

# Large Graph Mining: Patterns, Cascades, Fraud Detection, and Algorithms

Christos Faloutsos  
Carnegie Mellon University  
Pittsburgh, PA, USA  
christos@cs.cmu.edu

## ABSTRACT

Given a large graph, like who-calls-whom, or who-likes-whom, what behavior is normal and what should be surprising, possibly due to fraudulent activity? How do graphs evolve over time? How does influence/news/viruses propagate, over time? We focus on three topics: (a) anomaly detection in large static graphs (b) patterns and anomalies in large time-evolving graphs and (c) cascades and immunization.

For the first, we present a list of static and temporal laws, including advances patterns like 'eigenspokes'; we show how to use them to spot suspicious activities, in on-line buyer-and-seller settings, in FaceBook, in twitter-like networks. For the second, we show how to handle time-evolving graphs as tensors, how to handle large tensors in map-reduce environments, as well as some discoveries such settings.

For the third, we show that for virus propagation, a single number is enough to characterize the connectivity of graph, and thus we show how to do efficient immunization for almost any type of virus (SIS - no immunity; SIR - lifetime immunity; etc).

We conclude with some open research questions for graph mining.

## Categories and Subject Descriptors

E.0 General

## Keywords

Add your own selection; of keywords or terms here; separated with semi-colons

## Bio

Christos Faloutsos is a Professor at Carnegie Mellon University. He has received the Presidential Young Investigator Award by the National Science Foundation (1989), the Research Contributions Award in ICDM 2006, the SIGKDD Innovations Award (2010), nineteen “best paper” awards (including two “test of time” awards), and four teaching awards. He is an ACM Fellow, he has served as a member of the executive committee of SIGKDD; he has published over 200 refereed articles, 11 book chapters and one monograph. He holds six patents and he has given over 30 tutorials and over 10 invited distinguished lectures.

His research interests include data mining for graphs and streams, fractals, database performance, and indexing for multimedia and bio-informatics data.