Privacy Issues in Social Networking Websites: Is Facebook Revealing Our Social Life?

Somidip Dey
School of Computer Science
The University of Manchester
Manchester, United Kingdom.
Somidip.Dey@postgrad.manchester.ac.uk

Sayak Sen
Department of Computer Science
Sikkim Manipal University
Gangtok, India.
S.sayak.sen@gmail.com

ABSTRACT
With the proliferation and expansion of social networking websites like facebook, google+, myspace, etc., the most important question that arises in the mind of security analysts is that: "Are these social networking websites enhancing our lives or exposing them to others?” In this poster, the authors try to analyze the different privacy issues related to facebook over the years and try to answer the critical question asked before. As a conclusion, it can be inferred that a lot of information can still be revealed about an OSN (Online Social Network) user’s social life to third parties.

Categories and Subject Descriptors

General Terms

Keywords
Social Networks, Data Breaches, Web Crawling, Web Spider, Privacy.

1. INTRODUCTION
In our technologically saturated and digitally enhanced environment, almost everything can be sourced from the internet. Now, human beings live most part of their life in the virtual world of the internet, where we can post our opinion, chat or send messages to other people, who are also virtually connected. We can even connect with new people and befriend them, who are with the same interest or strata. All these features are profound in social networking websites for example facebook, google+, etc. and over the last few years it has been evident that these social networking websites are becoming more acceptable and usable by most of the people, mainly the youngsters. But with the rise of these social networking websites, there is also an increase in privacy concerns among the security analysts and socialists.

This poster paper focuses on the privacy issues of facebook, which is now the mostly used social networking website out there. According to statistics provided by Facebook Inc., there are more than 1.11 billion active users on facebook now (Facebook Inc., 2013).

2. FACEBOOK AND TYPE OF EXTRACTABLE DATA
According to research published on 2009 [1], most users belonged to at least one of the 531 regional networks, 9,764 university networks, 129,168 high school networks in the year 2009 and over the years these statistics have increased. So the amount of data available through facebook is huge and one of the top targets for potential data aggregators [2]. All these data can be collected using web-spidering or web-crawling [2] programs (scripts) and then these can be used in various purposes.

The data of interest that the potential aggregators may look for [2],[6],[7] are the following:

- **Profile Data**, which consist of the personal data, for example user’s name, user’s picture, location, contact information, educational history, job status, relationship status, personal preferences, interest, etc., which are uploaded by the user voluntarily.

- **The Social Graph**, which is a graph consisting of users (vertices) and their friendship links (edges), denoted by the following formula: \( G = (V, E) \), \( V \) is the set of users ('v' is each user), where ‘v \( \in \) V’ , and E is the set of friendships ('e' is each friendship), where ‘e \( \in \) E’. This graph reveals all the users connected to each user, even in common groups and hence this data can be used to efficiently collect user profiles and other data related to those profiles.

- **Traffic Data**, which consist of user's IP addresses, web-browser information, session information such as frequency and length of user logged in. This data is of least interest because it can not be aggregated by data aggregators.

3. DATA EXTRACTION METHODS
There are various methods of extracting data [1][2][3][5], which are the following:

- **Public Listings**, which do not require an account to access the data related to a user and is the easiest way for crawling. These data are in public view and can be seen by anyone.

- **False Profiles**, where a data aggregator may compromise a valid email address to gather more information connected to a user, than the one collected form the public listing. Sometimes sending a False Friendship Request can also be considered as the same method of having a False Profile.

- **Profile Compromise and Phishing**, where access to a user’s profile can be gained illegitimately either by operating system attacks such as malwares and key-loggers, or by phishing attacks. This method is more effective than creating a false profile in order to gain access to information of a user and its associates.

- **Malicious Applications**, where the malicious applications developed using facebook API, can gain access to
information of a user if the privacy settings permits them to do so and sometimes these applications can also be used for phishing purposes.

- **Facebook Query Language (FQL)**, where facebook permits to make queries using a subset of SQL database query language. Research (J. Bonneau, 2009) proves that FQL can be used to fetch UIDs (user IDs) and related data even from different set of groups, and it would take approximately 2000 machine days to exhaust the facebook’s allocated UID space of $2^{10^7}$.

4. FACEBOOK’S PRIVACY ISSUES IN 2013

Public listing itself reveals a lot of information even without the need of an authentic account on facebook. When public listing was first introduced, it used to reveal 10 friends of a user, and the original idea was to encourage users to create an account on facebook [6]. During a research in 2009 [1], it was revealed that on the public listing of a user, 8 friends related to that user are exposed for data aggregation. These 8 friends can also be from the same group to which a user belongs to on facebook. An example of public listing feature in 2009 can be found from Figure 1. Although the original idea was to encourage users to create an account on facebook, but at the same time these data can be used by potential data aggregators in different ways, sometimes dangerously also [5][7].

Now, in 2013, facebook has dealt with this privacy issue by not revealing related people of a user in the public listing mode. Instead the favorites of a user such as music, books, movies, television, games, etc., providing a list of top 4 selections of each favorite [4]. An example is shown in Figure 2. But now, although this has made the collection of information related to other users difficult through public access. But at the same time the choices of favorites gives any application developer to access the users’ profiles with same or related favorites’ choices using web-crawling or FQL. Since, the access to social graph is no more possible using public listing, now data aggregators can access graphs other than social one, with vertices such as choice of favorite and edges as the number of likes by users for that choice. Now this can be revealing different information about a user, which in case, can be used for the benefits of the data aggregators. For example: Using FQL, developers can access different users’ information from one user by using the Social Graph, then by using web-crawling/spidering their choices of favorites can be accessed and collected by an organization and then that organization can use that information for different purposes such as targeted phishing to other users with same choice preference or even designing products according to those choice preferences. This example is more justifiable in Figure 3.

5. LIST OF FIGURES

![Figure 1: Public Listing View of the account of J. Bonneh](image1)

![Figure 2: Public Listing View of the account of one of the authors](image2)

![Figure 3: Exploitation of Facebook User’s Personal Data](image3)

6. CONCLUSION

After reviewing all the privacy aspects and issues related to facebook, it is evident that, although facebook has changed their privacy policies over the years, but still there are plenty of opportunities for different data aggregators to extract data from a particular user or a group of users with same choices of preference. This proves that facebook, to some extent, is revealing OSN user’s social life to the third parties, which should be addressed and solved in order to secure the privacy of a user in true sense.

7. ACKNOWLEDGMENTS

SD would like to thank Juncheng Li and Shaohsuan Cheng for their valuable comments on this research.

8. REFERENCES


Privacy Issues in Social Networking Websites: Is Facebook Revealing Our Social Life?

Somdip Dey  
School of Computer Science,  
The University of Manchester.  
United Kingdom.

Sayak Sen  
Department of Computer Science,  
Sikkim Manipal University.  
India.

Introduction
In our technologically saturated and digitally enhanced environment, almost everything can be sourced from the internet, and OSN users spend most part of their life in social networking websites like facebook, google +, etc.

But at the same time these social networking websites leak different data related to OSN users and hence, jeopardising the privacy of the users.

Motivation
Facebook is now inextricably the mostly used online social networking website with more than 1.11 billion active users. So privacy of the users are of upmost importance.

Type Of Extractable Data on Facebook
- Profile Data
- The Social Graph
- Traffic Data

Data Extraction Methods
- Public Listings
- False Profiles
- Profile Compromise and Phishing
- Malicious Applications
- Facebook Query Language (FQL)

Conclusion:
Even after restricted public listing, there are still many ways to leak OSN user’s social data to third parties.

References: