Network Structure and Influence

A. Yazdiha[1] B. Donkers[2]

[1]Erasmus School of Economics
[2]Erasmus School of Economics

Department Lunch Meeting, 2018

Outline

Outline

1. Motivation
   - Introduction
   - Findings from Email data
2. Community Detection
   - Our View of the Problem
3. 1st Project: Community Detection
   - Community Detection
   - About the Data
   - Some Preliminary Findings
4. 2nd Project: Joint Model of Behavior and Structure
   - Resolution
5. 3rd Project: Network Evolution
   - Dynamic Evolution of Behavior and Structure

Networks and Influentials

- This is a very well known problem in networks and marketing in general
- Who is an influential?
- Abundant literature on identifying influence and detecting opinion leaders
- Among these the role of network metrics has been quite common

Previous Research

- A lot of attention is given to testing influence w.r.t individual positions at network level
  - i.e. Network centrality:degree, betweenness,closeness, prestige, etc
- Almost all of the previous works deals with the global view of the network
  - Watts & Dodds 2007, Goldenberg et al 2009, etc

Outline

Knowing the Ground Truth Communities(SNAP - Leskovec & Krevl, 2014)

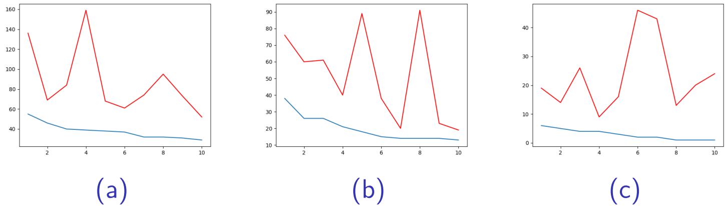


Figure: Degree centrality for the ground truth communities(blue line), and degree centrality for the same nodes at the network level(red line)

Outline

About the Network Structure

- Homophily: birds of a feather flock together
- Result: Networks end up with communities, where connections are:
  - Denser within vs. across
- As a matter of fact, networks are a multiplex of smaller sub-networks
- Communities as a signal for domains suggest a finer grained opinion leadership

Application of Community Detection in SNs

- Gain insights at the community level about individual positions rather than at the global network level, (Project 1)
- Allows for a good control when identifying influence by disentangling selection from causality, (Project 2)
- Enables dynamic monitoring of network and behavior evolution in a time-stamped data, (Project 3)

Outline

Big Picture

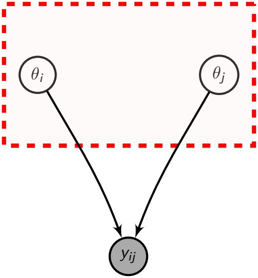


Figure: DGP for Networks

Generative Story of the Network Structure

- Individuals belong to a set of communities to different extents
  - membership probability vector
- Individuals take up different roles(activate communities) depending on whom they interact with
  - interaction based community indicator
- Individuals form a link to each other based on the chance of communication between announced communities
- This is mainly a mixed membership stochstic blockmodel, that allows for overlapping communities.

MMSB model Graphical Model

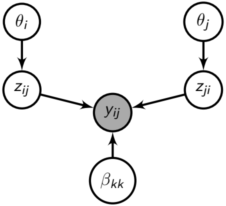


Figure: MMSB model

Inference Engine

- For all models we use Stochastic Variational Inference to get back paramaters
- Variational Inference(VI)
  - Approximating posterior through an easier tractable proposal distribution by minimizing the KL divergence between true posterior and proposal
- Stochastic Variational Inference(SVI)
  - Using Minibatch sampling to have a noisy estimate of the gradient with decreasing step size to optimize the variational objective
  - allows handling large datasets
- Speed up on 4 levels:
  - Variational engine
  - Stochastic minibatch sampling
  - Set informative update of parameters
  - Parallel implementation of batch updates(ongoing)

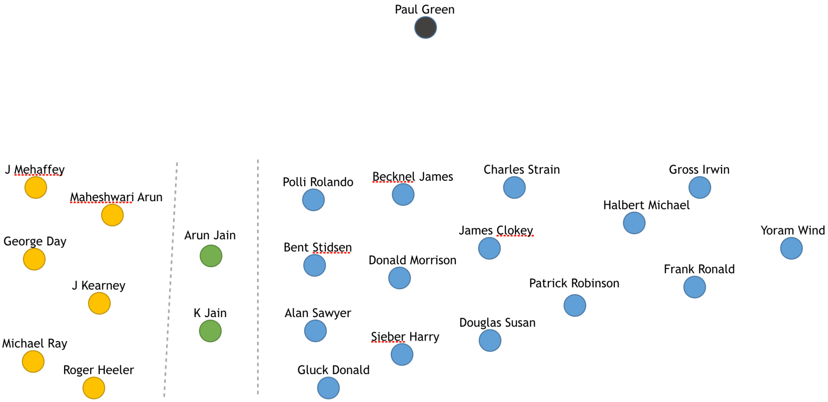
Outline

Co-authorship network (From Mkt Sci Database subimssions)

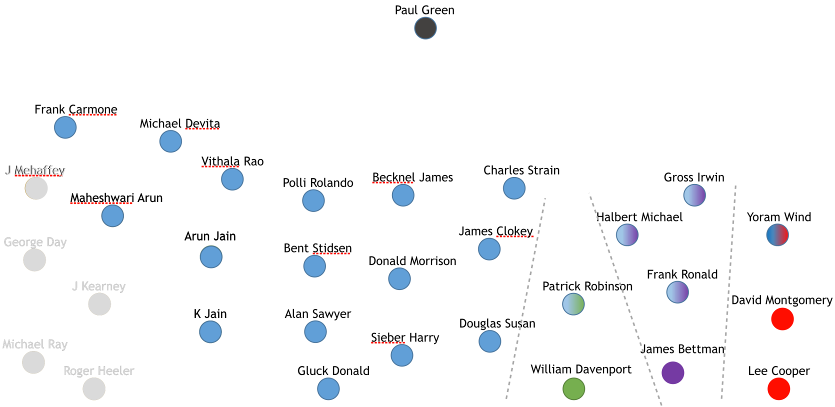
- A Network of authors in marketing that publish together(from 1973 to 2009)
  - authors are the nodes
  - co-authorships are edges
- including JM, JMR,JCR, J of Mark. Sci, etc
- more than 50000 authors

Outline

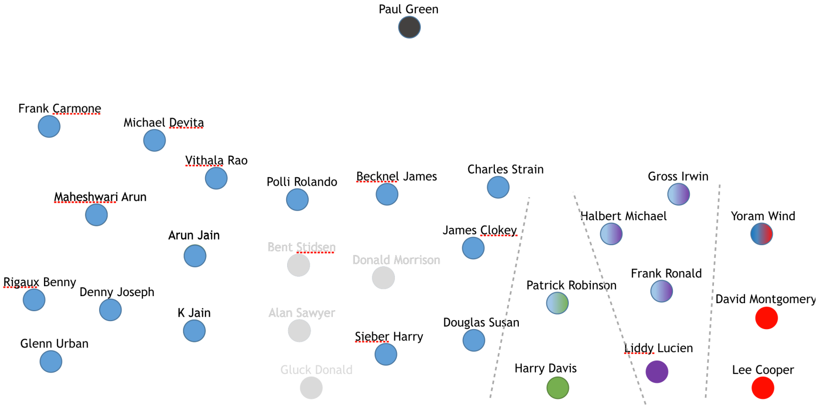
Co-authorship network(cont’d)



Co-authorship network(cont'd)



Co-authorship network(cont'd)



Motivation

- On top of communities, behavioral data can help us with better labels
- We can redefine opinion leadership as a function of community roles
- "Latent" Homophily if not accounted for, still can be a problem in identifying social influence

Big Picture

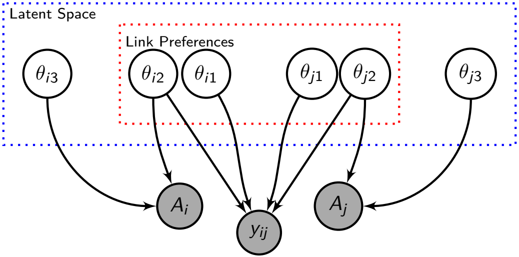


Figure: Graphical representation of network structure and behavior

When there is Social Influence

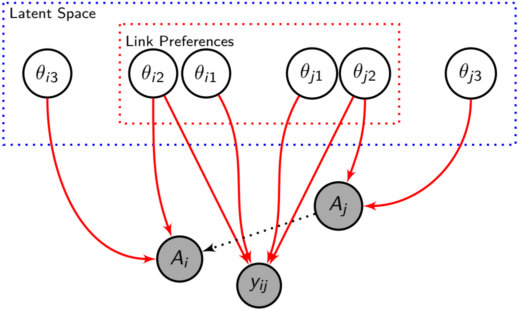


Figure: Graphical representation of network structure and behavior with influence-red signifies the open confounding path

Outline

Our Resolution

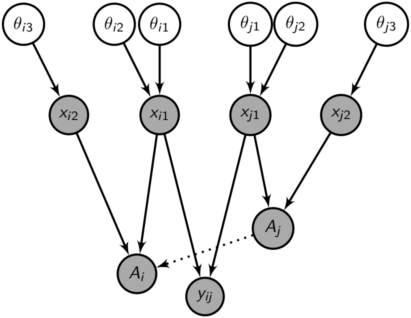


Figure: Solution to the confounding path

Outline

Motivation

- Static representation of networks can be naive for many networks
- ties form and severe
- As a result of endogenous network formation,
  - new behaviors may emerge
  - some behaviors may be discontinued

Big Picture

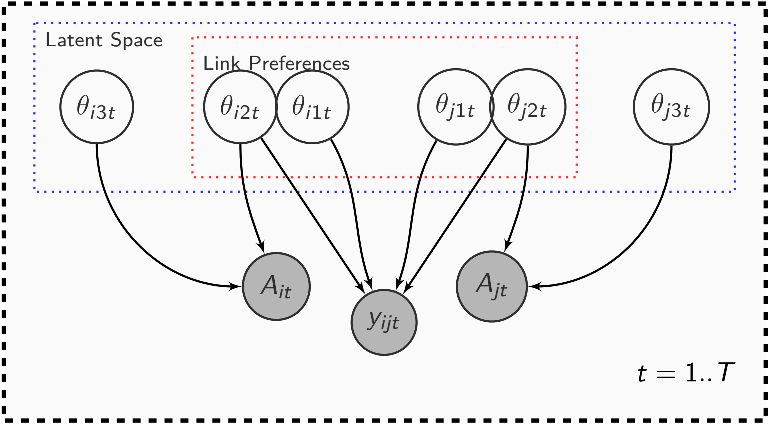


Figure: Dynamic evolution of network and behavior

Questions and Feedbakcs

- Directions to be focused more or less
- Positioning
- Job market suggestions