Network Analysis Reading Group (NARG)

1/26 - Terminology and Data Formats

Zachary McGee

Department of Government, University of Texas at Austin

Key Terms

Actor (node) - Social entities being studied, discrete individual units. (e.g. a person, group, department, agency, etc.)

Relational tie (edge) - The link between actors (e.g. friendship, transfer of resources, affiliation in a club, a road, etc.)

Dyad & Triad - A tie established between two (three) actors. Triadic analysis is especially important in directed networks.

Subgroup - Any subset of actors and all ties among them (e.g. House Freedom Caucus members and all of their LPAC disbursements)

Key Terms

Stars - A single actor connected to multiple other actors (sometimes with reciprocity in the ties)

Directed vs. (Undirected) - A tie has an origin and destination (or not)

Dichotomous vs. (Valued) - A relationship exists or does not (A relationship has strength or is weighted)

Structural variables - Measured on pairs of actors, these are the measurement of ties between different types of actors (e.g. friendships, transactions, etc.).

Composition variables - Actor attributes, similar to standard social science variables (e.g. race, gender, class etc.).

Key Terms - Figures via Robins (2011)

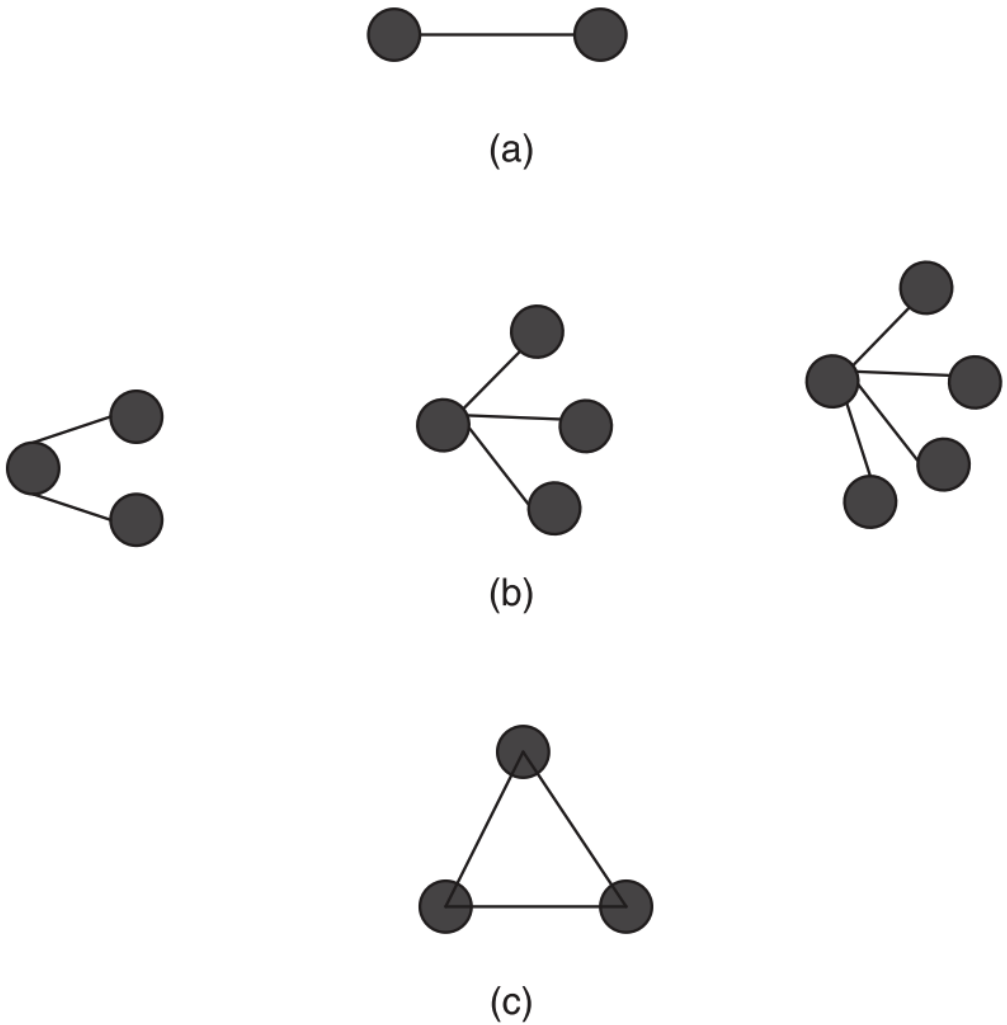


Figure 32.1 Configurations for Markov random graph model for undirected graphs (a) edge; (b) 2-star, 3-star, 4-star (higher-order star configurations may also be included); (c) triangles

Key Terms - Figures via Robins (2011)

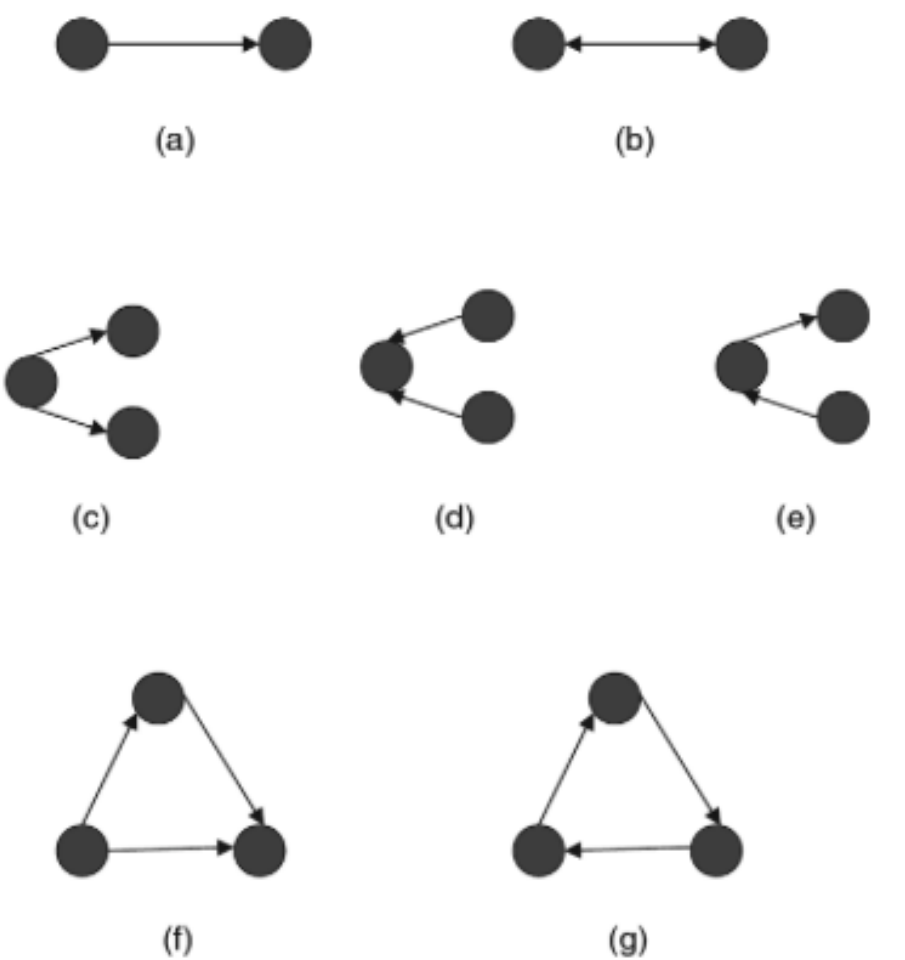


Figure 32.6 Some common Markov configurations for directed graphs. (a) Arc (b) reciprocated arc, (c) 2-outstar, (d) 2-instar, (e) 2-path, (f) transitive triad , and (g) cyclic triad

Types of Networks

One-mode Networks - All actors come from one set (e.g. nonprofits, members of a family, etc.).

Two-mode Networks - Two sets of actors that are distinct (e.g. Congressmen and their committees linked via their membership status, corporations and nonprofits linked via donations from one to the other, etc.).

Affiliation Networks - One set of actors is measured with respect to attendance at, or affiliation with, a set of events or activities (e.g. Political party donors and fundraising events, CEOs and country club memberships, etc.).

Ego-Centered Networks - Consists of a focal actor (the ego) as sets of alters who have ties to ego, and measurements on the ties among these alters (e.g. A voter and anyone he discussed the election with).

Data Formats

Incidence Matrix - A square (n x n) matrix with actors aligned to form a symmetric matrix. Relational links are binary.

Signed Matrix - A square (n x n) matrix with actors aligned to form a symmetric matrix. Relational links are -1 to 1.

Adjacency Matrix - A square (n x n) matrix with actors aligned to form a symmetric matrix. Relational links are valued.

Edge List - A two-column matrix, each row defines one edge. (e.g. Eric Cantor -> John Boehner, or, The Department of Government -> College of Liberal Arts, etc.)

Wasserman and Faust (1995) - “The SNA Bible”

Discussion:

Wasserman, Stanley, and Katherine Faust. Social Network Analysis: Methods and Applications. Cambridge University Press, 1995.

Porter et al. (2005) - Porter, Mason a, Peter J Mucha, M E J Newman and Casey M Warmbrand. 2005. “A network analysis of committees in the U.S. House of Representatives.” Proceedings of the National Academy of Sciences 102(20):7057–7062.

R Code

```
rm(list=ls())

library(igraph)
library(stringr)
library(openxlsx)

# modify your path as appropriate
# note that this a slow/inefficient function - CSV or similar format would be better
data <- read.xlsx("/Users/zacharymcgee/Dropbox/UT/Spring 2017/Misc/NARG/globalterrorismdb_061

# creating a small sample visualization
sample_edgelist <- as.matrix(head(data[,c('gname', 'country_txt')]), ncol=2)
sample_graph_obj <- graph_from_edgelist(sample_edgelist)
plot.igraph(sample_graph_obj)

# visualize attacks from top N groups
n <- 10

ordered <- unique(data$gname)
prevalence - sapply(ordered, function(x)sum(data$gname == x))
ordered <- ordered[order(prevalence, decreasing = T)]
edgelist <- as.matrix(data[data$gname %in% ordered[2:(n+1)],c('gname', 'country_txt')])

# shorten some names
to_shorten <- grep('\(', edgelist[,1])
edgelist[to_shorten, 1] <- str_match(edgelist[to_shorten, 1], '\(([A-Z]+)\)')[,2]
```

R Code

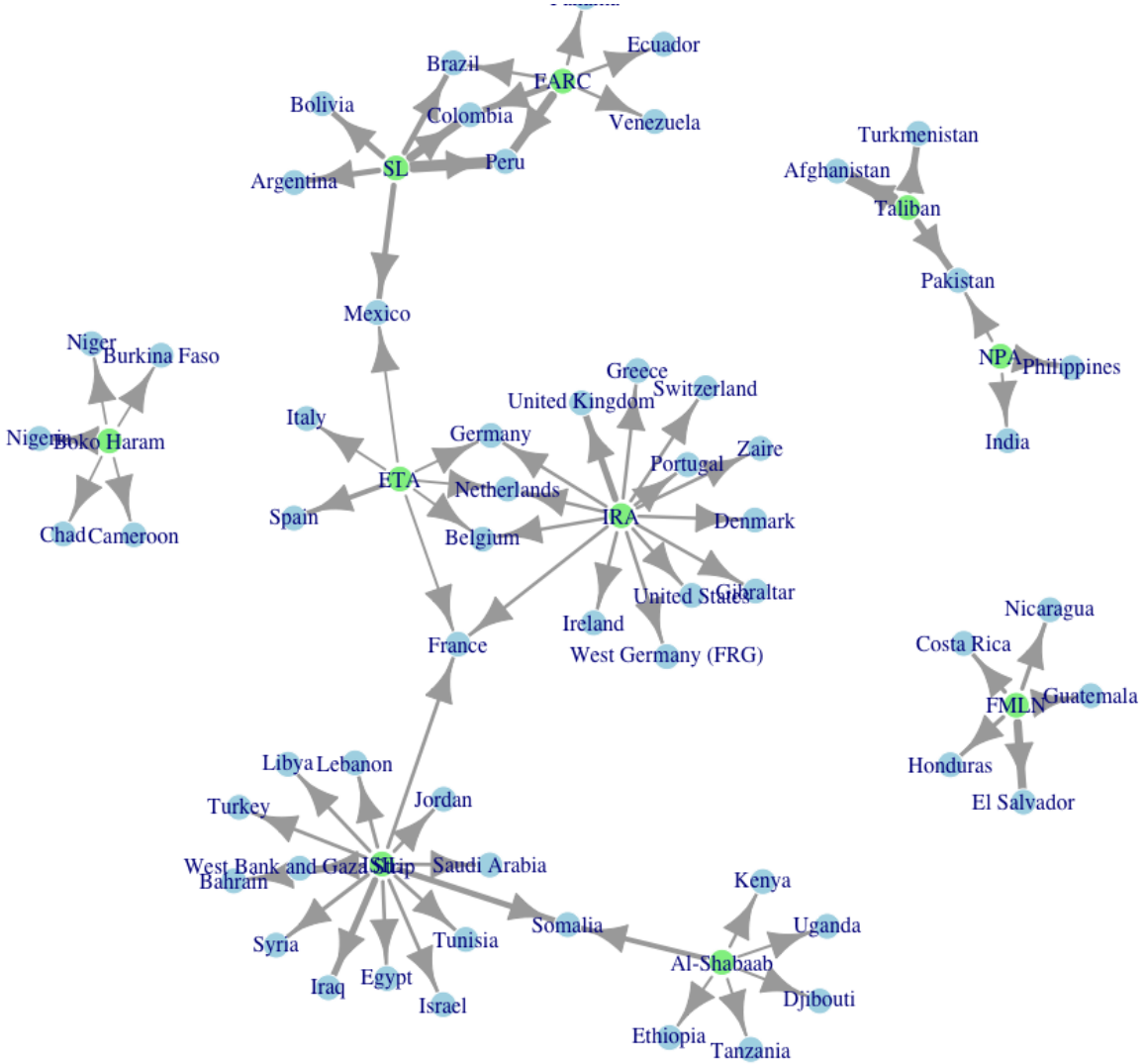
```
# collapse to a weighted graph of unique entries (could imagine separating by year, decade, et
weighted_edges <- unique(edgelist)
igraph_obj <- graph_from_edgelist(weighted_edges, directed=T)

# add a few attributes
V(igraph_obj)$type <- V(igraph_obj)$name %in% edgelist[,1]
E(igraph_obj)$weight <- apply(weighted_edges, 1, function(x)sum(edgelist == x))

# setting the edge weight for visual purposes (10 seemed good here)
edge_weight <- 10
E(igraph_obj)$weight <- edge_weight*E(igraph_obj)$weight/max(E(igraph_obj)$weight)

# plot, with some graphical parameters modified for legibility
plot.igraph(igraph_obj,
            vertex.size=5,
            vertex.frame.color='white',
            vertex.color=ifelse(V(igraph_obj)$type == 1, 'lightgreen', 'lightblue'),
            edge.width=E(igraph_obj)$weight,
            layout=layout_with_fr(igraph_obj),
            margin=-.3)
```

R Code


Ecuador
Brazil
FARC
Bolivia
Colombia
Venezuela
Peru
Argentina
SL
Mexico
Turkmenistan
Afghanistan
Taliban
Pakistan
NPA
Philippines
India
Niger
Burkina Faso
Nigeria
Boko Haram
Chad
Cameroon
Greece
Switzerland
United Kingdom
Italy
Germany
Portugal
Zaire
ETA
Netherlands
Spain
IRA
Denmark
Belgium
Gibraltar
United States
Ireland
West Germany (FRG)
France
Nicaragua
Costa Rica
FMLN
Guatemala
Honduras
El Salvador
Libya
Lebanon
Turkey
Jordan
West Bank and Gaza Strip
Saudi Arabia
Bahrain
Kenya
Somalia
Uganda
Syria
Tunisia
Al-Shabaab
Djibouti
Iraq
Egypt
Israel
Ethiopia
Tanzania

References

Porter, Mason a, Peter J Mucha, M E J Newman and Casey M Warmbrand. 2005. "A network analysis of committees in the U.S. House of Representatives." *Proceedings of the National Academy of Sciences* 102(20):7057–7062.
URL: *http://www.pnas.org/cgi/doi/10.1073/pnas.0500191102*

Robins, Garry. 2011. Exponential Random Graph Models for Social Networks. In *The SAGE Handbook of Social Network Analysis*, ed. Peter J Carrington and John Scott. SAGE pp. 484–500.

Wasserman, Stanley and Katherine Faust, eds. 1995. *Social Network Analysis*. Cambridge: Cambridge University Press.