may not be guaranteed as most OSN users expect. We provide remedies for OSN users to mitigate the risk of involuntary information leakage in OSNs. Finally, we discuss the costs and implications of resolving the privacy exploits.

Keywords: Online social network, privacy control, information flow, private information leakage, inherent exploit

1. Introduction

Online Social Network services (OSNs) have become an essential element in modern life for human beings to stay connected to each other. About 82% online population use at least one OSN such as Facebook, Google+, Twitter, and Linkedln, which facilitates building relationship, sharing personal experiences, and providing other services [1]. Via OSNs, massive amount of personal data is published online and accessed by users from all over the world. Prior research [2, 3, 4, 5] shows that it is possible to infer undisclosed personal data from *publicly* shared information. Nonetheless, the availability and quality of the public data causing privacy leakage are decreasing due to the following reasons: 1) privacy control mechanisms have become the standard feature of OSNs and keep evolving. 2) the percentage of users who choose *not* to publicly share information is also increasing [4]. In this tendency, it seems that privacy leakage could be *prevented* as increasingly comprehensive privacy control is in place. However, this may not be achievable according to our findings.

Instead of focusing on new attacks, we investigate the problem of *privacy leakage under privacy control* (PLPC). PLPC refers to private information leakage even if privacy rules are properly configured and enforced. For example, Facebook allows its users to control over who can view their friend lists on Facebook. Alice, who has Bob in her friend list on Facebook, may not allow Bob to view her complete friend list. As an essential functionality, Facebook recommends to Bob a list of users, called "*people you may know*", to help Bob make more friends. This list is usually compiled by enumerating the friends of Bob's friends on Facebook, which includes Alice's friends. Even though Alice doesn't allow Bob to view her friend list, Alice's friend list could be leaked as recommendation to Bob by Facebook.

We investigate the underlying reasons that make privacy control vulnerable from the perspective of information flow. We start with categorizing the personal information of an OSN user into three *attribute sets* according to *who the user is*, *whom the user knows*, and *what the user does*, respectively. We model the information flow between these attribute sets and examine the functionalities which control the flow. We inspect representative real-world OSNs including Facebook, Google+, and Twitter, where privacy exploits and their corresponding attacks are identified.

Our analysis reveals that most of the privacy exploits are inherent due to the underlying conflicts between privacy control and essential OSN functionalities. The recommendation feature for social relationship is a typical example, where it helps expanding a user's social network but it may also conflict with other users' privacy concerns for hiding their social relationships. Therefore, the effectiveness of privacy control may not be guaranteed even if it is technically achievable. We investigate necessary conditions for protecting against privacy leakage due to the discovered exploits and attacks. Based on the necessary conditions, we provide suggestions for users to minimize the risk of involuntary information leakage when sharing private personal information in OSNs.

We analyze the potentially vulnerable users due to our identified attacks through user study, in which we investigate participants' usage, knowledge, and privacy attitudes towards Face-

*Corresponding author

Email addresses: yan.li.2009@smu.edu.sg (Yan Li), yjli@smu.edu.sg (Yingjiu Li), qiang.yan.2008@smu.edu.sg (Qiang Yan), robertdeng@smu.edu.sg (Robert H. Deng)