Privacy Issues of Graph Based Search in Social Networking Sites

Deveeshree Nayak
The University of Memphis
Department of Computer Science
Memphis, Tennessee
dnayk@memphis.edu

Rosario Robinson
Anita Borg Institute, Programs
Atlanta, Georgia
rosarior@anitaborg.org

Summer Prince
Tennessee Technological University
Department of Computer Science
Cookeville, Tennessee
smolmstead21@students.tntech.edu

ABSTRACT
Social networks continue to be a primary interaction method in society. Supporting the front end of social networks are enormous backend databases containing large amounts of personal information about users. Needs for efficient searching of the social networking databases for useful information has driven the adoption of the graph-based search in the context of social networks. In this paper, we present a survey of graph based search engines in social networking and the security and privacy issues they introduce.

Categories and Subject Descriptors
H.2.0 [**Information Systems**]: General – *Security, integrity, and protection.*

General Terms
Security, Human Factors

Keywords
Graph Based Search, Social Networks, Security, Facebook

1. INTRODUCTION
In our present world, most individuals desire to stay connected with friends, relatives, and social groups via the Internet. Social networking sites offer a chance for individuals to make friendship and interconnect with each other. In spite of modern technologies and advanced searching algorithms, it is still difficult for users to search and get relevant results using the social networking search engines. Google provides search results of interlinked data through Google search, but is largely unable to provide search results correlated directly to the relationship of an individual. Social networking sites such as Facebook, Google+, Tumblr, Twitter, provide search engines similar to Google search, where data will be searched based on the exact word. Users of these websites have grown to billions but the adopted traditional search methods are insufficient to manage such big data. Database designers continue to be challenged to determine the best method for handling and querying big data. This paper provides a detailed survey of privacy issues of the use of graph-based search in social networking sites. More specifically, this survey paper will focus on Facebook (FB) graph search.

2. GRAPH BASED SEARCH
Graph based search is a style of semantic search, which was designed to reply to the user's queries. This search methodology is helpful for the users who have diverse interest. Facebook (FB) graph search is a brand of semantic search engine that was presented in March 2013 [4]. FB adopted this search technique to respond to the user queries on the big data gathered from its billions of users. The FB graph search algorithm searches for the information inside the user's friend network and provides numerous filters for the results, for instance, searching for friends with similar interest like software engineering and operating system. A simple example of using FB graph search is shown in Figure 1.0 where a user searches for friends and friends of friends that like sushi in New York. While FG graph search offers users a new search feature, graph search also introduces major privacy issues. FB graph search provides the general public with information about user profiles that disclose details about the users' personal lives [3].

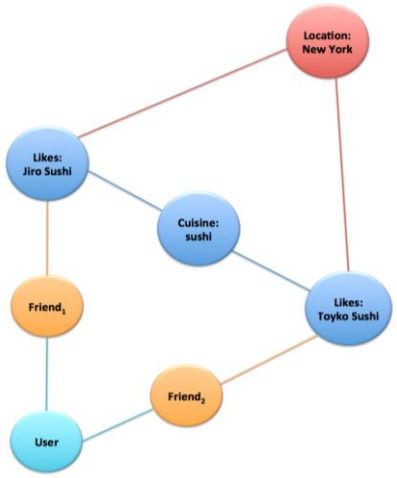

Figure 1.0 Example of Facebook Graph Search

3. GRAPH SEARCH PRIVACY ISSUES
When a user is searching for a desired topic, FB graph search will provide suggestions related to the user's query. While FB graph search was designed to guard the privacy of its users through existing settings, many privacy issues remain unsolved. For example, if a correspondent wants to interview a person then they can visit that person's FB profile to gather information. Additionally, the information is collected through searching photos, locations, and links with people based on interests. FB graph search provides results that a user has permissions to see elsewhere in FB. An example is seen when a friend has uploaded a photo and set it to be publicly viewable, while one of the friends in the photo has their privacy setting on friends only. In this case the user has to request their friend to change the privacy setting or remove the photo. All content published as public will be available through FB graph search regardless of the user's privacy settings. Photo comments are available in FB for an extended and undefined period of time. For example, availability of a user's bachelorette party photo album containing inappropriate photographs and/or comments user may create problem in her personal life. Discovery of the bachelorette photo album and comments by people outside of the user's approved network can be termed as privacy by anonymity. Another example is seen by the case of an intruder tracking a user using FB graph search to learn about the target user's interests, hobbies, and locations. FB provides the following granular privacy settings and tools:

- Who is able to see a user's future posts and review all user's posts things they're tagged in
- Limit the audience for posts
- Who is able to send friend requests and messages are filtered in the Inbox
- Who is able to see user's personal information such as email address and phone number
- Allowing search engines to link to a user's timeline

Despite of all the security features and tools offered by FB, the security and privacy of users are breached. FB has default settings that all user profiles are searchable. It means a user is not able to hide their profile information from the graph search engine. FB and Google save the entire history of every user's search, which is useful in the case of a terrorism investigation, so the investigation teams can utilize the terrorist's history including locations, interests, and accessing times. But these advantages for tracking malicious users are a threat to the privacy of a normal user. Google has a setting where users have a choice whether they want to save their chat history or not. In this instance, the privacy of an individual is not compromised, but all other history of user behavior is saved by Google. Hence, it is a major privacy concern to individuals. An example of the increase in the exposure of a Facebook user's private information by their privacy settings is shown in Figure 2.0. If the user has a privacy setting to share information with only their network of friends, the exposure is significantly less than if the user has their privacy setting set to the network of friends of friends.

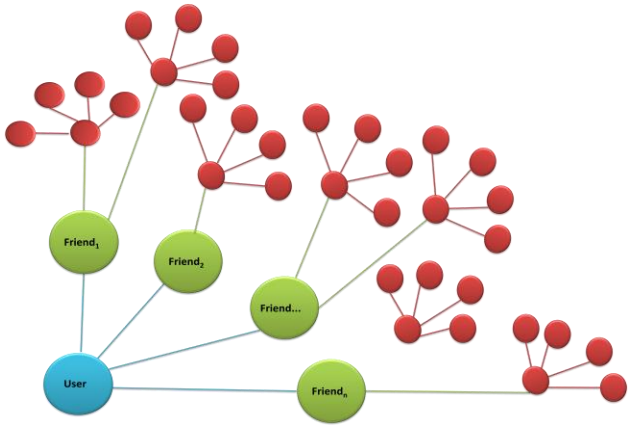

Figure 2.0 Increase of Exposure in a Network through Private Information Exposure by Privacy Settings

4. ADDITIONAL FEATURES INTRODUCED IN GRAPH BASED SEARCH

Unicorn [4] is a newly designed social graph indexing system which is used for searching the edges between tens of billions of users and entities on thousands of commodity servers. This technology is utilized for information recovery. It is built to answer billions of user queries per day at latencies in the hundreds of a millisecond, and it acts as an infrastructural building block for FB graph search. The paper published by Facebook, *Unicorn: A System for Searching the Social Graph* explains the design and query language maintained by Unicorn. Additionally, it explains the origin of how it converted the primary backend for FB's search engines. FB discloses a "public view" of user profiles to search engines that adds eight of the user's friendship links, and the results of the work of Bonneaua et. al validates that it is not safe to disclose incomplete information about a user's social network [5].

5. CONCLUSION

Security and privacy changes with the rapid advancement of technologies used on user's private information in social networking sites such as FB. Researching on new technologies and their impact on privacy helps us to develop new methods to protect the privacy of users on social network sites. In *Retrospective Privacy: Managing Longitudinal Privacy in Online Social Network*, a survey of 193 Facebook user participants were asked to explain their sharing priorities and purposes to posts which were available in dissimilar amounts of time [6]. The results indicated that a user was significantly less willing to share the information as time increased from the publishing of the information. But with graph based search technology the information shared in the past can easily be searched and can affect the privacy of a person. Privacy in FB graph search is an ongoing area of research.

6. REFERENCES

[1] Haynes, Jonathan and Perisic, "Mapping Search Relevance to Social Networks," Proceedings of the 3rd Workshop on Social Network Mining and Analysis (SNA-KDD '09), USA, 2009, ISBN: 978-1-60558-676-2.

[2] Hyoungshick Kim and Joseph Bonneau, "Privacy-Enhanced Public View for Social Graphs," Proceedings of the 2nd ACM workshop on Social web search and mining (SWSM '09), Pages: 41-48, USA, 2009, ISBN: 978-1-60558-806-3.

[3] Kumar, Ravi, et.al, "The Web and Social Networks," Published in ACM journals, Volume 35, Issue 11, Nov 2002, Pages: 32-36.

[4] Michael Curtiss, et.al, "Unicorn: A System for Searching the Social Graph," in proceeding of VLDB Endowment, Volume 6, Issue 11, Aug 2013.

[5] Joseph Bonneau, et.al, "Eight Friends Are Enough: Social Graph Approximation via Public Listings," Proceeding of second ACM EuroSys Workshop on social Network System, Pages 13-18, USA, 2009.

[6] Oshrat Ayalon, et.al, "Retrospective Privacy: Managing Longitudinal Privacy in Online Social Networks," Proceedings of the Ninth Symposium on Usable Privacy and Security Article No. 4, USA, 2013.