Advanced Network Analysis

Olga Chyzh [www.olgachyzh.com]

Today's Class

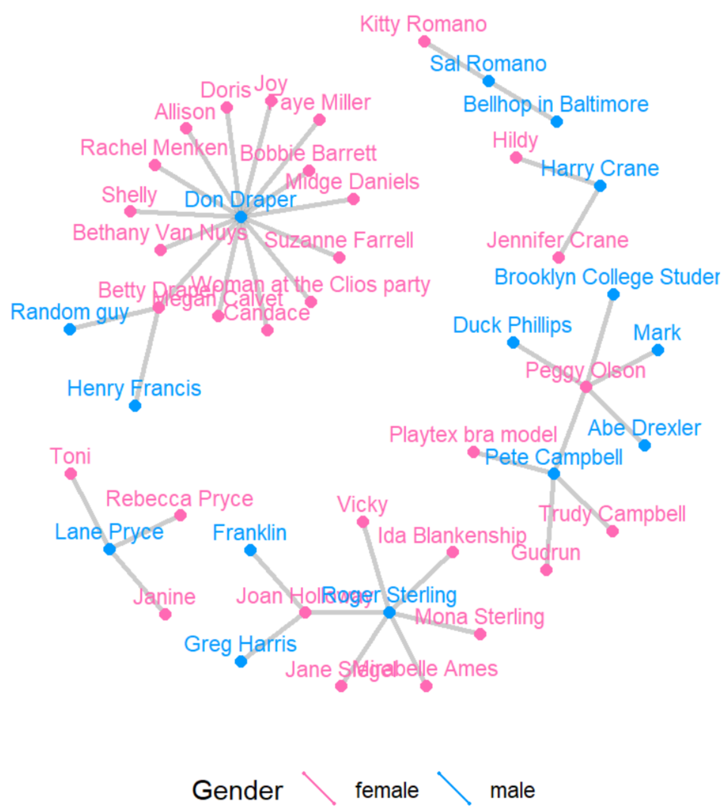
Network Analysis: Getting Started with the super basics

1. Definitions. What is network analysis?
2. Network Science: Origins
3. Network features and measurements
4. Collecting network data
5. What does this all mean for Political Science?
6. Processing network data in R

Definitions. What Is Network Analysis?

What is a network (i.e., a graph)?

A set of **nodes** and **relation(s)** defined on them



Defining Network Features & Measurements: What's a node?

- A **node** can be defined as an entity that can form relations with other entities.

Synonyms:

- actor: from sociometry, common terminology in sociology and psychology
- vertex: from graph theory (i.e., math), common terminology in mathematics and physics

Term node is common in statistics and applied sciences outside of soc and psych.

Examples of Nodes

- Individuals (Mad Men characters, legislators)
- Families (Padgett and Ansell's reading)
- Human Rights NGOs
- Countries

Defining Network Features & Measurements: What's a relation?

- A **relation/tie** defines the existence of an attribute relating nodes.

Synonyms:

- link: common in computer science (e.g., huge lit on “Link Prediction”) and social sciences
- edge: graph theoretic terminology common in physics and math, but also elsewhere

Ties can have characteristics:

- Weight
- Qualitative attributes
- Direction

Let's brainstorm ties to go with our nodes

- Romantic relationship, marriage, friendship
- Business relationship
- Cooperation/conflict

Network Science: Origins

Origins and History of Network Analysis

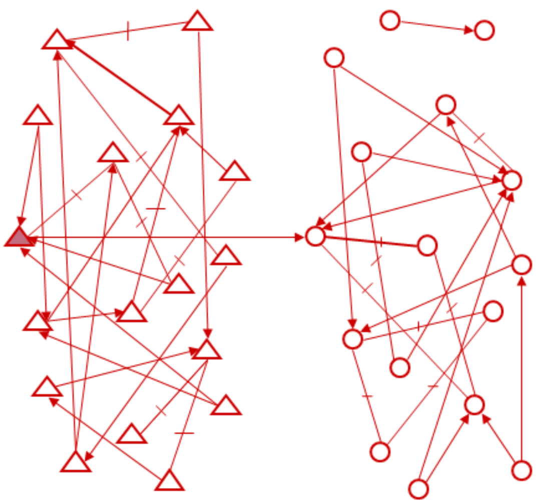
Early Puzzles

- Sociologists began using the term as early as 1887 and early 1990s
- Emile Durkheim, Jacob Moreno, and later Harrison White (among others) were interested in understanding social patterns and the relations between members of a system.
    - How do people feel towards one another? Why might this matter?

Early study of network analysis

Early Puzzles: Individuals inside social groups

- Example: In 1932 there was a pandemic of runaways at Brooklyn public and private schools: within two weeks 14 girls ran away, which was 30 times more than the average number
  - Moreno's finding: position in network predicted whether the girl would run away



Early study of network analysis

Kathleen Carley, Another Early Puzzle (1980-90s): Group Stability and Organization

Early study of network analysis

Kathleen Carley, Another Early Puzzle (1980-90s): Group Stability and Organization

- Example: What makes a group stable? Application: founders/leaders of businesses

Early study of network analysis

Kathleen Carley, Another Early Puzzle (1980-90s): Group Stability and Organization

- Example: What makes a group stable? Application: founders/leaders of businesses
- Shared knowledge between actors matters as much, if not more, than context for group stability

Early study of network analysis

Kathleen Carley, Another Early Puzzle (1980-90s): Group Stability and Organization

- Example: What makes a group stable? Application: founders/leaders of businesses
- Shared knowledge between actors matters as much, if not more, than context for group stability
- i.e., characteristics outside the group versus inside of the group

Defining Network Features & Measurements

How can we capture these relationships? *The Sociomatrix.* example: *PONIES*

Ranks and Relationships in Highland Ponies and Highland Cows by **Brock et. al 1976**

Sociomatrix: Pony threats

Directed, targeted behavior

Table 3: Threat relationships between ponies on the hill. The table shows the frequency with which each pony threatened each other individual. Animals were ranked according to the method used by SCHEIN and FOHRMAN (1955). Values which are underlined indicate the two ponies which each animal threatened relatively most except in the four lowest ranking ponies where only the identity of the most threatened individual is shown

Threatener	WT	WH	WS	GA	BR	BA	TD	WG	PM	CA	GD	DA	2 B	2 D	2 G	2 S	TA	Total threats	Number of different ponies threatened
WT		2	8	6	8	10	8	15	5	12	6	14	9	15	4	3	9	134	16
WH			6	8	6	1	2		3	7	9	4	4		2	4	2	58	13
WS				1	9	8	7		1	9	11	10	7	1	6	1	4	75	13
GA					3	1	2		2	4	3	8	5	3	1	2		34	11
BR				1		2	3	12	4	9	11	6	6	13	5	3	7	82	13
BA		6			1		2		1	3	7	5	4	0	3		5	38	10
TD					1			7	1	4	4	1	3	6	3	5	3	38	11
WG		2	3				1			8	4	3	5	3	5	4	3	41	11
PM										6	7	6	9	9	7	9	6	59	8
CA									1		8	5	10	9	9	5	8	55	8
GD										2		18	8	4	6	8	5	51	7
DA									1	2	2		5	4	8	7	4	33	8
2 B											4	4		5	6	7	10	36	6
2 D													1		1	5	4	11	4
2 G																6	2	8	2
2 S												1		1	2		4	8	4
TA															1			1	1

Sociomatrix: Pony grooming

Undirected, mutual behavior

Table 5: Grooming relationships between ponies. The table shows the frequency with which each individual groomed each other pony. Animals were ranked according to their position in the threat hierarchy. Values which are underlined indicate the two ponies which each individual groomed with relatively most (see p. 210)

Grooms with	WT	WH	WS	GA	BR	BA	TD	WG	PM	CA	GD	DA	2 B	2 D	2 G	2 S	TA	Total grooming session	Number of different ponies groomed with
WT			1			1		$\underline{8}$			2							12	4
WH			5		$\underline{33}$	$\underline{25}$	2		2	4	5	2	1					79	9
WS	1	$\underline{5}$			1	$\underline{5}$		1			1			1				15	7
GA					$\underline{11}$	1	3	1		2	2	3	2					25	7
BR		$\underline{33}$	1	11		4	4	1	4	4	$\underline{23}$	4	1	4	2	3	2	100	15
BA	1	$\underline{25}$	5	1	4		$\underline{14}$			12	3	4	2	1	1			73	12
TD		2		3	4	$\underline{14}$								3				26	5
WG	$\underline{8}$		1	1						1								11	4
PM		2			4					6	$\underline{12}$	$\underline{9}$	1	2	2	2		40	9
CA		4		2	4	$\underline{12}$		1	6		$\underline{8}$	2	1		1	1	2	44	12
GD	2	5	1	2	$\underline{23}$	3			12	8		$\underline{21}$	4	1	3	2		87	13
DA		2		3	4	4			9	2	$\underline{21}$		1	3	6	6	1	62	12
2 B		1		2	1	2			1	1	4	1		2	$\underline{15}$	3	3	36	12
2 D			1		4	1	3		2		1	3	2		4	1	4	26	11
2 G					2	1			2	1	3	6	$\underline{15}$	4		$\underline{12}$	7	53	10
2 S					3				2	1	2	$\underline{6}$	3	1	$\underline{12}$		2	32	9
TA					2					2		1	3	4	$\underline{7}$	2		21	7

How might we look at this in R?

- Our class has an R package that will contain all the datasets.
- To load our class package (you have to do this each time you need to access the data):

```
library(devtools)
install_github("ochyzh/networkdata")
library(networkdata)
```

- Now, load the specific dataset:

```
data(highlandPonies)
ponies<-as.matrix(highlandPonies[1:17, 2:18])
```

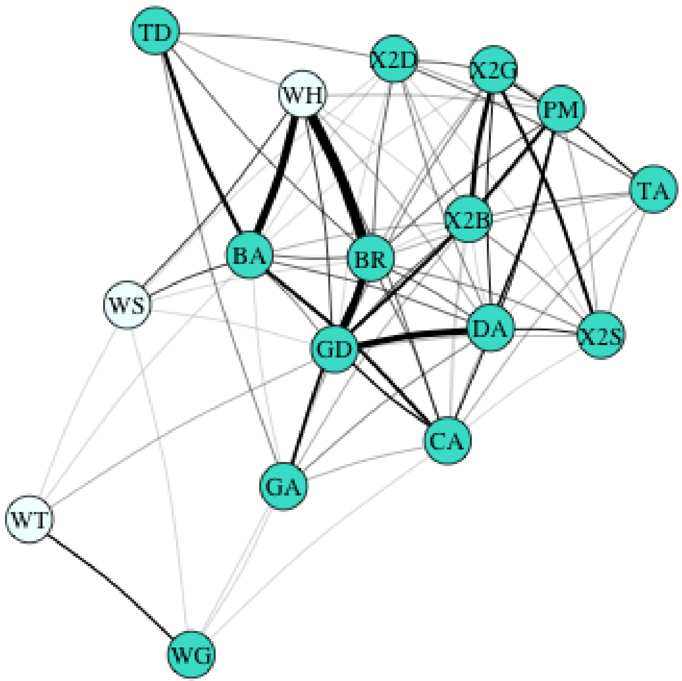
Create a graph object:

```
library(igraph)
pGraph <- graph_from_adjacency_matrix(ponies, weighted=TRUE, mode="ur
                                      diag=FALSE)
```

Plot using the *igraph* package:

```
V(pGraph)$color <- ifelse(V(pGraph)$name %in% c("WT", "WH", "WS"), "a
ponyPlot<- plot(pGraph,
     edge.arrow.size=.2,
     edge.color="black",
     vertex.frame.color="black",
     vertex.label=V(pGraph)$names,
     vertex.label.color="black", layout=layout_with_fr,
     edge.width=E(pGraph)$weight/5, edge.curved=.08)
```

Mutually supportive behavior between ponies



Other Examples of Neworks

- Survey data (Mexico violence)
- Text (co-occurrence matrices)
- Event data (conflict between actors, shared behavior between actors)
- Membership data (subcommittees)

Collecting network data

Activity 1: Build your own sociomatrix of shared experience

Break-out groups interview

1. Brainstorm a question you can ask each member of the group.
2. Keep the question simple yet specific enough to get variation in this small sample, i.e. "have you ever visited Texas" or "have you ever graphed a social network before?"
3. Record this observational data where in each observation (person) is a row in the data and their response (the variable you measured) is a column.
4. Now transform this data, with pen and pencil, into a sociomatrix.

Activity 1: Undirected sociomatrix of Shared Interests

Step 1: A simple Data set

Step 2: A simple, undirected, sociomatrix

Name	Variable, 1=yes
Iris	1
Ash	1
Chris	0
Kym	1

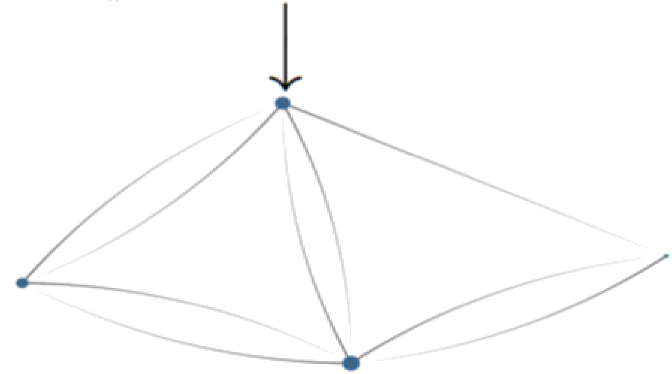
	Iris	Ash	Chris	Kym
Iris		1	0	1
Ash	1		0	1
Chris	0	0		0
Kym	1	0	0	

Summary: data processing

Sender	Receiver	Event
i	j	y_{ij}
$\vdots$	k	y_{ik}
	l	y_{il}
j	i	y_{ji}
$\vdots$	k	y_{jk}
	l	y_{jl}
k	i	y_{ki}
$\vdots$	j	y_{kj}
	l	y_{kl}
l	i	y_{li}
$\vdots$	j	y_{lj}
	k	y_{lk}

$\longrightarrow$

	i	j	k	l
i	NA	y_{ij}	y_{ik}	y_{il}
j	y_{ji}	NA	y_{jk}	y_{jl}
k	y_{ki}	y_{kj}	NA	y_{kl}
l	y_{li}	y_{lj}	y_{lk}	NA



Networks in Political Science

Dyads

- Introduced by the use of dyads, largely in International Relations literature

Networks in Political Science

Dyads

- Introduced by the use of dyads, largely in International Relations literature
- Early work in IR focused on the behavior and policies of individual states (for example, Morgenthau 1948).

Networks in Political Science

Dyads

- Introduced by the use of dyads, largely in International Relations literature
- Early work in IR focused on the behavior and policies of individual states (for example, Morgenthau 1948).
- Analysis of pairs of countries (trade, war, democracy, political ties).

Networks in Political Science

Dyads

- Introduced by the use of dyads, largely in International Relations literature
- Early work in IR focused on the behavior and policies of individual states (for example, Morgenthau 1948).
- Analysis of pairs of countries (trade, war, democracy, political ties).
    - Example:
        - US-Iraq 2003: War
        - US-Iran 2003: No War
        - Iran-Iraq 2003: No War

Networks in Political Science

Dyads

- Introduced by the use of dyads, largely in International Relations literature
- Early work in IR focused on the behavior and policies of individual states (for example, Morgenthau 1948).
- Analysis of pairs of countries (trade, war, democracy, political ties).
    - Example:
        - US-Iraq 2003: War
        - US-Iran 2003: No War
        - Iran-Iraq 2003: No War

This image again (or why logits are not who you think they are)

Sender	Receiver	Event
i	j	y_{ij}
$\vdots$	k	y_{ik}
	l	y_{il}
j	i	y_{ji}
$\vdots$	k	y_{jk}
	l	y_{jl}
k	i	y_{ki}
$\vdots$	j	y_{kj}
	l	y_{kl}
l	i	y_{li}
$\vdots$	j	y_{lj}
	k	y_{lk}

$\longrightarrow$

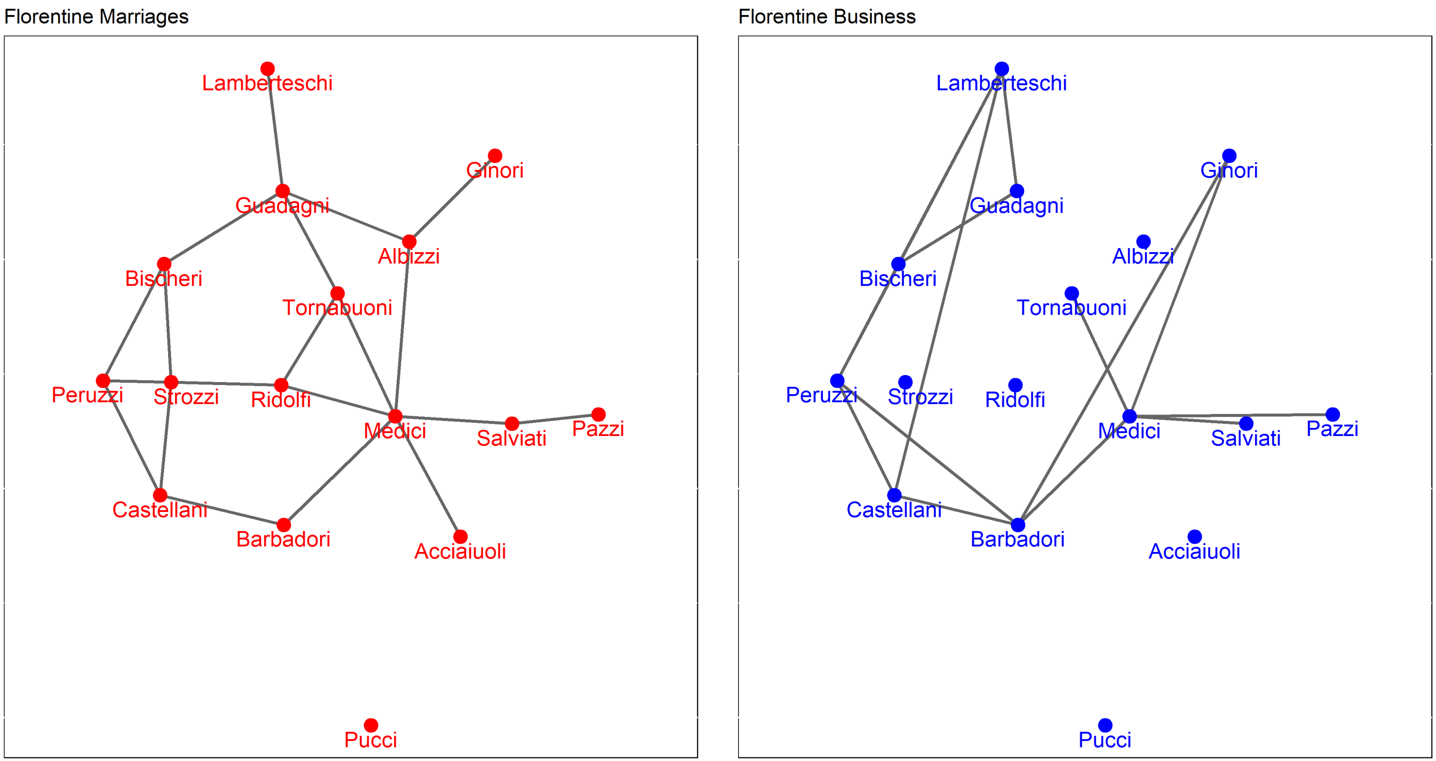
	i	j	k	l
i	NA	y_{ij}	y_{ik}	y_{il}
j	y_{ji}	NA	y_{jk}	y_{jl}
k	y_{ki}	y_{kj}	NA	y_{kl}
l	y_{li}	y_{lj}	y_{lk}	NA

Networks in Political Science

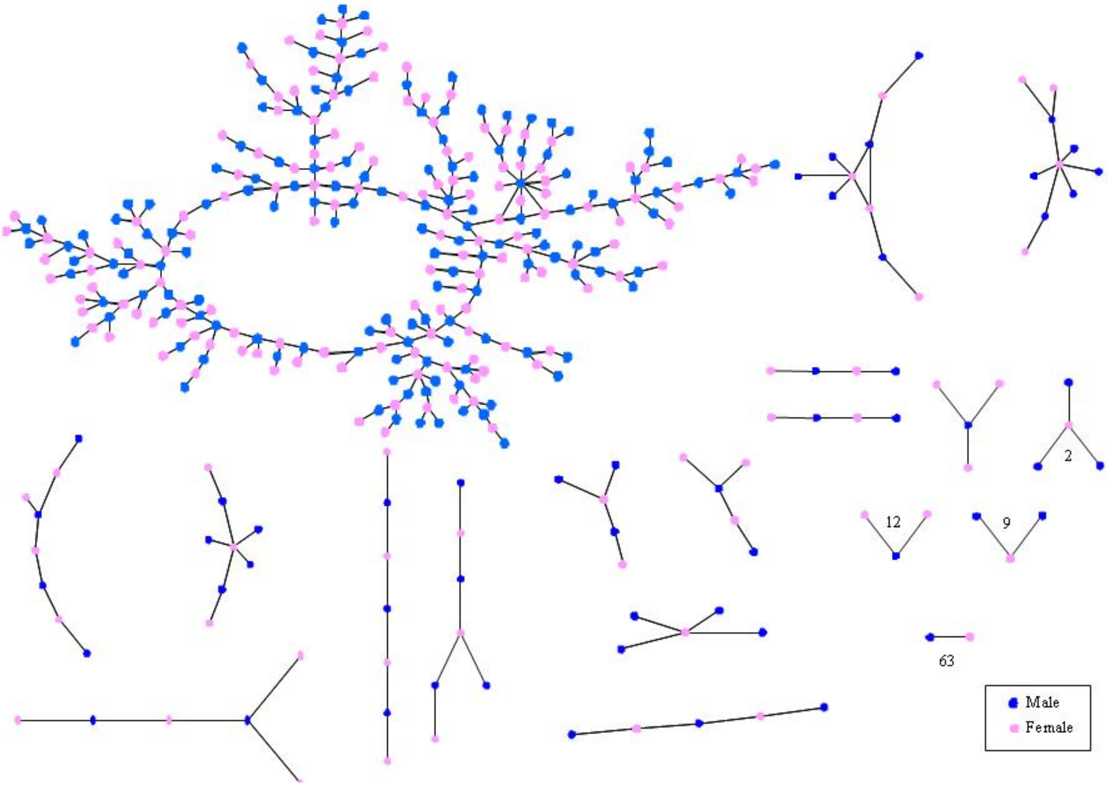
Today: Systems (Dyads --> Networks)

- Researchers recognize that dyads cannot be studied independently
- Network analysis is seen in a wide variety of applications both within and beyond Political Science:
    - geography
    - spatial analysis
    - conflict studies
    - peer-networks
    - congressional voting

Network graphs can reveal important structures

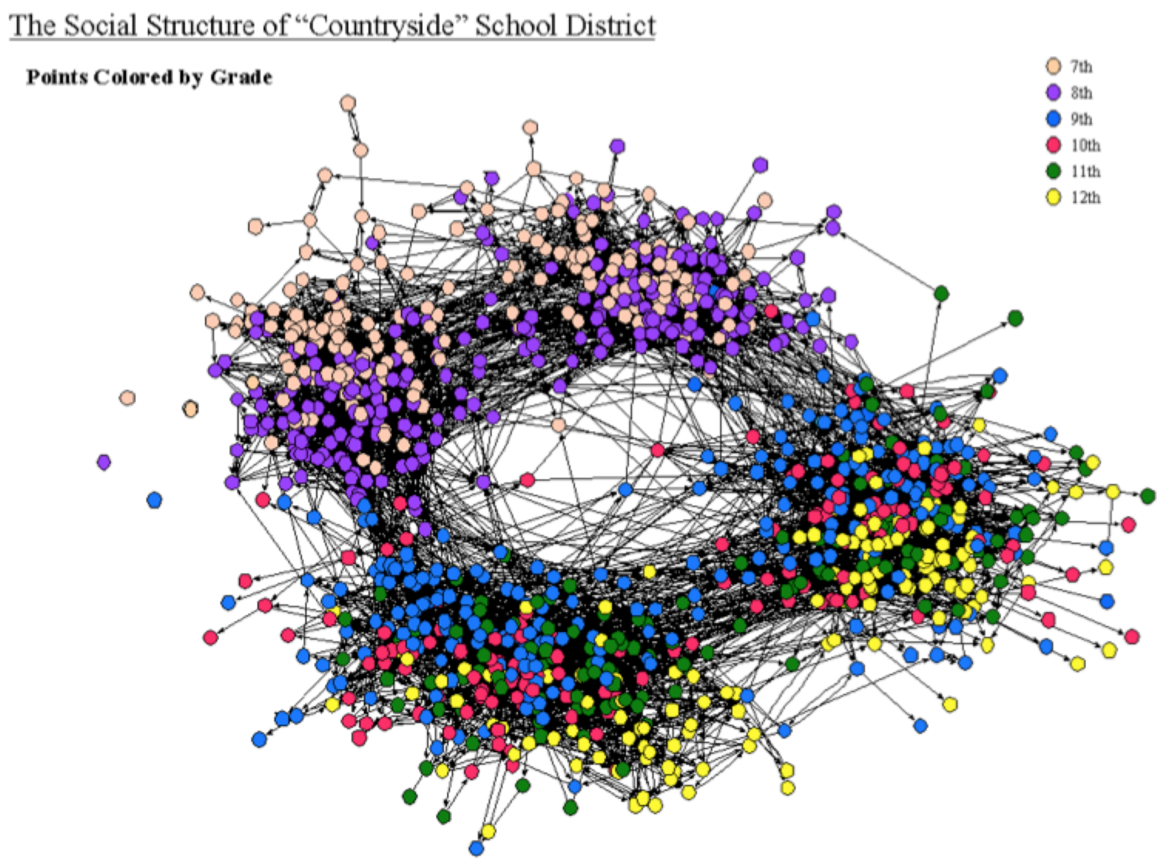


Adolescent romantic and sexual networks

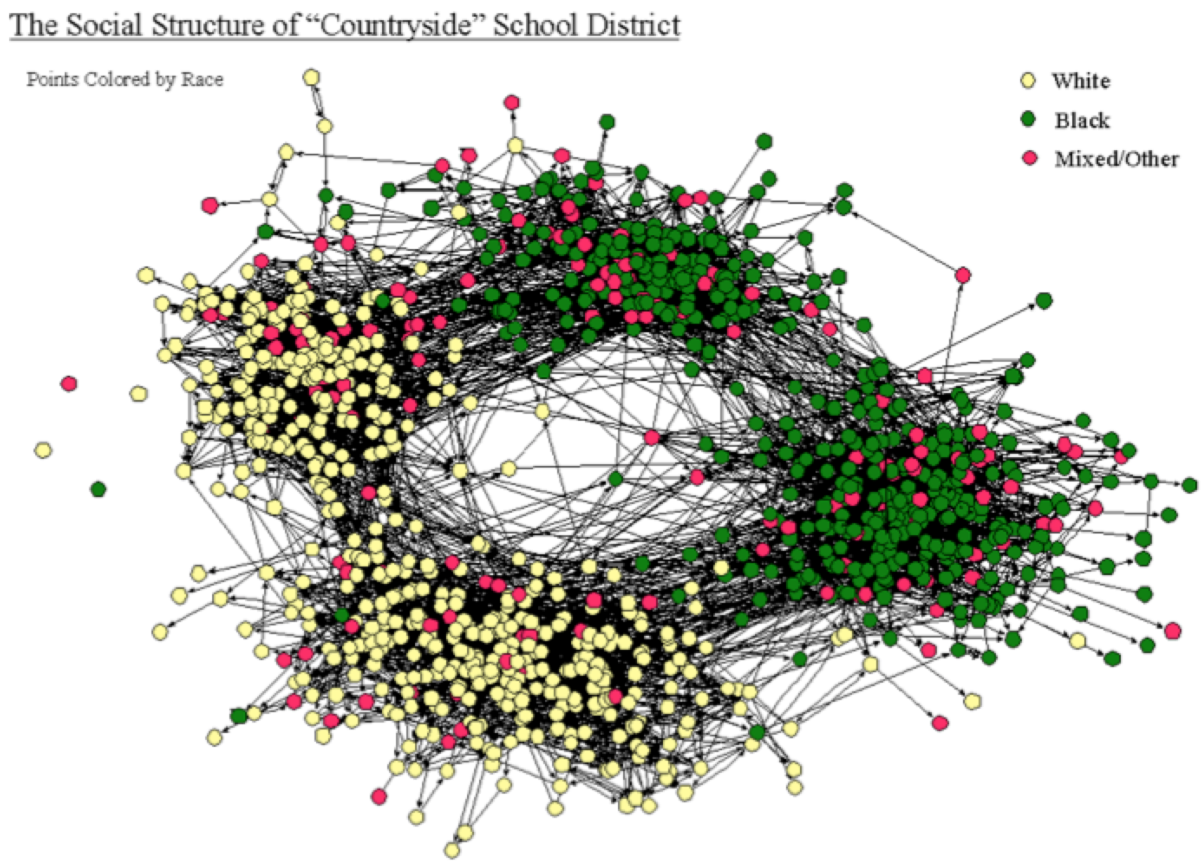


Bearman, Moody and Stovel

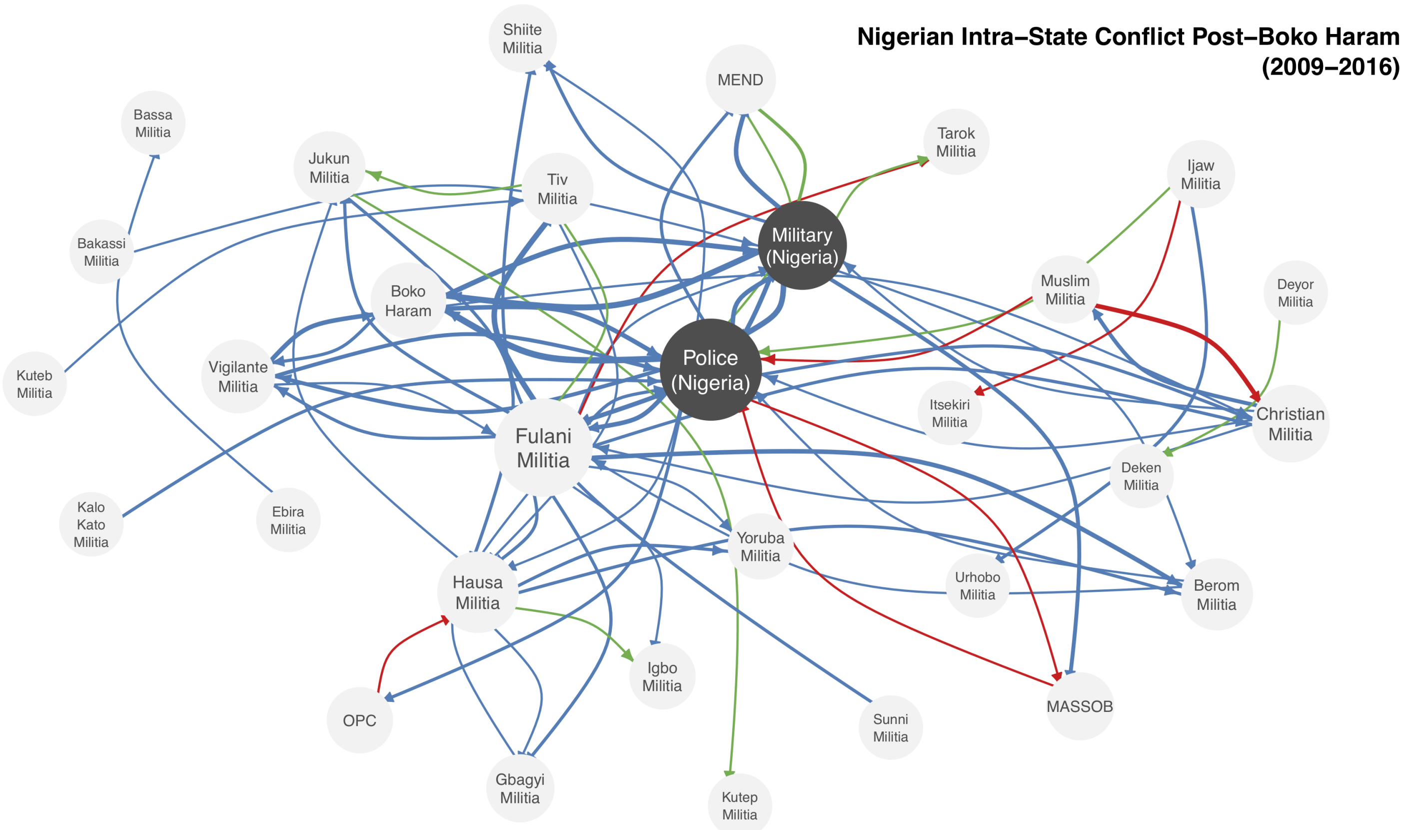
Adolescent Social Structure by **Jim Moody**



Adolescent Social Structure by **Jim Moody**



Networks of Violence in Nigeria



Networks of Violence: Predicting Conflict in Nigeria by **Dorff, Gallop, & Minhas**

Structures do hide in hairballs …

International Conflict Event Warning System (ICEWS): Material Conflict by **Minhas, Hoff, & Ward**

Processing Network Data:

Dealing with Data

You might begin with either a matrix or information stored separately about edges and nodes. This depends on your data collection strategy.

Useful terminology for working in *R*:

- Matrices: the adjacency matrix
- Edges: linkages between actors or nodes
- Vertices: nodes (or actors) in your system

How can we go beyond dyads?

How do we restructure a dyadic data frame such as alliances from COW into a matrix format?

```
data(defAlly)
head(defAlly)
```

```
##   ccode1 ccode2   ij defAlly year
## 1      2     20 2_20       1 2012
## 2      2     31 2_31       0 2012
## 3      2     41 2_41       0 2012
## 4      2     42 2_42       0 2012
## 5      2     51 2_51       0 2012
## 6      2     52 2_52       0 2012
```

Your Turn (5 min)

Can you transform `defAlly` into a matrix form as shown below?

```
##     2 20 31 41 42 51 52 53 54 55 56 57 58 60 70 80 90 91 92 93
## 2  NA  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## 20  1 NA  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## 31  0  0 NA  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## 41  0  0  0 NA  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## 42  0  0  0  0 NA  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## 51  0  0  0  0  0 NA  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## 52  0  0  0  0  0  0 NA  0  0  0  0  0  0  0  0  0  0  0  0  0
## 53  0  0  0  0  0  0  0 NA  1  1  1  1  1  1  0  0  0  0  0  0
## 54  0  0  0  0  0  0  0  1 NA  1  1  1  1  1  0  0  0  0  0  0
## 55  0  0  0  0  0  0  0  1  1 NA  1  1  1  1  0  0  0  0  0  0
## 56  0  0  0  0  0  0  0  1  1  1 NA  1  1  1  0  0  0  0  0  0
## 57  0  0  0  0  0  0  0  1  1  1  1 NA  1  1  0  0  0  0  0  0
## 58  0  0  0  0  0  0  0  1  1  1  1  1 NA  1  0  0  0  0  0  0
## 60  0  0  0  0  0  0  0  1  1  1  1  1  1 NA  0  0  0  0  0  0
## 70  0  0  0  0  0  0  0  0  0  0  0  0  0  0 NA  0  0  0  0  0
## 80  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 NA  0  0  0  0
## 90  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 NA  0  0  0
## 91  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 NA  0  0
## 92  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 NA  0
## 93  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 NA
```

What We've Learned So Far

- Network analysis has applications to many fields, from animal science to economics and political science.
- Network data are most often stored as a sociomatrix.
- Networks may contain important structures that are missed if treated as dyadic data.
- Network data are easily visualized using `igraph`.
- Convert long-form (dyadic) data into a matrix in R.