Statistical Analysis and Data Mining: The ASA Data Science Journal Explore this journal >

Volume 10, Issue 5
October 2017
Pages 290–311

View issue TOC
**Special Issue:**
CoDA 2016 Special Issue: Selected Papers from the Conference on Data Analysis 2016 – Part III
ORIGINAL ARTICLE

# Context-aided analysis of community evolution in networks

Giuliana Pallotta, Goran Konjevod, Jose Cadena, Phan Nguyen

**First published:**
15 September 2017 Full publication history

**DOI:**
10.1002/sam.11354 View/save citation

**Cited by (CrossRef):**
0 articles Check for updates Citation tools

Funding Information

**Funding Information** U.S Department of Energy, DE-AC52-07NA27344.

## Abstract

We are interested in detecting and analyzing global changes in dynamic networks (networks that evolve with time). More precisely, we consider changes in the activity distribution within the network, in terms of density (ie, edge existence) and intensity (ie, edge weight). Detecting change in local properties, as well as individual measurements or metrics, has been well studied and often reduces to traditional statistical process control. In contrast, detecting change in larger scale structure of the network is more challenging and not as well understood. We address this problem by proposing a framework for detecting change in network structure based on separate pieces: a probabilistic model for partitioning nodes by their behavior, a label-unswitching heuristic, and an approach to change detection for sequences of complex objects. We examine the performance of one instantiation of such a framework using mostly previously available pieces. The dataset we use for these investigations is the publicly available New York City Taxi and Limousine Commission dataset covering all taxi trips in New York City since 2009. Using it, we investigate the evolution of an ensemble of networks under different spatiotemporal resolutions. We identify the community structure by fitting a weighted stochastic block model. We offer insights on different node ranking and clustering methods, their ability to capture the rhythm of life in the Big Apple, and their potential usefulness in highlighting changes in the underlying network structure.