An attack on de-anonymization of social networks

Sri Vivek Datta

February 8, 2016

Nowadays many social networking sites are being introduced, which is showing an impact on the individual's privacy. there are many websites naming some like Facebook, twitter, Xing. These create a way for stealing the data of the data which do basically record the personal data of an user such as Name, DOB, Place and additional information like the browsing location and the type of system and operating system one uses. compared to other sites with the privacy issues, social sites do track more information.

There are many ways of tracking one's identity i.e, by using the cookies , history, internet and also browser ID string. But users privacy is so strong on the social sites, but they all have at least one group membership which can be used to break their privacy. Hackers are making use of these sites to exploit the individual's identity using de-anonymization attack which is stealing the information about an individual by redirecting them through other sites which can be done using the user's group membership information.

A social network is modeled using a set of nodes and edges where nodes are considered as users and edges as relationships between them. They are related using the following equation,

$G = (V, E)$, V as nodes(users) and E as edges(relationships).

And social sites do have groups are interrelated to each other. A browsers history is the weak point of all which is recorded and after certain time interval removed. this is the time which allows an attacker to steal the information. The attacker basically has two assumptions, one is what are the pages that an user is accessing and the second assumption is learning about the members of the group and their activities. Attackers do use this weak point because the user groups are basically two types:
1. Public
2. Closed.

The Web applications which are related to these social sites are sometimes vulnerable as they are linked using Hyperlinks connected to web servers which helps in stealing the information from the sites which are still being developed like Xing and some other small websites. For example, a simple malware attack is sending a message to an user like this Hello, your computer is infected Please clean it. Which is a major way of infecting an user by phishing the personal information. That is the reason why the browser warns the user you are entering an harmful website do you wish to continue.

The attacker first performs a basic step of information collection and then compares those to the previous data like the browser used or is it accessed previously or how many times it is hit. Basically when we click on a user ids hyperlink it is numerical so it is easy to predict. In general, the group size in a social site is limited, so the threshold is high enough to find all the members in the group.

As the attacks are harder these days, an user has to be directed to a malicious website which allows the user to steal the personal information.

My Github repository https://github.com/srivivekdatta/security-algorithms.git