CHIPTS Method Seminar, UCLA

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

UTHealth | School of Public Health
The 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 behavioures**: 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

✓1st mode → a set of students

✓2nd 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**

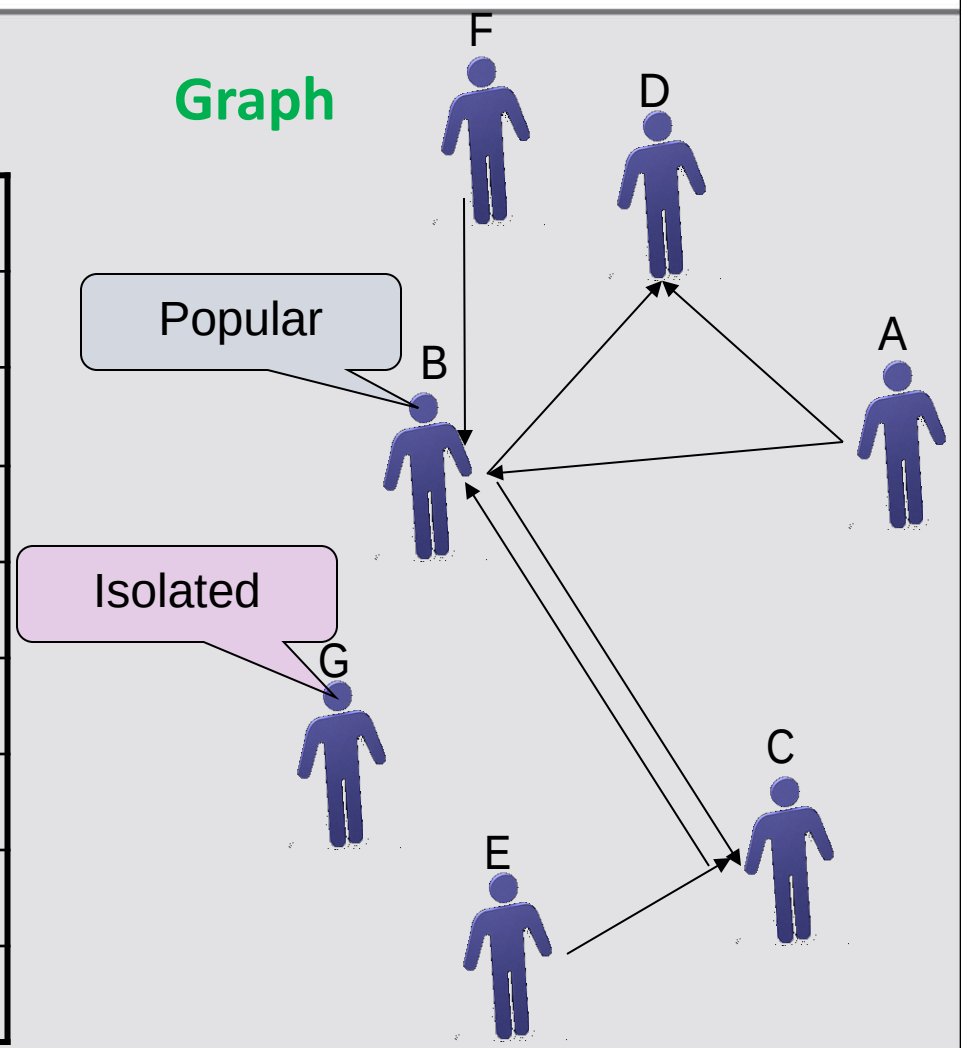
Example of one-mode network (7 students with friendship ties)

Matrix form

	A	B	C	D	E	F	G	∑
A	-	1	0	1	0	0	0	2
B	0	-	1	1	0	0	0	2
C	0	1	-	0	0	0	0	1
D	0	0	0	-	0	0	0	0
E	0	0	1	0	-	0	0	1
F	0	1	0	0	0	-	0	1
G	0	0	0	0	0	0	-	0
∑	0	3	2	2	0	0	0	7

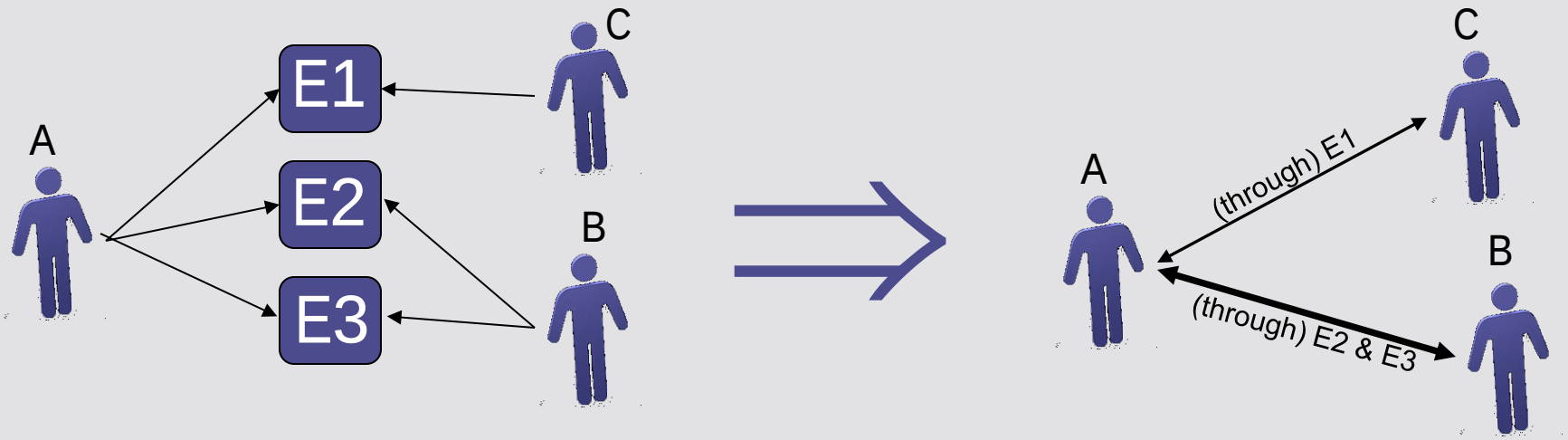
Indegree = popularity (summing across rows for each column)

Graph



Example of two-mode network

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



Two-mode graph

Projected one-mode graph

Projected actor-by-actor matrix

$$AA^T = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \underline{\underline{C}}$$

Social network methods

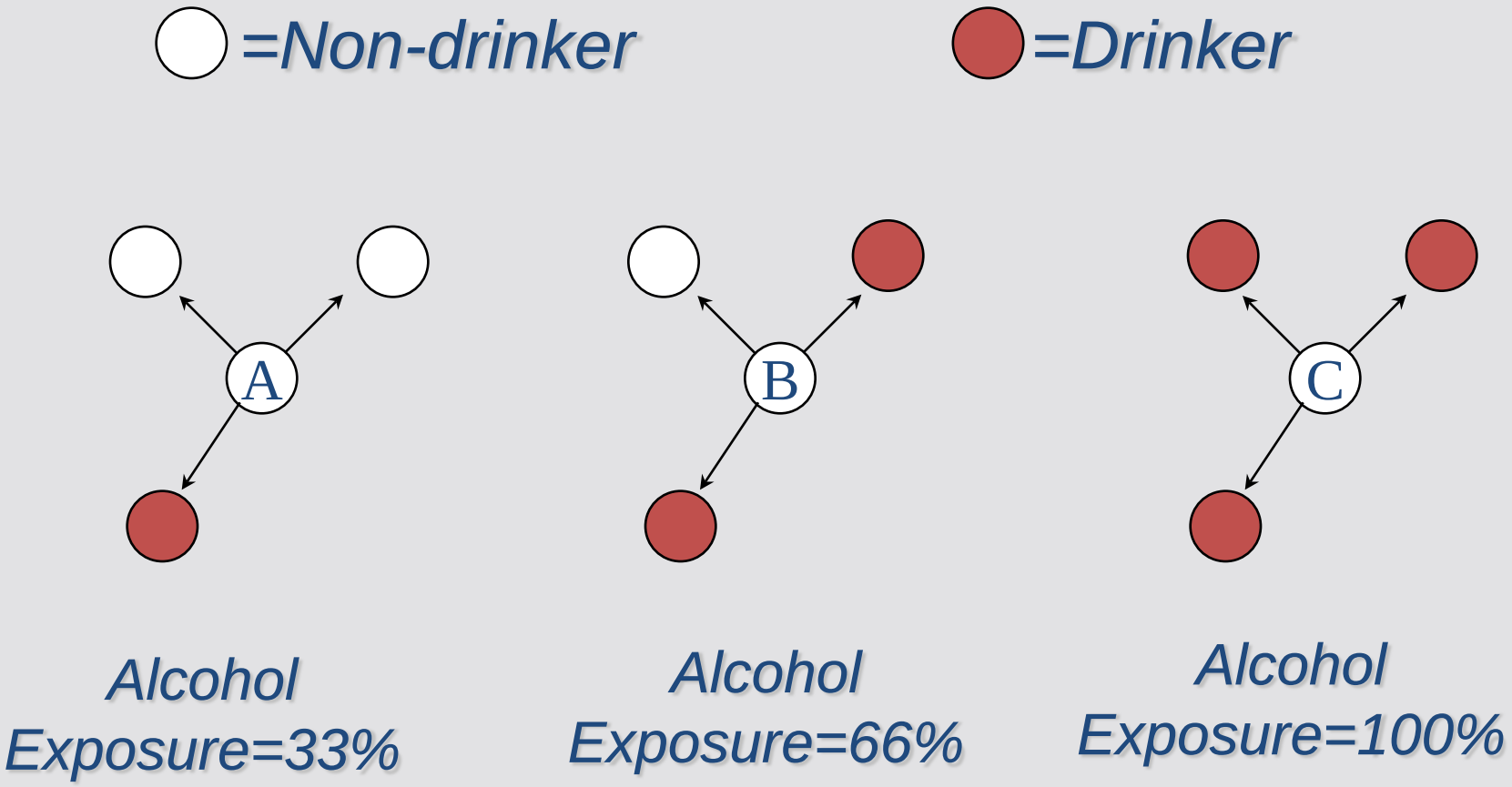
Network Exposure Model (Valente, 1995; 2005; Burt, 1987)

Background: Network exposure model

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



Source: Valente (1995)

Network Exposure Model (Valente, 1995, 2005)

$$\underline{E} = \frac{\sum_j w_{ij} y_j}{\sum_i w_i}$$

- **W** is the **friendship matrix**
- **y** is vector of **friend's drinking**

Example of computing **exposure to friend drinkers E**

$$\underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 1 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}}_{Friendship\,Matrix,\,\underline{\underline{X}}} \times \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ \vdots \\ 0 \end{bmatrix}}_{Friend'sdrinking,\,\underline{y}} \overset{\text{(element-wise)}}{\div} \underbrace{\begin{bmatrix} 4 \\ 5 \\ 5 \\ 3 \\ \vdots \\ 4 \end{bmatrix}}_{total\#ofnomination} = \underbrace{\begin{bmatrix} .15 \\ 0 \\ .20 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\underline{E}}$$

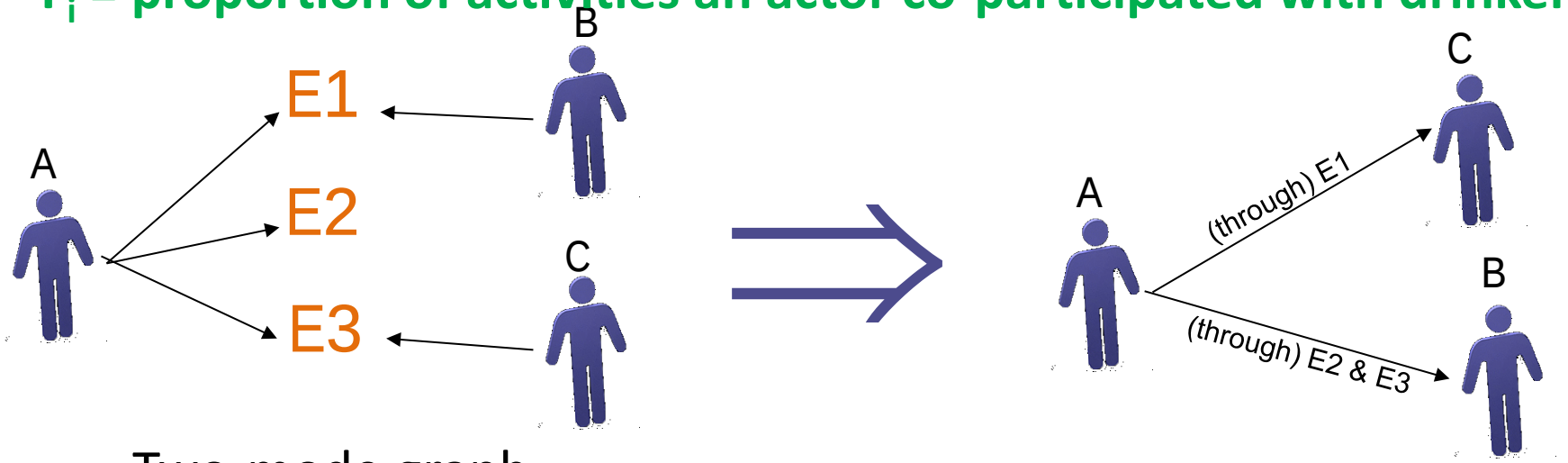
E_i = proportion of friends who drink alcohol

Affiliation Exposure Model (Fujimoto & Valente, 2011, 2012)

$$\underline{F} = \frac{\sum_j c_{ij} y_j}{\sum_i c_i}$$

- **C** is the **co-membership matrix** = $A*A^T$
- **y** is vector of **alters' behavior**

F_i = proportion of activities an actor co-participated with drinker(s)



Two-mode graph

Projected one-mode graph

Regression model
With Affiliation Exposure Term

$$y = \rho \boxed{Wy} + X\beta + \gamma D + \varepsilon, \quad \varepsilon \sim N\left(0, \sigma^2 I\right)$$

Affiliation Exposure Terms, F

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 <u>directly connected ties</u>** ☜

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

Logistic regression model	Drinking
Affiliation exposure based on sports activities	**1.23*** (0.12)
Number of sports participated in	0.99 (0.03)
Exposure to **Friends' drinkers**	**1.42***** (0.04)

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**

Fujimoto, Wang, Ross, & Williams (2015) in *American Journal of Public Health*

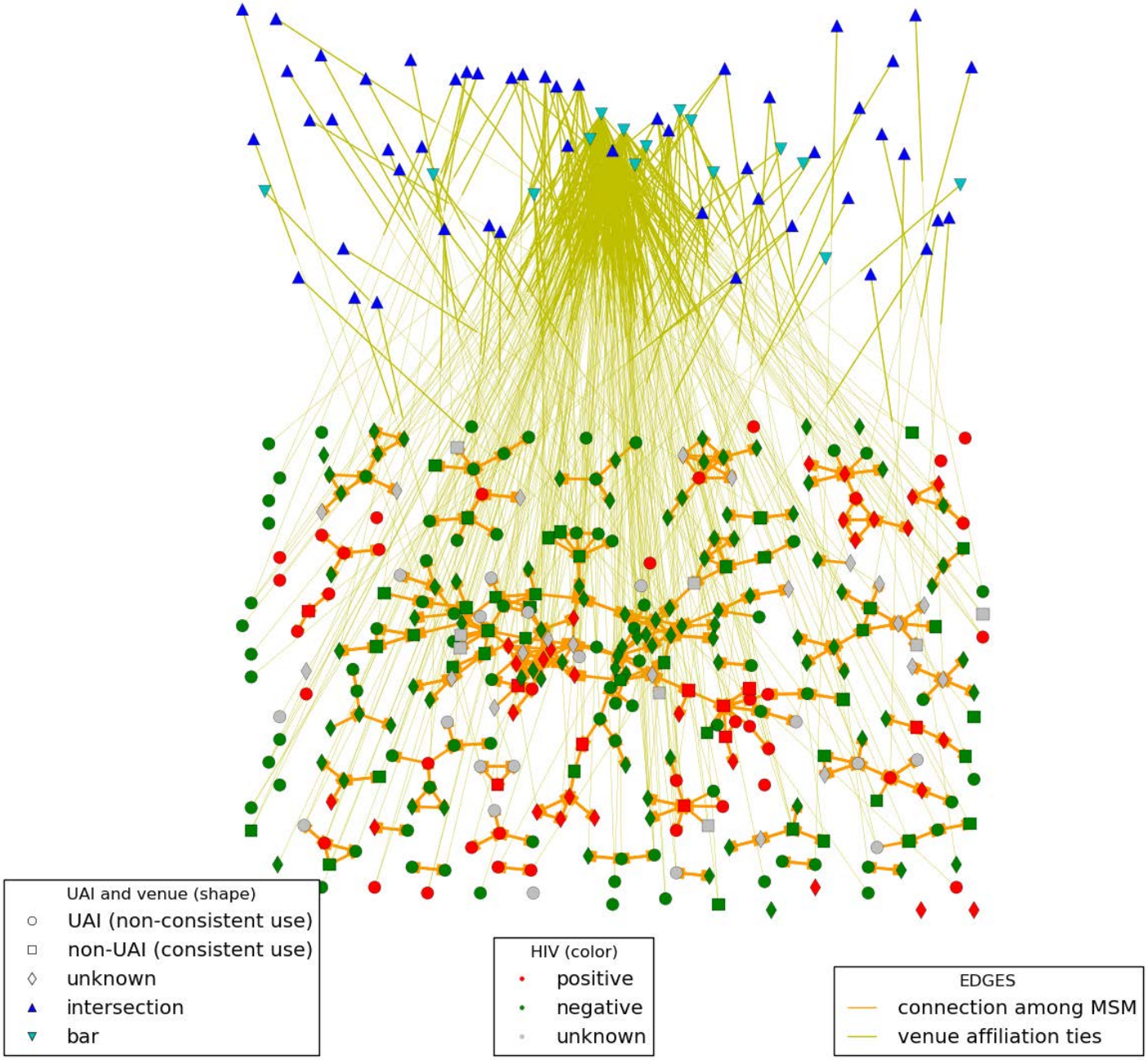
Study 2: Data and Aim

- **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

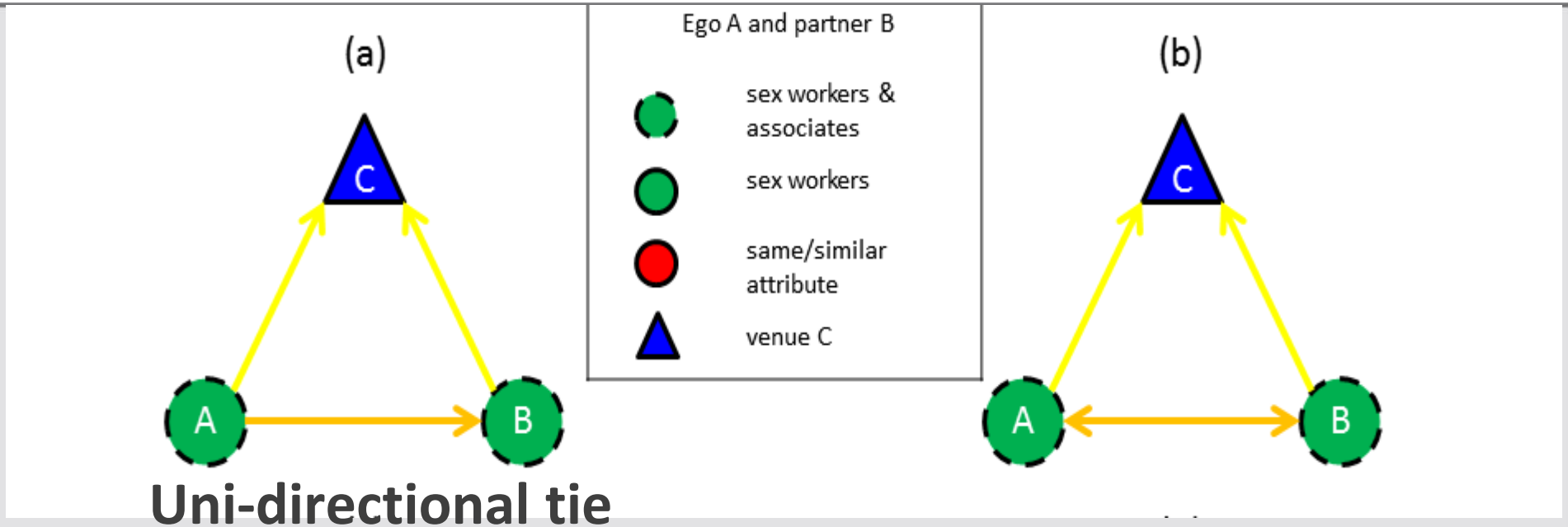
Weak ties facilitate the **diffusion of information** (Granovetter, 1983)

- 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



Uni-directional tie (Weak tie)

Bi-directional tie (Strong tie)

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

YMAP: Young Men's Affiliation Project of HIV Risk and Prevention Venue

NIH/NIMH 1R01MH100021

(PI: Fujimoto, K. & Schneider, J.A.)

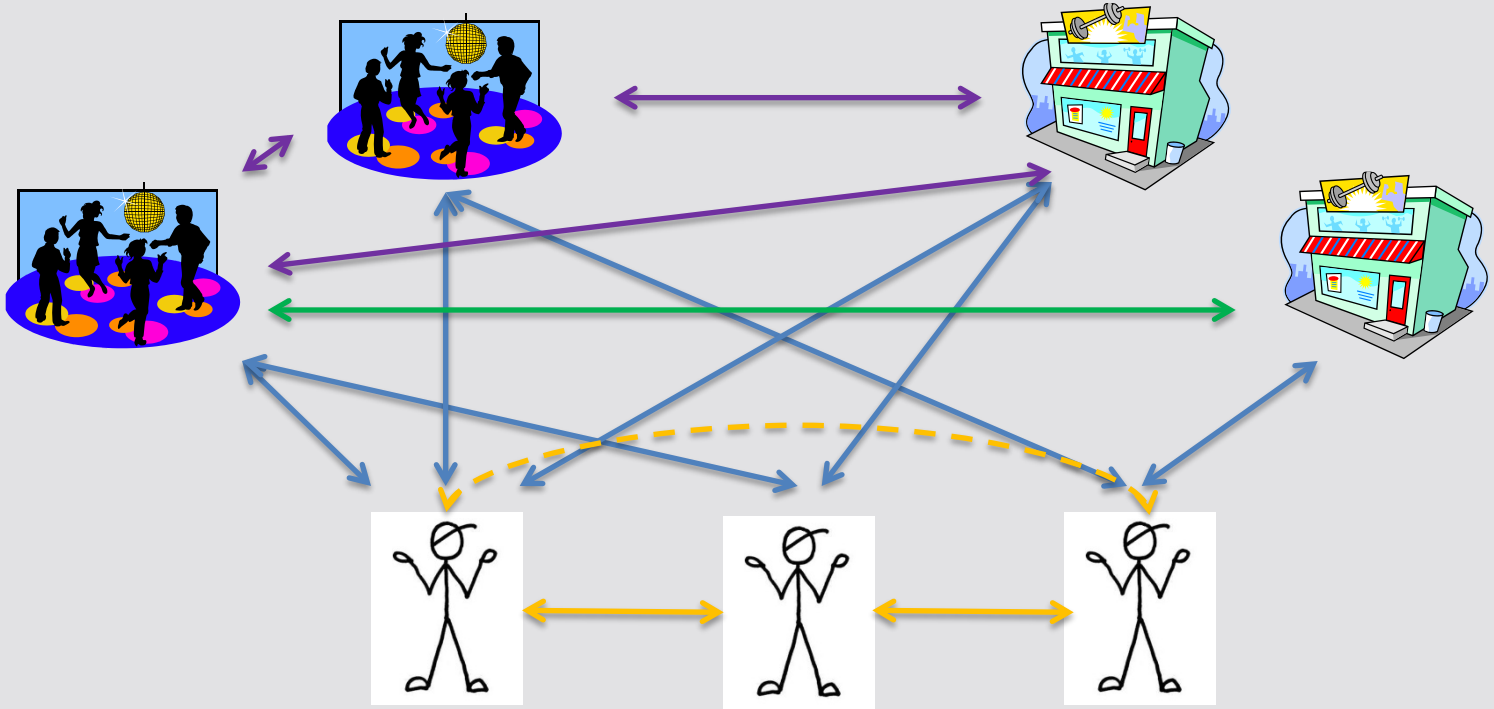
2013-2018

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**
     2. Wave 2 (2015)
     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).

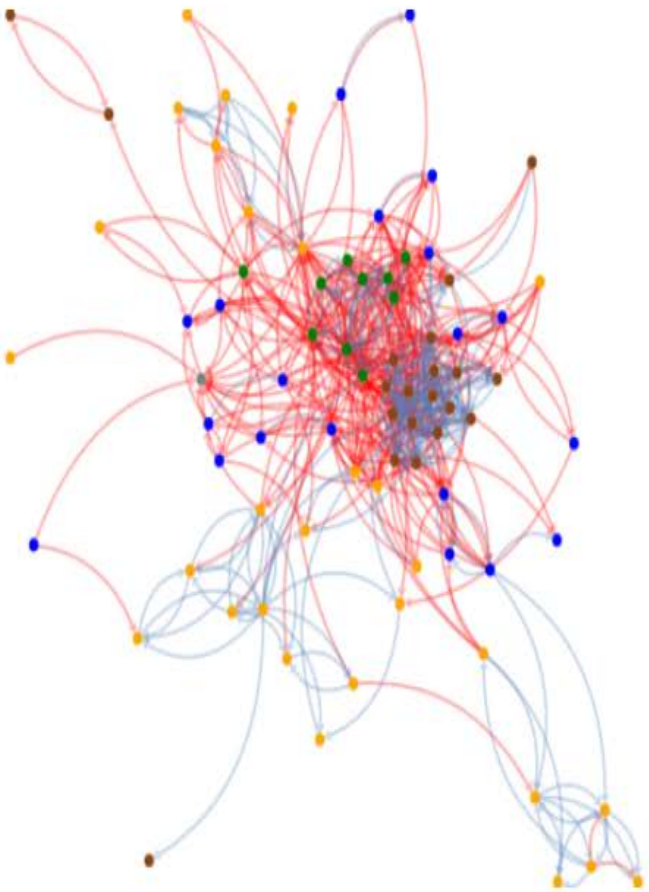
3 different types of relationships

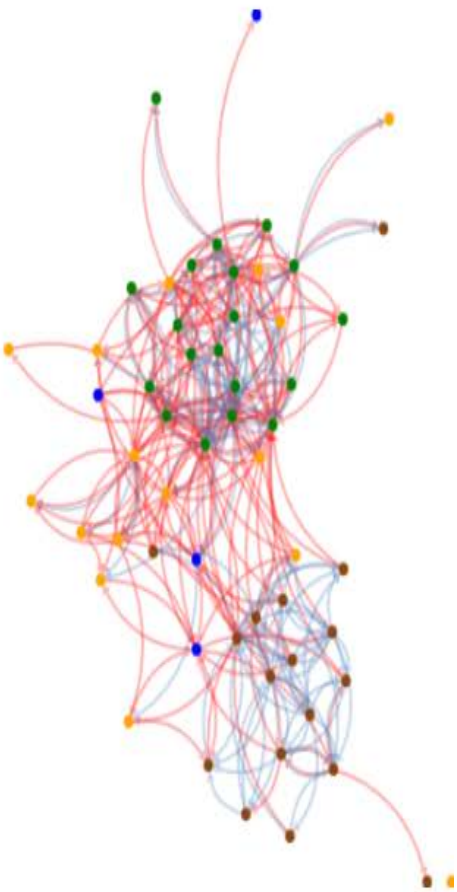
- **Collaboration**: Worked together on any activity, project, or event for a shared goal or purpose, **formally** or **informally**.
- **Competition**: **competed for resources** such as revenue, clientele or membership, employees, or other resources.
- **Sponsoring relation**: financial support, grant 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

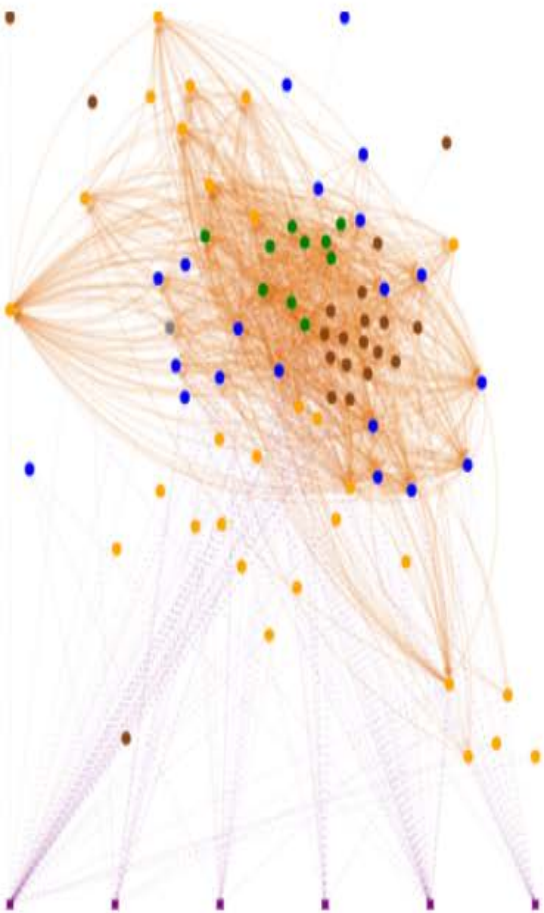


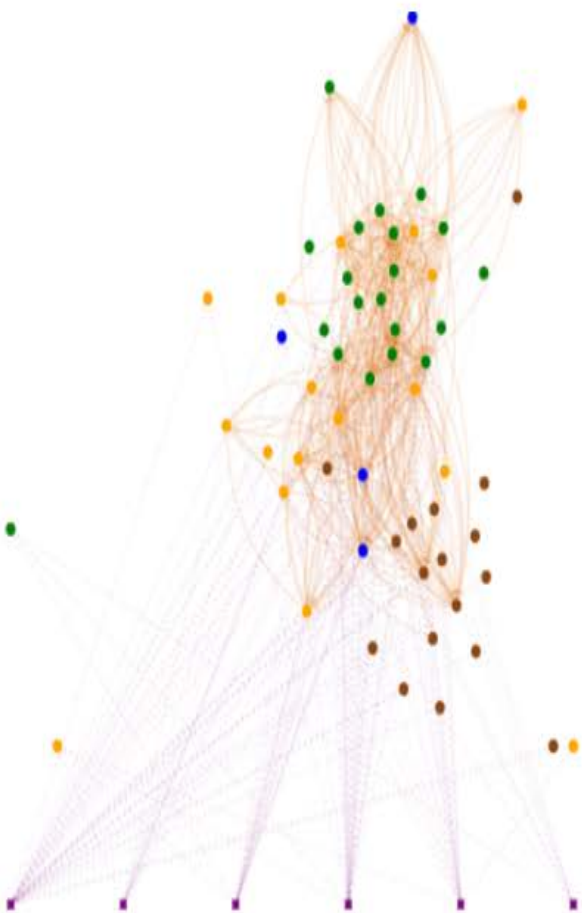


Houston

Chicago

Multiple networks of common funders and funding types





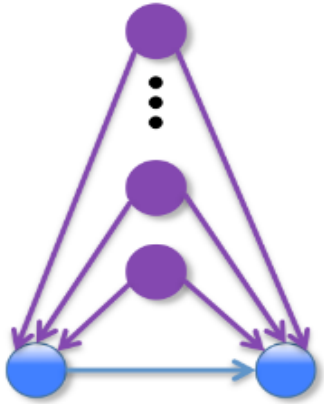
Houston

Chicago

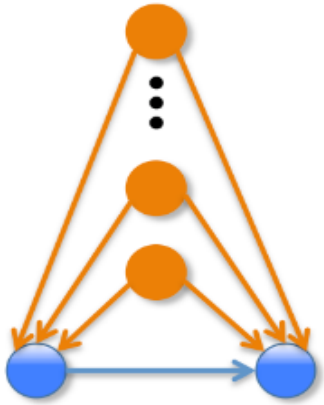
Network configurations for ERGM specifications

H1a: competition through same venue type

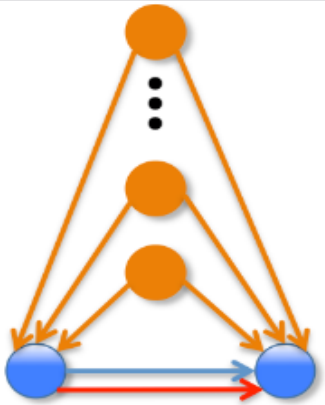
H1b: competition through geographic proximity



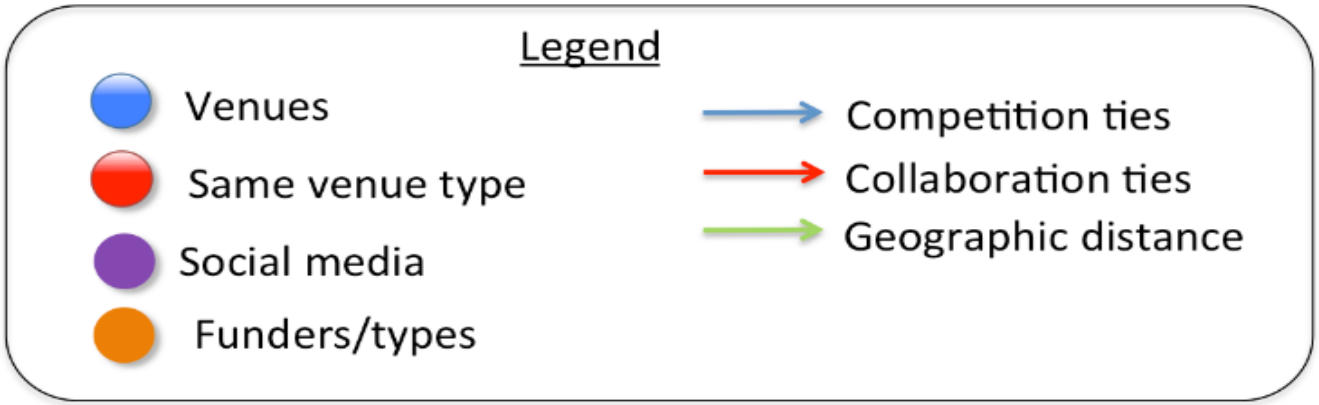
H1c: competition through shared social media use



H1d: competition through shared funding sources



H3: entrainment through shared funding sources



Study findings

1) **Risk venues** tend to compete one anther only for Houston
2) **Venues with shared media use** tend to compete one anther only for Chicago
3) **Venues with shared funder** tend to compete one anther 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

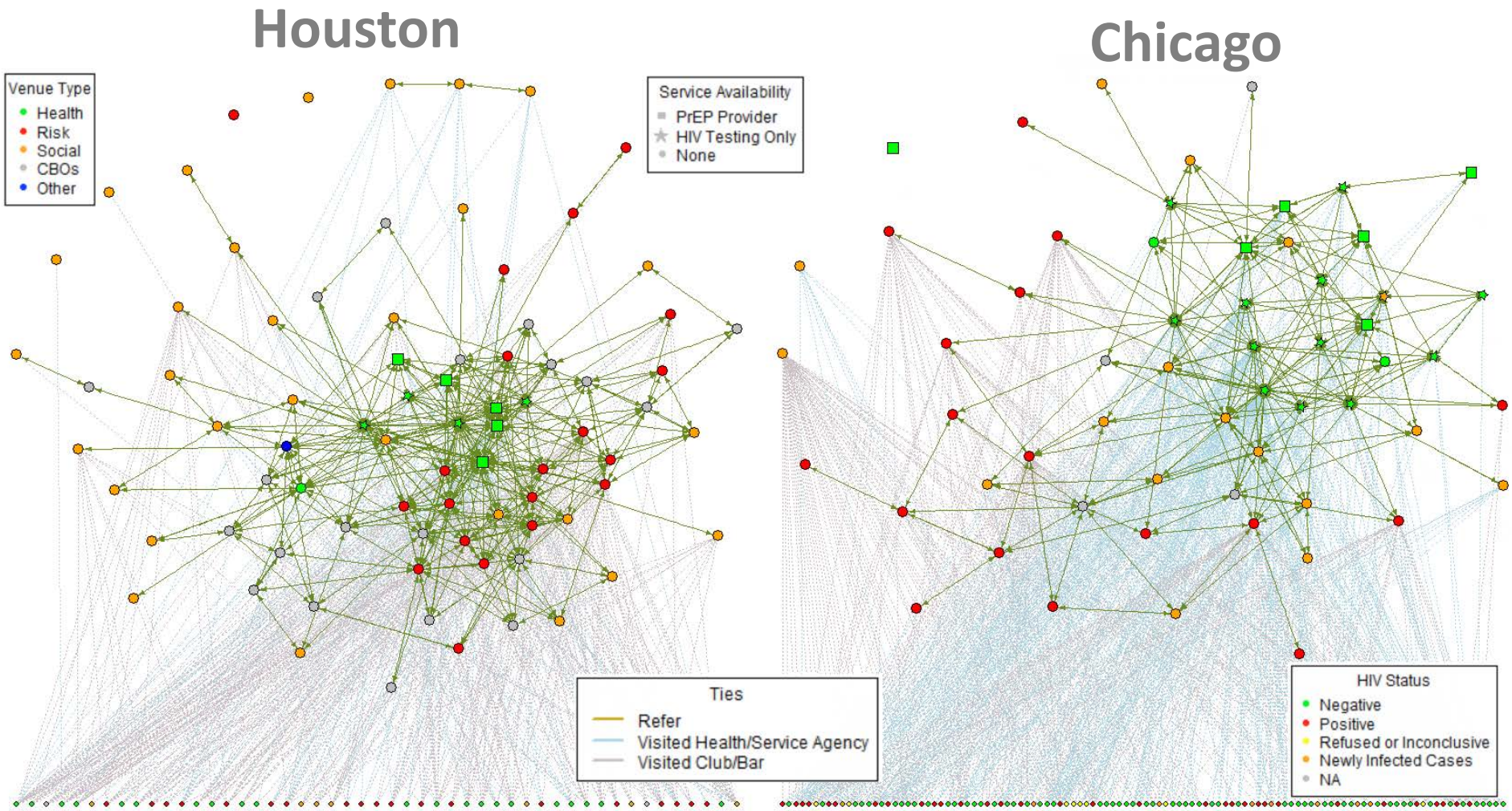
2) Implementation of a common funding program:

- Encourage all grantees to have **partnership** with otherwise **competing grantees**

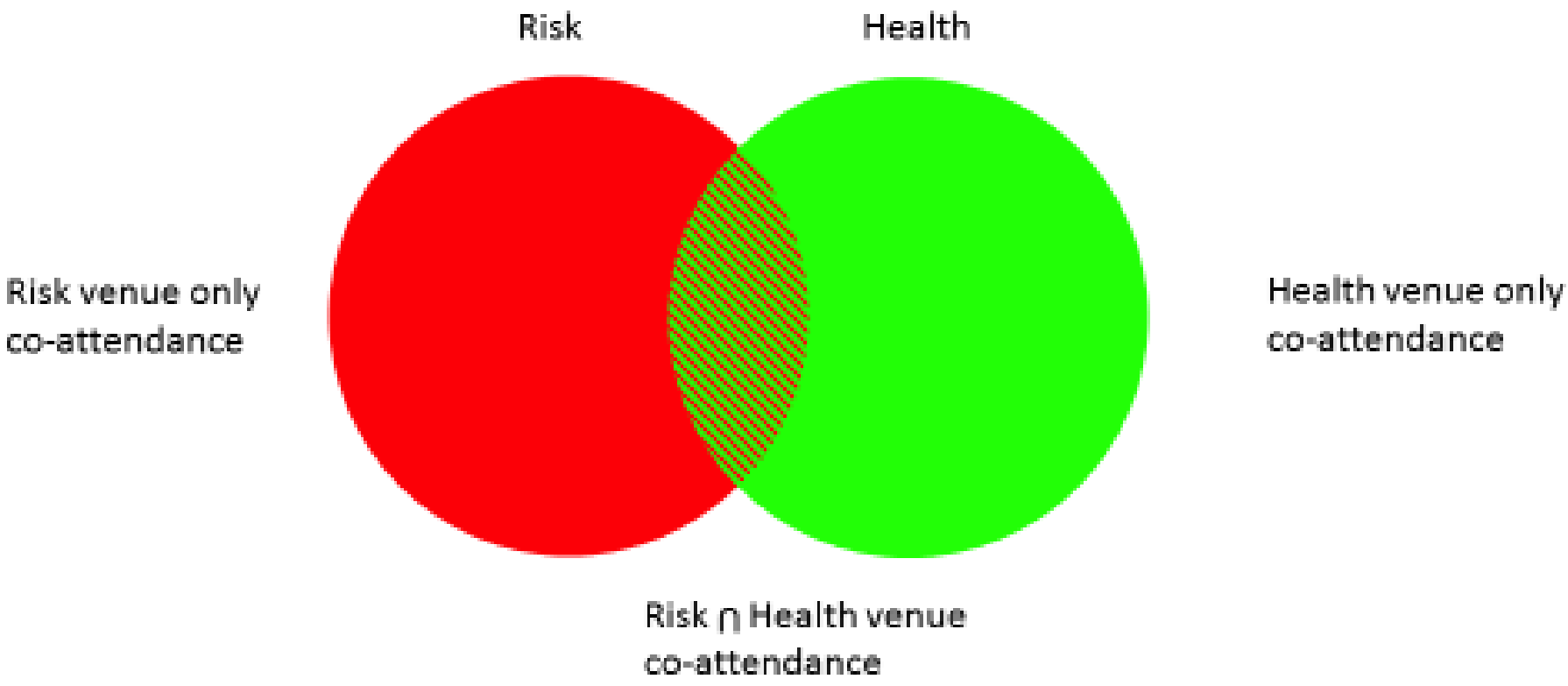
Study design—Phase II

1) **Interview with YMSM:** Sample of **younger MSM (17-29)**
   - 400 YMSM in Houston & 400 YMSM in Chicago
     1. **Wave 1 (2014-15) ← report today**
     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

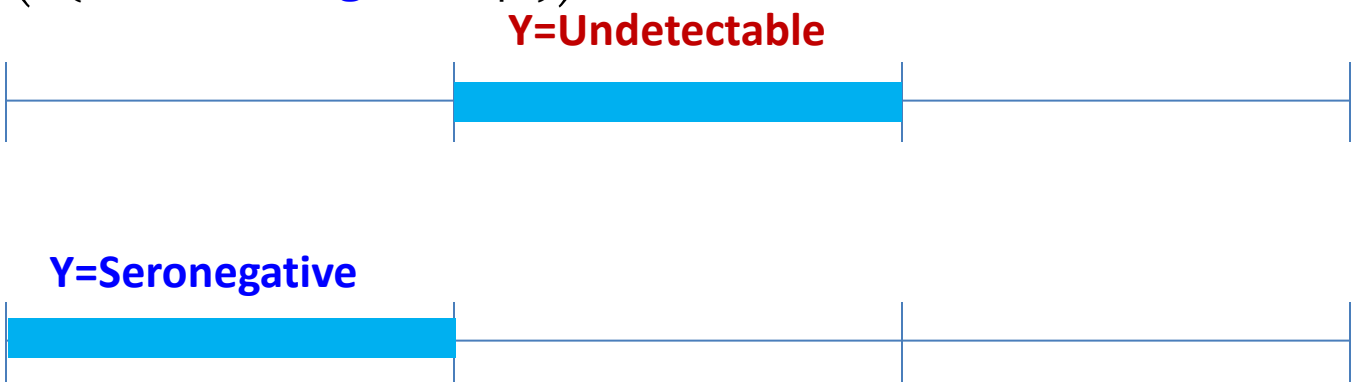
- **Viral load level (RNA)**
  1) 0-Seronegative: 182 (62%)
  2) 1-Seropositive AND VL undetectable: 65 (22%)
  3) 2-VL > 400: 48 (16%)

Multinomial logistic regression

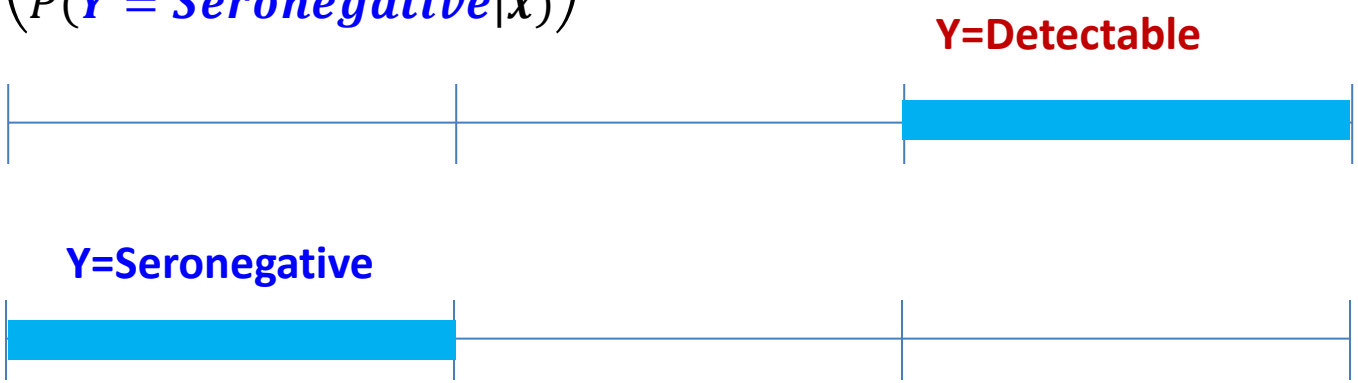
$$(1)\ g_1(x) = log\left(\frac{P(\boldsymbol{Y=Undetectable}|x)}{P(\boldsymbol{Y=Seronegative}|x)}\right) = \boldsymbol{x}'\boldsymbol{\beta}_1$$

$$(2)\ g_2(x) = log\left(\frac{P(\boldsymbol{Y=Detectable}|x)}{P(\boldsymbol{Y=Seronegative}|x)}\right) = \boldsymbol{x}'\boldsymbol{\beta}_2$$

$$g_1(\boldsymbol{x}) = log\left(\frac{P(\boldsymbol{Y = Undetectable}|\boldsymbol{x})}{P(\boldsymbol{Y = Seronegative}|\boldsymbol{x})}\right)$$



$$g_2(\boldsymbol{x}) = log\left(\frac{P(\boldsymbol{Y = Detectable}|\boldsymbol{x})}{P(\boldsymbol{Y = Seronegative}|\boldsymbol{x})}\right)$$



Adjusted Odds Ratio of Logit Models

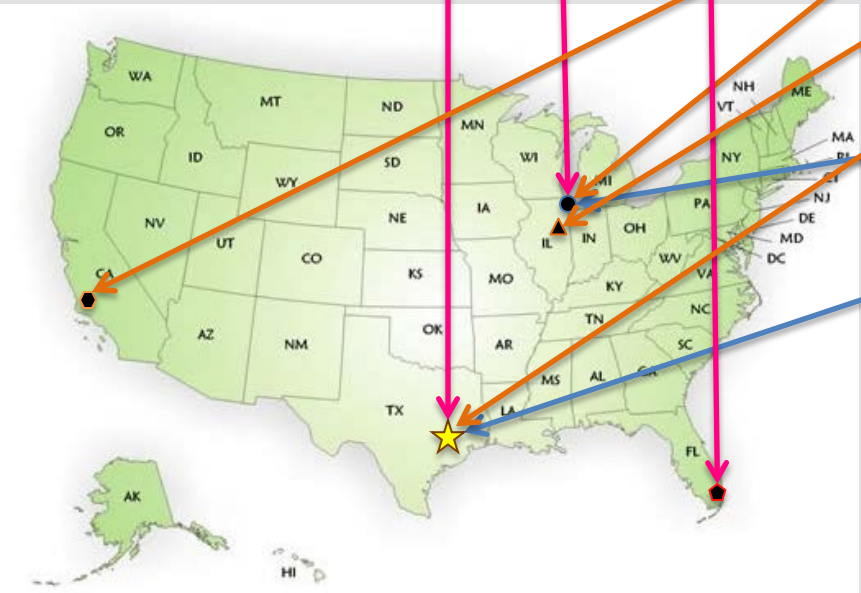
Multinomial regression model	Undetectable vs seronegative	Detectable vs seronegative
Risk and health venue co-attendance	--	--
Only risk venue co-attendance	--	5.13*
Only health venue co-attendance	--	--
# of risk and health venue attended	.86***	.90*
# of risk venue attendance		
# of health venue attendance		

Extension of YMAP to other projects

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)

Chicago (n=56)

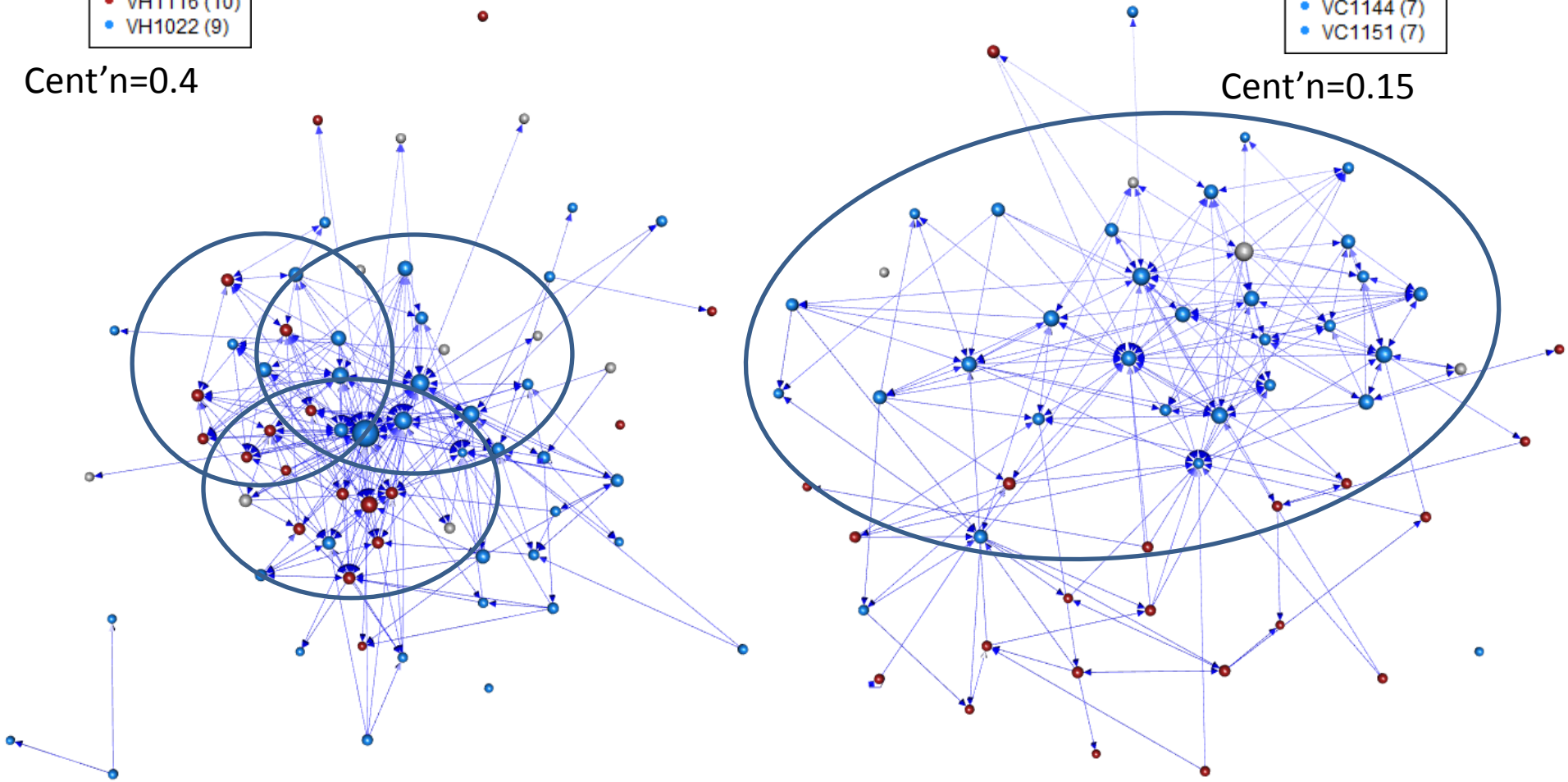
Top 10 in degree
- VH1011 (22)
- VH1014 (21)
- VH1001 (17)
- VH1007 (13)
- VH1018 (12)
- VH1012 (11)
- VH1036 (11)
- VH1108 (11)
- VH1116 (10)
- VH1022 (9)

Risk Classification
- Risk
- Protective
- Neither

Top 10 in degree
- VC1003 (21)
- VC1005 (17)
- VC1148 (12)
- VC1006 (9)
- VC1145 (8)
- VC1014 (7)
- VC1016 (7)
- VC1041 (7)
- VC1144 (7)
- VC1151 (7)

Cent'n=0.4

Cent'n=0.15



Competition network

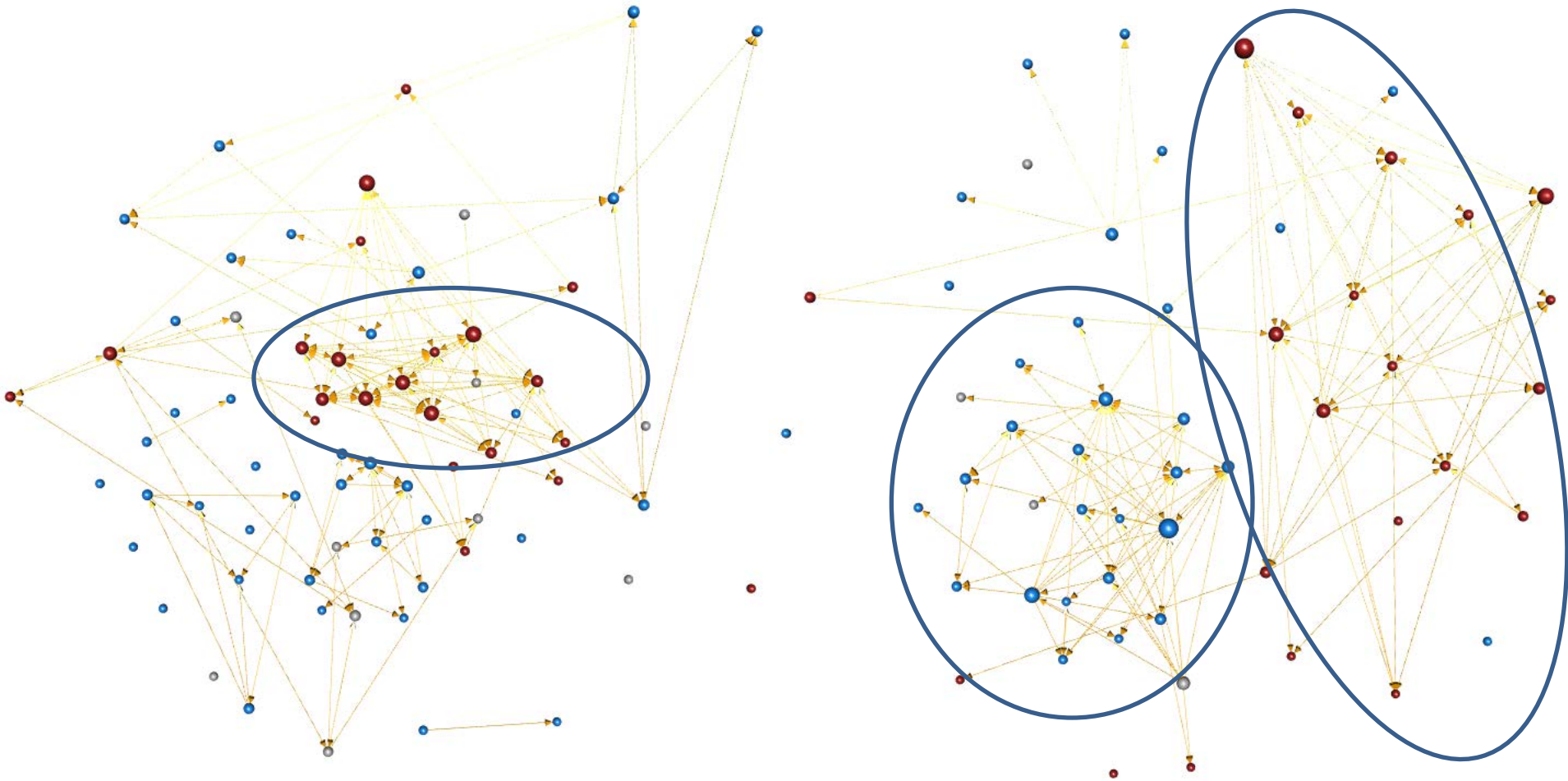
Houston (n=70)

Top 10 in degree
- VH1038 (10)
- VH1032 (9)
- VH1048 (9)
- VH1052 (8)
- VH1066 (8)
- VH1033 (7)
- VH1036 (7)
- VH1049 (7)
- VH1007 (6)
- VH1041 (6)

Risk Classification
- Risk
- Protective
- Neither

Chicago (n=56)

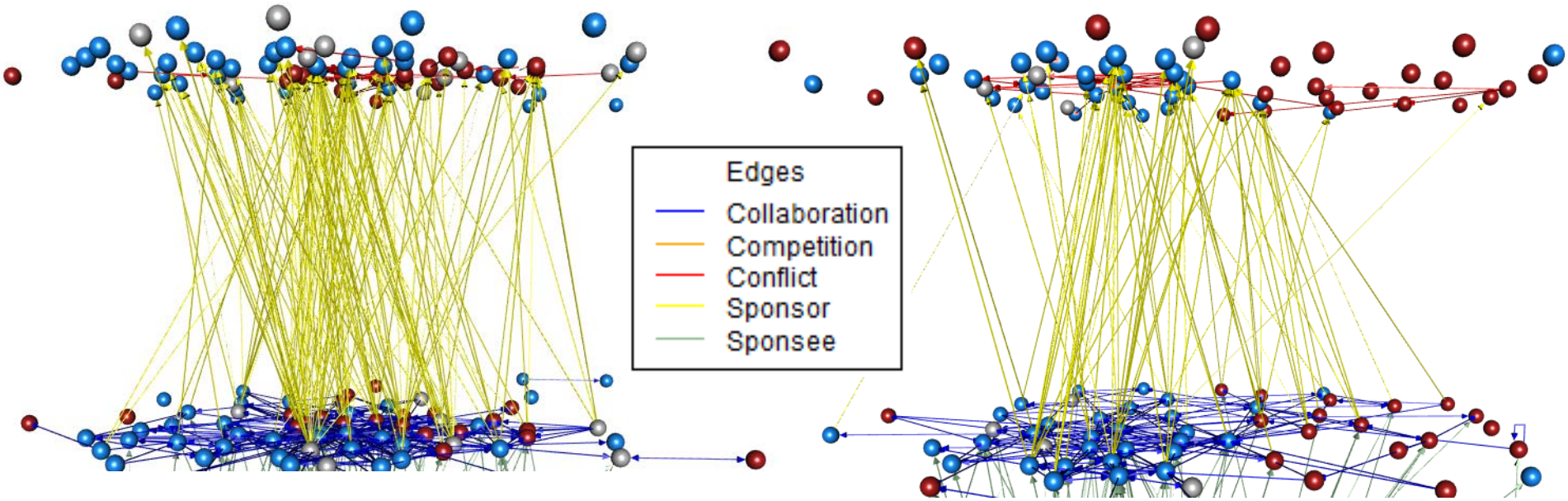
Top 10 in degree
- VC1003 (12)
- VC1006 (7)
- VC1027 (7)
- VC1005 (6)
- VC1002 (5)
- VC1010 (5)
- VC1070 (5)
- VC1009 (4)
- VC1014 (4)
- VC1028 (4)



Houston (n=70)

Sponsor networks

Chicago (n=56)



Risk Classification
- Risk
- Protective
- Neither