

What online social networks may know about non-members

April 30 2012

(Phys.org) -- What can social networks on the internet know about persons who are friends of members, but have no user profile of their own? Researchers from the Interdisciplinary Center for Scientific Computing of Heidelberg University studied this question. Their work shows that through network analytical and machine learning tools the relationships between members and the connection patterns to non-members can be evaluated with regards to non-member relationships. Using simple contact data, it is possible, under certain conditions, to correctly predict that two non-members know each other with approx. 40 percent probability.

For several years scientists have been investigating what conclusions can be drawn from a computational analysis of input data by applying adequate learning and prediction algorithms. In a social network, information not disclosed by a member, such as sexual orientation or political preferences, can be “calculated” with a very high degree of accuracy if enough of his or her friends did provide such information about themselves. “Once confirmed friendships are known, predicting certain unknown properties is no longer that much of a challenge for machine learning”, says Prof. Dr. Fred Hamprecht, co-founder of the Heidelberg Collaboratory for Image Processing (HCI).

Until now, studies of this type were restricted to users of social networks, i.e. persons with a posted user profile who agreed to the given privacy terms. “Non-members, however, have no such agreement. We therefore studied their vulnerability to the automatic generation of so-

called shadow profiles”, explains Prof. Dr. Katharina Zweig, who until recently worked at the Interdisciplinary Center for Scientific Computing (IWR) of Heidelberg University.

In an online social network, it is possible to infer information about non-members, for instance by using so-called friend-finder applications. When new Facebook members register, they are asked to make available their full list of e-mail contacts, even of those people who are not Facebook members. “This very basic knowledge of who is acquainted with whom in the social network can be tied to information about who users know outside the network. In turn, this association can be used to deduce a substantial portion of relationships between non-members”, explains Ágnes Horvát, who conducts research at the IWR.

To make their calculations, the Heidelberg researchers used a standard procedure of [machine learning](#) based on network analytical structural properties. As the data needed for the study was not freely obtainable, the researchers worked with anonymised real-world Facebook friendship networks as a test set of basic data. The partitioning between members and non-members was simulated using a broad possible range of models. These partitions were used to validate the study results. Using standard computers the researchers were able to calculate in just a few days which non-members were most likely friends of each other.

The Heidelberg scientists were astonished that all the simulation methods produced the same qualitative result. “Based on realistic assumptions about the percentage of a population that are members of a social network and the probability with which they will upload their e-mail address books, the calculations enabled us to accurately predict 40 percent of the relationships between non-members.” According to Dr. Michael Hanselmann of the HCI, this represents a 20-fold improvement compared to simple guessing.

“Our investigation made clear the potential social networks have for inferring information about non-members. The results are also astonishing because they are based on mere contact data”, emphasises Prof. Hamprecht. Many social network platforms, however, have far more data about users, such as age, income, education, or where they live. Using this data, a corresponding technical infrastructure and other structural properties of network analysis, the researchers believe that the prediction accuracy could be significantly improved. “Overall our project illustrates that we as a society have to come to an understanding about the extent to which relational data about persons who did not provide their consent may be used”, says Prof. Zweig.

The results of the research were published in “*PLoS ONE*”.

More information: *PLoS ONE* 7(4): e34740.
[doi:10.1371/journal.pone.0034740](https://doi.org/10.1371/journal.pone.0034740)

Provided by Heidelberg University

Citation: What online social networks may know about non-members (2012, April 30) retrieved 26 April 2024 from <https://phys.org/news/2012-04-online-social-networks-non-members.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.