

Scaling Graph Computations at Facebook

Johan Ugander
Cornell University
Ithaca, NY
jhu5@cornell.edu

ABSTRACT

With over a billion nodes and hundreds of billions of edges, scalability is at the forefront of concerns when dealing with the Facebook social graph. This talk will focus on two recent advances in graph computations at Facebook. The first focus concerns the development of a novel graph sharding algorithm — Balanced Label Propagation — for load-balancing distributed graph computations. Using Balanced Label Propagation, we were able to reduce by 50% the query time of Facebook’s ‘People You May Know’ service, the realtime distributed system responsible for the feature extraction and ranking of the friends-of-friends of all active Facebook users. The second focus concerns the 2011 computation of the average distance distribution between all active Facebook users. This computation, which produced an average distance of 4.74, was made possible by two recent computational advances: HyperANF, a modern probabilistic algorithm for computing distance distributions, and Layered Label Propagation, a modern compression scheme suited for social graphs. The details of how this computation was coordinated will be described. The talk describes joint work with Lars Backstrom, Paolo Boldi, Marco Rosa, and Sebastiano Vigna.

Categories and Subject Descriptors

G.2.2 [Mathematics of Computing]: Discrete Mathematics—*Graph Theory, Graph Algorithms*

General Terms

Algorithms, Measurement

Keywords

graph partitioning, social networks, degrees of separation

Biography

Johan Ugander is a Ph.D. Candidate in Applied Mathematics at Cornell University, and a frequent research collaborator with the Data Science team at Facebook. His collaborations with Facebook have been published in the Proceedings of the National Academy of Sciences (PNAS), WebSci, WSDM, and WWW. His work has received two Best Paper/Best Student Paper Awards. His research examines large-scale computational problems at the intersection of graph theory, probability theory, optimization, and algorithm design, aiming to better understand the structure of social systems, social networks, and human decision making.