Privacy Issues on Social Media: A Tool for Raising Privacy Awareness on Social Media

Bum Mook Oh[1], HeeYoon Byun[1], and Arvindram Krishnamoorthy[1]

[1] University of Washington, Seattle WA 98105, USA

Abstract. Personal information on the web must be handled very delicately for its exposure can leave the victim with their identity stolen and destitute. Although security measures are normally implemented to prevent such breaches, targeted hacks continuously occur. However, the authors noticed that among 36 of the Facebook groups they frequent, 70% in aggregate willingly exposed their emails and other identifying details in exchange for detailed job recruitment reports and other related resources. First, we created a search engine that returned all the public posts where an individual commented their email. We segmented these commenters into three groups each with different treatment. We conducted a pre-post study over the course of five months that incorporated clickstream analysis and quantitative records for all groups using the search engine to determine whether both it and the survey in tandem measurably reduced the number of times the user submitted personal information after being confronted with the grave ramifications such behavior could have on their social, mental, and monetary wellbeing. From this, we determined how best to persuade commenters to reduce the number of posts they make in the future through evaluating the cogency of our search engine.

Keywords: Privacy, Social Media, Awareness.

1 Introduction

The purpose of this paper is to raise public awareness of the risk of personal information exchange on social media. Hackers exposed 110 million Americans accounts during 2014 [1]. By this we mean that access to an account is gained without express permission from or awareness of the user. While this enormous number of people are struggling with privacy issues that may result in phishing attacks, where a malicious individual employs various tricks to obtain a user's credentials, and identity theft, which is the successful acquisition of the aforementioned credentials for monetary benefits, we found that some users on social media, Facebook voluntarily reveal their personal information, especially their email address in the comments for detailed job recruitment and free resources posts. This prompts questions concerning the main factors of such behavior and how existing security measure can handle them. This article will examine and investigate why the unusual and dangerous behavior of a user attempting personal information trade-off occurs by conducting a user survey, and will

also explore effective methods of raising public awareness and preventing potential hacking attacks.

2 Research Method

2.1 Method

A Following the creation of Facebook search engine, we will test and measure the percentage of users who decreased in the number of commented emails on Facebook threads after sending out a notification email. Initially, we will separate the commenters into three groups: "inform," "persuade," and "control." The "inform" group will receive the notification that their email was discovered from public Facebook threads. The "persuade" group will receive the same notification, as well as a link to a survey. If the commenter takes the survey, it will inform the user the dangers of having their email on the public Facebook feed. The "control" group will remain unmanipulated from our researchers.

2.2 Facebook Crawler

A python crawler was created to scan comments on Facebook groups. When an email was found in the comments, it was stored in the private Google firebase database along with any other data about the post where it was commented. To eliminate any threats to the database itself, all data was hashed before being stored. The following code is a portion of the crawler. The full code can be found at [2].

2.3 List of Crawled Facebook Groups

A total of 36 Facebook groups were found which asked users to comment emails on posts in order to receive resources. The chosen 36 pages were found by searching related terms such as ‘comment your email’, ‘we will email you’ on Facebook’s search engine. The following is the list which the crawler examined comments.

Table 1. Crawled Facebook Groups and their Provided Resources.

Provided Resources	Facebook Groups
Users were promised packaged cover letters, resumes of past successful applicants for companies.	Jobuniv, publicdangi, specupcafe, naverdakchi, dokchi4, gongchisa, betomorrow.co.kr, incruit, jobcheatkey, jobshopper, dokkmsa, jobplanetkorea
Shares preformatted PowerPoint slides targeted for college students.	Powerpnt82, pptparty, dauernPPT, indesual, 2ggamang
Fan clubs for musicians, sharing live footage of the musician’s performance.	Beautifulintwice, MINA0324, infitspirit, 907737702604366, ioihistory, TwiceTTOnce

Independent entertainment companies sharing their work to individual fans for promotional purposes.	Gamseonglowkick, kiwiproductions
Independent musicians sharing their work with fans for promotional purposes.	MassiveDitto, MoonRecordsLabel, djdoublejkorea
Shares educational content for students interested in Universities or higher education.	Gogosusi, yiersanChina, sinagong, englishrangcafe, schoolchcafe, yiersanChina, sinboonup, brucenfriends

2.4 User Interface

The user interface where users can search their emails. The search engine has minimal design for the sole functionality of the engine is to search the user's information (See Fig. 1). When a search is made, a call is made to the server to validate the request and searches the database. The results are then displayed on the results section. By clicking on the post, the user is directed to the actual post on Facebook.

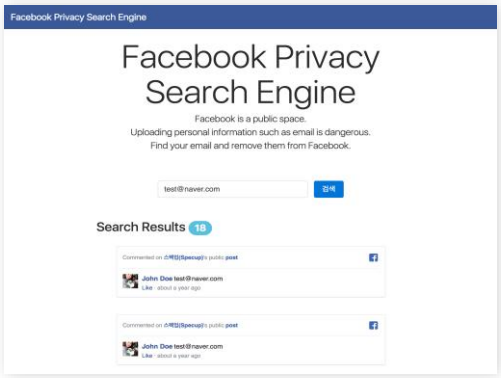

Fig. 1. Search Engine webpage, when a user searches their email, the results are shown in a list

3 Findings

3.1 Crawling Results

A total of 674,889 posts were examined and total of 19,350,900 comments of these posts were examined. In total, 407,774 users were found who have posted emails in the comments.

3.2 User Survey

A survey including demographic and qualitative questions is displayed under the search results after the user retrieves the Facebook threads by entering their email address. We are currently in the process of collecting the insightful survey results that will be used for final analysis.

3.3 Pre-Post Study Results

One week after the user completes the survey, we will restart the crawler. We intend for the email notifications to impress upon the users that unnecessarily leaking of their email can have grave consequences. We will measure the percentage of users who continue to comment their emails after this time period, hoping to see a decrease.

Another session of crawling will be made to compare the before and after the user is aware of their email addresses which are public. We hope to measure the percentage of users who decrease the number of commenting emails.

4 Conclusion

Currently, most social media, including Facebook, do not seem to have a particular solution to the privacy issue arising from the exchange of personal information. To investigate the root problem of users' insensitive behaviors, we conducted a user survey with 45 participants who commented on a thread with an email address. As we decided that raising awareness is the key, we asked the research participants to find their comments on actual posts by using a search engine, python crawler, that scans comments on Facebook groups. Throughout post-study of clickstream analysis and measuring quantitative records, it transpired that this solution alerted people´s attention to how easily their information could be located on a public post. Finally, we proved that increasing awareness is another tool to tackle this social issue and we hope that social media become more proactive about developing alternative solutions for this phenomenon.

References

1. Pagliery, Jose. “Half of American adults hacked this year.” CNNMoney, Cable News Network, money.cnn.com/2014/05/28/technology/security/hack-data-breach/index.html. Accessed 18 Sept. 2017.
2. Oh, Bum Mook. fb-privacy-project, (2017), GitHub repository, https://github.com/5tiger-jelly/fb-privacy-project/blob/master/crawler.py