

Scientists Develop World's Fastest Program to Find Patterns in Social Networks

June 29 2010



(PhysOrg.com) -- As social networks like Facebook, Flickr, Youtube and Twitter increasingly make it possible to access appropriate information within their networks, a whole host of new applications become possible. For individuals, search engines could better differentiate "friends" and suggest groups with more closely matched interests or concerns. Businesses could search allowed information to offer products or services better matched to customers. And national security and counter-terror analysts, with appropriate court authorization, could look for "groups of people within social networks that match certain characteristics.

However, a technical obstacle to all of these is the difficulty inherent in being able to find all parts of the social network that match a given query network pattern. This essential first step (called the "subgraph matching"

step by [computer scientists](#)) is often succeeded by many other application-specific steps. The subgraph matching problem is enormously challenging and has long been known to be computationally very difficult, rising exponentially in complexity with the size of the network increases.

University of Maryland grad student Matthias Broecheler working with Computer Science Professor V.S. Subrahmanian and University of Calabria (Italy) Professor Andrea Pugliese have recently unveiled a new mathematically-based [computer program](#), or [algorithm](#), called [COSI](#) (Cloud Oriented Subgraph Identification) that will support subgraph pattern matching in very large social networks containing hundreds of millions, even billions, of links.

In a paper that has been accepted for presentation at the 2010 Advances in Social Network Analysis and Mining [conference](#) to be held in Denmark in August, Broecheler, Pugliese and Subrahmanian leveraged a key insight - it is possible to split the social network into a set of almost independent, relatively small sub-networks, each of which is stored on a computer in a cloud computing cluster in such a way that the probability that a query pattern will need to access two nodes is kept as small as possible. Using knowledge of past queries and a [complex](#) set of calculations to compute these probabilities, their paper reports algorithms and experiments to answer social network subgraph pattern matching queries on real-world social network data with 778 million edges (which may denote relationships or connections between individuals) in less than one second. More recent results not contained in the paper are able to efficiently answer queries to social network databases containing over a billion edges.

"These new algorithms for subgraph matching make it practical for the first time to implement many desirable functionalities previously only practical for small networks," said Anil Nerode, who is Goldwin-Smith

professor of mathematics and computer science and former director of the Mathematical Sciences Institute at Cornell University. "We can expect a profound influence of these algorithms on extending the capabilities of social networks," said Nerode, who is not involved in the work.

Professor Subrahmanian, one of the inventors, said: "An innovative mix of cloud computing and smart thinking, COSI shows how exact social network pattern matching on complex query patterns can be efficiently implemented. It is a significant advance, not only in answering complex queries over large social networks, but also for answering queries over the Semantic Web. This advance could have a significant impact for individual users of social networks as well as for the national security and business communities and is yet another innovative mix of cloud computing and smart thinking.

"The next challenge for COSI will be to perform matching of similar, but not exact, patterns," he said.

Provided by University of Maryland

Citation: Scientists Develop World's Fastest Program to Find Patterns in Social Networks (2010, June 29) retrieved 19 April 2024 from <https://phys.org/news/2010-06-scientists-world-fastest-patterns-social.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.