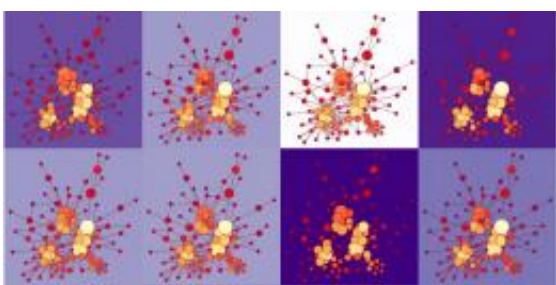


Facebook (and Systems Biologists) Take Note: Network Analysis Reveals True Connections

December 7 2009, By Megan Fellman



A new universal method of network analysis can predict friendships in a social network and protein-protein interactions within a cell. The method also can separate interactions likely to be spurious from those likely to be correct. In the illustration, each square represents a slightly different reconstruction of a protein interaction network. The background color indicates the reliability of the reconstruction, with the darkest color indicating the best representation of the original network.

(PhysOrg.com) -- Facebook figures out that you know Holly, although you haven't seen her in 10 years, because you have four mutual friends -- a good predictor of direct friendship. But sometimes Facebook gets it wrong. "Hey, I don't know Harry!"

Roger Guimera and Marta Sales-Pardo, a husband-wife research team at Northwestern University, have developed a universal method that can accurately analyze a range of complex networks -- including social

networks, protein-protein interactions and air transportation networks. Although the datasets they used were much smaller than Facebook's, the researchers demonstrated the great potential of their method.

Guimera and Sales-Pardo had wondered if one technique, exploiting the fact that all networks have groups in them and those groups are connected in many different ways, could be used to predict both friendships in a social [network](#) and protein-protein interactions within a cell. They applied their mathematical and computational framework to five different networks, ranging from a group of dolphins to a network of neurons, and found one method indeed could reliably analyze all.

The details of their algorithm, which can predict missing and spurious interactions in a system, will be published in the Dec. 7 Early Edition by the [Proceedings of the National Academy of Sciences](#) (*PNAS*).

"The way the flu spreads, for example, is based on an underlying network, and it's important to understand the critical patterns," said Guimera, a research assistant professor of chemical and biological engineering in the McCormick School of Engineering and Applied Science. "Using available data, our method tries to find the best description of the network being analyzed, no matter what kind of network."

In the study, Guimera and Sales-Pardo tested their method on a range of five known "true" networks: a karate club, a social network of dolphins, the neural network of the worm *C. elegans*, the air transportation network in Eastern Europe and the metabolic network of *E. coli*. These networks have between 34 nodes (members of a karate club) and 604 nodes (metabolites in a metabolic network).

"Our method separates wheat from chaff, the signal from the noise," said Sales-Pardo, also a research assistant professor of chemical and

biological engineering. "There are many ways to map nodes in a network, not just one. We consider all the possible ways. By taking the sum of them all, we can identify both missing and spurious connections."

A more accurate method of network analysis could help [Facebook](#), for example, identify truly relevant connections -- with 350 million Facebook users the number of mistakes can add up quickly. Systems biology could benefit, too. The project to obtain a complete map of the millions of human protein-protein interactions has a projected cost of \$1 billion but relies on techniques with accuracies (estimated in 2002) to be below 20 percent.

The central idea behind Guimera and Sales-Pardo's method is that, even though each network has unique characteristics (depending on its functional needs and evolutionary history), all networks share a remarkable property: their nodes can be classified into groups with the nodes connecting to each other depending on their group membership. In a social network, for example, people can be grouped by age, occupation, political orientation and so on. The method proceeds by averaging all possible groupings of the nodes, giving each grouping a weight that reflects its explanatory power.

For each of the five true networks, the researchers introduced errors and applied their algorithm to the distorted network. Each time, the algorithm produced a new network that reliably separated interactions likely to be spurious from those likely to be correct, without the aid of any additional information (such as the type of network or the amount of errors). Each new network reconstruction was closer to the original true network than the network containing errors and omissions.

"The flexibility of our approach, along with its generality and its performance, will make it applicable to many areas where network data reliability is a source of concern," the authors wrote.

The *PNAS* paper is titled "Missing and Spurious Interactions and the Reconstruction of Complex Networks."

Source: Northwestern University ([news](#) : [web](#))

Citation: Facebook (and Systems Biologists) Take Note: Network Analysis Reveals True Connections (2009, December 7) retrieved 24 April 2024 from <https://phys.org/news/2009-12-facebook-biologists-network-analysis-reveals.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.