MATH 100 Topic

Analyzing Your Social Network Data

Lulu Kang

E1-105B, lkang2@math.iit.edu

1. Introduction

With the growth of websites such as Facebook, Google Plus and LinkedIn, social networking is enjoying an increasingly large public profile. There is also an increasing amount of time and money being spent by researchers in statistics, computer science, mathematics and even biology on the collection and analysis of network data.

Network data is data on the interactions between things (whether people, companies, websites, proteins or whatever), rather than on the things themselves. We refer to these interacting things as nodes and the interactions as links. Most commonly, these links are binary measurements; either there is a link or there is not between any two nodes. So, in the Facebook example, you are either friends on Facebook or you are not.

2. Examples

1. Facebook social network: "Analyzing my facebook friends", Significance Magazine, Vol 9, Issue 4, 2012.
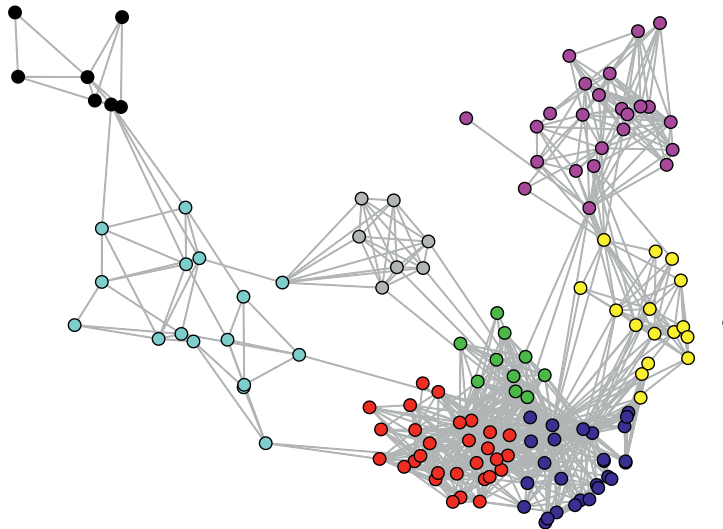

Figure 1. The author's Facebook network. A statistical analysis of the link pattern (grey lines) reveals information on the Facebook friends (circles). The algorithm colours the circles; examination of the colouring reveals that the friend groupings are easily interpretable

"The groups in my network are school/home-place (magenta), my family (yellow), college halls (red), college ski trip (green), my fiancée's friends and family (navy), Ireland-based statisticians (cyan), social network statisticians (black) and Norwegians (grey)."

2. LinkedIn Professional Network

http://inmaps.linkedinlabs.com

3. Lists of Tasks

1. Acquiring Data

- Facebook Network data: apps NameGenWeb and NameGenDev
  http://cyborganthropology.com/How_to_Download_Facebook_a_Social_Graph_with_NameGenWeb
  (Note that these methods used to be working, but now seem to be disabled by facebook. In the worst scenario, you can manually input the data.)
- Google Plus follower network:
  http://mashe.hawksey.info/2011/12/google-plus-network-info-to-nodexl
- LinkedIn Professional Network
- Others

2. Analyzing Data:

- Knowledge Base: clustering modeling, or some other statistical modeling techniques.
- Software: R package VBLPCM, etc.

3. Conclusions:

- Basic: finding the structure of your social networks.
- More Advanced Goal: making some statements on the behaviors of your friends.

Reference:

1. Salter-Townshend M. (2012) Analyzing my facebook friends. *Significance*, **9**(4), 40-42.

2. Christakis, N. A. and Fowler, J. H. (2007) The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, **357**(4), 370-379.
3. Handcock, M. S., Raftery, A. E. and Tantrum, J. M. (2007) Model-based clustering for social networks. *Journal of the Royal Statistical Society, Series A*, **170**(2), 1-22.
4. Salter-Townshend, M. (2012) *VBLPCM: Variational Bayes Latent Position Cluster Model for Networks*, R package version 2.0.