Application of social network analysis to HIV/AIDS research

January 22, 2016

Kayo Fujimoto, PhD
Assistant Professor
Department of Health Promotion & Behavioral Sciences
School of Public Health, University of Texas Health Science Center at Houston
Outline of today’s presentation

1. **Introduction to Social Network Analysis (SNA)**
   - Concept: actors, relations, and mode
   - Network data: stored, presented, measured

2. **Social Network Methods**
   - Network exposure model
     - Modeling social influence
   - Exponential Random Graph Models (ERGM)
     - Identifying structural characteristics (statistical inference)

3. **Empirical application to health behavioral research**

4. **Ongoing projects (NIH R01, R21)**
   - YMAP (“Young Men’s Affiliation Project”)
   - iMAN (“integrated Molecular and Affiliation Network analysis”)
   - BARS (“Building Agent-based models for a Racialized criminal-justice System”)


Introduction to Social network analysis

Basic definitions and concepts
Utilities of Social Network Analysis

Social network analysis permits us to:

Understand patterns of social relationships

1) Visualizing patterns
2) Describing patterns
3) Making statistical inference about how relations are patterned
The basic idea

- **Social network** is composed of:
  - A set of actors (nodes):
    - individuals, organizations, venues, students, etc.
  - Relations (ties) with one another:
    - friendship, advice-seeking, discussion, drug-use, sexual/drug-injecting relationship
- **Actors** have characteristics:
  - **Socio-demographic**: gender, race/ethnicity, age
  - **Health/Disease status**: obesity level, HIV seropositivity
  - **Risk behaviours**: substance use, risky sexual behavior
Any networks are composed of:

- A distinct set of entities → “Mode”
- **One-mode** network → a single set of actors
  - A set of students, a set of organizations, etc.
- **Two-mode** network → two distinct sets of actors
  - 1\textsuperscript{st} mode → a set of students
  - 2\textsuperscript{nd} mode can be:
    - A set of events attended
    - Club membership (club, company board)
    - Crowd Identity (jocks, nerd, popular, etc)
    - Intervention sessions
Handling social network data

• Social network data:
  • Recorded by matrix representation
  • Visualized by a graph
Example of one-mode network (7 students with friendship ties)

Matrix form:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Σ</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

Graph:

- Popular
- Isolated

Indegree = popularity (summing across rows for each column)
Example of two-mode network

- Convert **two-mode network** into **one-mode network** *(projection)*
  - Each pair of actors are connected if they share at least one common event.

Two-mode graph  
Projected one-mode graph
Projected actor-by-actor matrix

\[ AA^T = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} = C \]
Social network methods

Network Exposure Model (Valente, 1995; 2005; Burt, 1987)
1) **Diffusion of innovation**: examines *how new ideas and behaviors are spread through a network*

1) **Network exposure model** provides us with a method of *measuring the level of exposure to other members who adopted innovation*
Graphical example of computing the level of exposure to friends’ drinkers

Non-drinker

Drinker

Alcohol Exposure=33%

Alcohol Exposure=66%

Alcohol Exposure=100%

Source: Valente (1995)
Network Exposure Model
(Valente, 1995, 2005)

\[
E = \frac{\sum_{j} w_{ij} y_j}{\sum_{i} w_i}
\]

- \(W\) is the friendship matrix
- \(y\) is vector of friend’s drinking

Example of computing exposure to friend drinkers \(E_i\)

\[
\begin{bmatrix}
0 & 1 & 0 & 0 & \cdots & 0 \\
1 & 0 & 1 & 0 & \cdots & 0 \\
1 & 0 & 0 & 1 & \cdots & 1 \\
0 & 0 & 0 & 0 & \cdots & 1 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 1 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
1 \\
0 \\
1 \\
0 \\
\vdots \\
0
\end{bmatrix}
\]

\[
= \begin{bmatrix}
4 \\
5 \\
5 \\
3 \\
\vdots \\
4
\end{bmatrix}
\]

\[
E_i = \text{proportion of friends who drink alcohol}
\]
Affiliation Exposure Model  
(Fujimoto & Valente, 2011, 2012)

\[
F = \frac{\sum_{j} c_{ij} y_j}{\sum_{i} c_i}
\]

- **C** is the co-membership matrix, \(C = A^*A^T\)
- **y** is vector of *alters’ behavior*

\(F_i = \text{proportion of activities an actor co-participated with drinker(s)}\)

Two-mode graph  

Projected one-mode graph
Regression model
With Affiliation Exposure Term

\[ y = \rho W y + X \beta + \gamma D + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I) \]
Study that applies affiliation exposure models to empirical data

Study 1: Alcohol peer influence and adolescent alcohol use

Fujimoto & Valente (2013) *Health Psychology*
Background: Peer influence in SNA

- Peer influence is based on friendship network
  - Friends’ influence is measured by:
    - i.e., the extent to which an adolescent is exposed to friends who use substance
    - Analyze its association with self use

🔗 Limited to using directly connected ties
Background: Peer influence in SNA

- Operationalize “affiliation-based peer influence”
  - through jointly participating in organized activities with drinkers
- Used Affiliation Exposure Model
- Tested the effect of affiliation exposure on self-drinking (outcome variable)
  - by controlling for friends’ influence (i.e., having friends who drink alcohol)
## Adjusted Odds Ratio of Logit Models

<table>
<thead>
<tr>
<th>Logistic regression model</th>
<th>Drinking</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affiliation exposure</strong> based on sports activities</td>
<td>1.23* (0.12)</td>
</tr>
<tr>
<td><strong>Number of sports participated in</strong></td>
<td>0.99 (0.03)</td>
</tr>
<tr>
<td>Exposure to <strong>Friends’ drinkers</strong></td>
<td>1.42*** (0.04)</td>
</tr>
</tbody>
</table>
Utilities of Affiliation Exposure Model

• Only require the collection of information about affiliation/participation in specific activities/events
  • Simply include unique activities (events) in a survey
    • Alcoholic anonymous
    • Online community membership
    • Intervention sessions
Exponential Random Graph Modeling (ERGM)

1) **Statistical method** of directly modeling observed network

2) **Identify** relational characteristics in the observed network

   1) By assessing if these **characteristics** are more likely to be observed in the network than randomly generated networks
Empirical application of ERGM to HIV/AIDS research

• Study 2: Venue-mediated weak ties in multiplex HIV transmission risk networks among drug-using male sex workers and associates

Study 2: Data and Aim

- **Data:**
  - 376 male sex workers and their associates (incl. men who have sex with me (MSM), female sex workers/non-sex workers)

- **Aim:** identify relational characteristics of HIV risk networks by modeling:
  - Multiple relationships of two types
    - One-mode network of sexual/drug-using/social relationships
    - Two mode network of venue affiliation (66 bars & street intersections)
Study 2: Method

1) Visualized the pattern of HIV risk network
   1) Sex/drug-use/social relationships (N=376)
   2) Venue affiliation (66 bars & street intersections)

2) Modeled observed network with respect to:
   1) Socio-demographic characteristics (race/ethnicity, age, homeless, etc.)
   2) Risky sexual behavior (unprotected sexual behavior, etc.)
   3) HIV status

   ➢ Used exponential random graph models (ERGMs)
HIV transmission network among 378 MSWs & associates with venue affiliation
Network concept: Strength of weak tie theory

**Weak ties** facilitate the **diffusion of information** (Granovetter, 1983)

- People obtain **diverse information** and avoid **information redundancy**

**Bar-based opinion leader HIV prevention** (Kelly, et al, 1991)

- **Venues** facilitate the **diffusion of information** (HIV prevention messages)
Hypothesis 1: Weak ties are more likely than strong ties to be observed linked to venues.
Study findings

1) Venue-mediated weak ties supported
   - Individuals formed weak ties by affiliating with the same venue.

2) Reciprocity
   - Individuals mutually connected to their partners, indicative of close relationships.

3) Homophily
   - Individuals choose their partners based on similarity in HIV-positive status, age, race/ethnicity, homelessness, bisexual orientation, the number of casual sex partners, and protective sex behavior.
Implications for network intervention

1) Importance of the **weak ties** suggest that **venues** could be promising **settings for diffusing HIV prevention messages**.

1) **Multiple characteristics** (HIV status, age, race/ethnicity), rather than a single factor, may identify an **individual** as a **peer**.
YMAP: Young Men’s Affiliation Project of HIV Risk and Prevention Venue
NIH/NIMH 1R01MH100021
(PI: Fujimoto, K. & Schneider, J.A.)
2013-2018
Background

• Younger Men who have Sex with Men (MSM) have the highest risk of HIV infection in the U.S.

• Disease incidence and risk behaviors are associated with where YMSM congregate
  – Social/risk venues
  – Preventive venues

• Imply->risk networks are formed through venue attendance pattern
Description of YMAP

Employs social network analysis to study HIV transmission networks among YMSM aged 17-29.
A multisite study conducted in two major cities

1) Houston – UT School of Public Health

2) Chicago

More **racially segregated** with distinct areas composed of:

1. Minority MSM on the **South** and **West** sides – University of Chicago

2. White MSM on the **North side** (Boys Town) – Lurie Children’s Hospital of Chicago (Northwestern University)
Study design—Phase I

1) Interview with venue representatives: sample of social/risk & preventive venues
   - 85 venues in Houston & 65 venues in Chicago
     1. Wave 1 (2014) ← report today
     3. Wave 3 (2016)
Venue attribute: 3 types

• **Health venue**: Physical venue that has a **health** focus

• **Risk venue**: Physical venue with a risk for **sexual or drug activity** (i.e., bars, clubs, bath houses, cruising spots, porn theaters).

• **Social/other venue**: Any **other physical venue**. May be neutral (e.g., retail store) or social (e.g., restaurant, café).
3 different types of relationships

- **Collaboration**: Worked together on any activity, project, or event for a shared goal or purpose, **formally** or **informally**.

- **Competition**: competed for resources such as revenue, clientele or membership, employees, or other resources.

- **Sponsoring relation**: financial support, grant funding
Next step for venue networks

• Model **multiplex networks structures** by conducting **ERGM analysis** of:
  • **Collaboration**
  • **Competitions**
  • **Funding relations**
  • in relation to **venue type, media use, physical proximity, venue size, common funders**, etc.

• Translate results into **network intervention program**
Multiple networks of competition and collaboration

- Model multiplex networks structures by conducting ERGM analysis of:
  - Collaboration
  - Competitions
  - Funding relations
- in relation to venue type, media use, physical proximity, venue size, common funders, etc.

- Translate results into network intervention program.

Houston

Chicago
Multiple networks of common funders and funding types

- Model multiplex network structures by conducting ERGM analysis of:
  - Collaboration
  - Competitions
  - Funding relations
  - in relation to venue type, media use, physical proximity, venue size, common funders, etc.

- Translate results into network intervention program.
Network configurations for ERGM specifications

H1a: competition through same venue type

H1b: competition through geographic proximity

H1c: competition through shared social media use

H1d: competition through shared funding sources

H3: entrainment through shared funding sources

Legend
- Venues
- Same venue type
- Social media
- Funders/types

- Competition ties
- Collaboration ties
- Geographic distance
Study findings

1) **Risk venues** tend to compete one another **only for Houston**

2) **Venues with shared media use** tend to compete one another **only for Chicago**

3) **Venues with shared funder** tend to compete one another **for both cities**

4) Competing venues **are less likely to collaborate** when they **share the same funders/funding types**
Implications for network intervention

1) Change funding structures:
   - To mobilize the community by collective efforts to better serve the local MSM community

2) Implementation of a common funding program:
   - Encourage all grantees to have partnership with otherwise competing grantees
1) Interview with YMSM: Sample of younger MSM (17-29)

- 400 YMSM in Houston & 400 YMSM in Chicago
  2. Wave 2 (2015-16)
YMSM’s affiliation with risk vs health venues
Venn diagram of co-attendance of risk/health venues with peers
Independent variables

○ Affiliation exposure to:
  ○ Co-attendees with high viral load through health and risk venues
  ○ Co-attendees with high viral load through health venues only
  ○ Co-attendees with high viral load through risk venues only
Outcome variables

- Viral load level (RNA)
  1) 0-Seronegative: 182 (62%)
  2) 1-Seropositive AND VL undetectable: 65 (22%)
  3) 2-VL > 400: 48 (16%)
Multinomial logistic regression

(1) \( g_1(x) = \log \left( \frac{P(Y=Undetectable|x)}{P(Y=Seronegative|x)} \right) = x' \beta_1 \)

(2) \( g_2(x) = \log \left( \frac{P(Y=Detectable|x)}{P(Y=Seronegative|x)} \right) = x' \beta_2 \)
\[ g_1(x) = \log \left( \frac{P(Y = \text{Undetectable} | x)}{P(Y = \text{Seronegative} | x)} \right) \]

\[ g_2(x) = \log \left( \frac{P(Y = \text{Detectable} | x)}{P(Y = \text{Seronegative} | x)} \right) \]
<table>
<thead>
<tr>
<th>Multinomial regression model</th>
<th>Undetectable vs seronegative</th>
<th>Detectable vs seronegative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk and health venue co-attendance</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Only risk venue co-attendance</td>
<td>--</td>
<td>5.13*</td>
</tr>
<tr>
<td>Only health venue co-attendance</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td># of risk and health venue attended</td>
<td>.86***</td>
<td>.90*</td>
</tr>
<tr>
<td># of risk venue attendance</td>
<td></td>
<td></td>
</tr>
<tr>
<td># of health venue attendance</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Extension of YMAP to other projects

**YMAP**
YOUNG MEN’S AFFILIATION PROJECT

**BARS (NIDA R01)**
Building Agent-based models for Racialized criminal-justice system

**iMAN (NIGMS R21)**
integrated Molecular and Affiliation Network analysis

- Argonne Nat. Lab., IL
- Athens, Greece
- Chicago, IL
- Houston
- Miami, FL
- UCLA, CA
Thank you for your attention

Any questions?

Contact: Kayo.Fujimoto@uth.tmc.edu
Collaboration network

Houston (n=66)
- VH1011 (22)
- VH1014 (21)
- VH1001 (17)
- VH1007 (13)
- VH1018 (12)
- VH1012 (11)
- VH1036 (11)
- VH1108 (11)
- VH1116 (10)
- VH1022 (9)

Risk Classification
- Red: Risk
- Blue: Protective
- Gray: Neither

Cent’n=0.4

Chicago (n=56)
- VC1003 (21)
- VC1005 (17)
- VC1148 (12)
- VC1006 (9)
- VC1145 (9)
- VC1014 (7)
- VC1016 (7)
- VC1041 (7)
- VC1144 (7)
- VC1151 (7)

Cent’n=0.15
Sponsor networks

Houston (n=70) Chicago (n=56)