Community Structures

Erika Fille T. Legara, Ph.D.  
Scientist, IHPC-A*STAR  
@eflegara | http://eflegara.github.io

I P[y]:  
IPython  
http://tiny.cc/commdet2017  
http://tinyurl.com/commdet2017
Walkthrough

• Network structure and function
• Communities
• Community detection algorithms
  • Divisive
  • Agglomerative
• Application: Community detection in framing analysis
• Hands-on  http://tiny.cc/commdet2017
  http://tinyurl.com/commdet2017
Network Structure

- Path length
- Diameter
- Clustering coefficient
- Degree distribution
- Centrality measures

What would the behaviour of networked systems be based on their structural properties?
Communities

• Cluster, module, group
• A group of nodes that have a higher likelihood of connecting to each other than to nodes from other communities
Example of Communities

- World Wide Web
  - Communities: Sites on related topics

- Friendship network
  - Communities: Group formation among people

- Disease-disease relationships
  - Communities: Functional modules

- Collaboration network
  - Communities: Research fields
Social Networks: Politics, Ideologies

The political blogosphere and the 2004 U.S. election: divided they blog
by Lada Adamic, Natalie Glance
Biological Network: Ageing

Health and Disease Phenotyping in Old Age Using a Cluster Network Analysis
JF Valenzuela, C Monterola, JC Tong, T.P. Ng, and A. Larbie
How does one find the Communities?

```python
import community
partition = community.best_partition(G)
```
How do I find the Communities?

Community Structures | EF Legara | 2017 NTU Winter School on Complexity Science
How does one find the Communities?

```python
import community
partition = community.best_partition(G)
```
Zachary’s karate club

34 members of a karate club
78 pairwise links between members who interacted outside the club

Community Structures | EF Legara | 2017 NTU Winter School on Complexity Science
Methods

**Agglomerative**

- Cut 3 classes
- Cut 4 classes
- Cut 6 classes

**Divisive**
Modularity

- Measure the quality of a partition

\[ M_c = \sum_{c=1}^{n_c} \left[ \frac{L_c}{L} - \left( \frac{k_c}{2L} \right)^2 \right] \]

“the fraction of edges that fall within communities, minus the expected value of the same quantity if edges fall at random without regard for the community structure”
Modularity

\[ M_c = \sum_{c=1}^{n_c} \left[ \frac{L_c}{L} - \left( \frac{k_c}{2L} \right)^2 \right] \]

- Number of communities
- Total number of links in Community c
- Total node degrees in Community c
- Total number of links
Modularity

\[ M_c = \sum_{c=1}^{n_c} \left( \frac{L_c}{L} - \left( \frac{k_c}{2L} \right)^2 \right) \]

Community 1:
\[ \frac{6}{20} - \left( \frac{14}{40} \right)^2 \]

Community 2:
\[ \frac{7}{20} - \left( \frac{16}{40} \right)^2 \]

Community 3:
\[ \frac{4}{20} - \left( \frac{10}{40} \right)^2 \]

\[ M_c = 0.50 \]
Modularity

\[
M_c = \sum_{c=1}^{n_c} \left[ \frac{L_c}{L} - \left( \frac{k_c}{2L} \right)^2 \right]
\]

Number of communities: \( n_c \)
Total number of links in Community \( c \): \( L_c \)
Total node degrees in Community \( c \): \( k_c \)
Total number of links: \( L \)

Community 1:
\[
\left[ \frac{2}{20} - \left( \frac{8}{40} \right)^2 \right]
\]

Community 2:
\[
\left[ \frac{9}{20} - \left( \frac{22}{40} \right)^2 \right]
\]

Community 3:
\[
\left[ \frac{4}{20} - \left( \frac{10}{40} \right)^2 \right]
\]

\( M_c = 0.35 \)
Modularity

\[ M_c = \sum_{c=1}^{n_c} \left[ \frac{L_c}{L} - \left( \frac{k_c}{2L} \right)^2 \right] \]

- Number of communities
- Total number of links in Community c
- Total node degrees in Community c
- Total number of links

\[ M = 0 \]
Modularity

\[ M_c = \sum_{c=1}^{n_c} \left[ \frac{L_c}{L} - \left( \frac{k_c}{2L} \right)^2 \right] \]
Quantifying Community Structures

Divisive
- Girvan-Newman Algorithm (edge betweenness)

Agglomerative
- Hierarchical clustering (distance matrix)
- Louvain Method (modularity-maximization)
Community Detection in Framing Analysis

A Network of Words

Framing - the “central organizing idea for making sense of an issue or conflict and suggesting what is at stake”

Framing in the News

- 346 news articles (randomly sampled)
- Population, family planning, reproductive health, and contraception published between 1987 and 2007
Every article assigned to a single dominant frame

- Population and Development
- Family Planning as Conflict between Government and Church
- Women’s and Reproductive Health
- Population Management Threatens Morals and Values
- Population Growth and Demographic Trends
- Others

Levels of discourse in the RH debate

By: Fr. Joaquin G. Bernas S. J.
02:26 AM May 30th, 2011

The debate on the RH bill appears often frustrating and sometimes verging on the chaotic. It is largely because the participants in the debate frequently communicate on different levels of discourse thereby evading real engagement. Many rarely make the necessary distinctions, nuances and clarifications. The result can be like listening to people arguing on different radio frequencies.

But this is largely inevitable for at least two reasons: first, the disagreements can legitimately be on different topics and, second, there is no director who can manage the ordering of topics for discussion. Thus the debate can mix basic constitutional issues, family planning, population control and national development issues, drug control or regulation and even criminal process in a hodgepodge of confusion.

The debate on the constitutional level alone already offers abundant material for levels of disagreement. The core issues arise from the non-establishment clause and the free exercise clause.

The non-establishment clause, popularly but not always precisely referred to as separation of church and state, can mean different things for different people. The core meaning is that it prohibits the establishment of a state religion. Historically for the Philippines, it means the denial to the Catholic church of the privileged position it occupied under Spanish sovereignty. Corollary to the cutting down of the privileged position of the Catholic church has been the recognition of the equal position of other religions.

Beyond the prohibition of a state religion, non-establishment also means the prohibition of the use of public resources for the support or for the prohibition of religion. But public resources may be used for a legitimate secular purpose even if incidental benefit to religion might arise. The use of public money, for instance, for making safe contraceptive devices available to the poor falls under this aspect of the non-establishment clause.
A Network of *Words*

Collected Articles

Delete stop words and replace with placeholder “xxx”

Generalize conceptually similar terms

Obtain high-frequency words

family_planning advocacy xxx xxx xxx world xxx xxx campaign
xxx enemy territory xxx xxx xxx xxx five-day meeting xxx
xxx xxx roman catholic_church xxx.

issues xxx artificial contraception xxx genital mutilation xxx wife
beating xxx legalize abortion xxx expect xxx xxx talk
xxx nation xxx xxx london-based
international_planned_parent_hood_federation (ippf).

xxx xxx xxx xxx xxx xxx xxx home nation xxx xxx' xxx oldest roman
catholic_church, xxx xxx xxx xxx xxx fight xxx xxx death xxx
xxx xxx promote artificial contraception xxx abortion.

prolife xxx xxx xxx picket xxx conference xxx xxx
large non-government promote xxx abortion, contraception,
sterilization xxx liberal sex_education xxx xxx xxx.

xxx accuse conference xxx xxx xxx xxx xxx xxx
liberal xxx abortion law xxx xxx xxx.'

ippf xxx xxx manila xxx xxx ideal xxx xxx philosophy,
xxx xxx check xxx explosion xxx xxx xxx' xxx population xxx ensure
xxx health xxx woman.

xxx choice xxx xxx xxx xxx venue xxx xxx xxx xxx xxx xxx
xxx xxx government xxx xxx xxx xxx xxx beyond xxx population
debate xxx xxx xxx xxx xxx xxx xxx xxx problem xxx enforce
population xxx development program," xxx ippf secretary_general
ingar brueggemann.
Map words: add edge weight (1u) between nodes if words co-exist in a scanning window

Create co-occurrence network

Perform cluster analysis
F1: Population and Development
F2: Population Growth and Demographic Trends
F3: Women’s and Reproductive Health
F4: Family Planning as Conflict between Government and Church
F5: Population Management Threatens Morals and Values
F6: Others
A Brief Introduction to Clustering and Community Detection

Author: Erika Filie Legara

The MIT License (MIT)

Copyright (c) 2016 Erika Filie Legara

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Description

Real-world networks reveal clustering behaviour, which is exhibited in the formation of communities/clusters/partitions (used interchangeably in this notebook) in the graph structure. When studying networks, their structure and function, it is therefore crucial to identify these groupings. The concept of clustering and community detection oftentimes go hand-in-hand, and also with other concepts such as cliques and connected components, to name a few. In this notebook, we explore these different structures with a strong focus on clustering and community detection.

Note: This notebook is a supplementary material to Michael Lees and Debraj Roy’s A Short Practical Introduction to NetworkX for the 2016 NTU Winter School on Complex Systems.

http://tiny.cc/commdet2017
http://tinyurl.com/commdet2017