Network Analysis using R

Bomi Lee

Outline

- Centralities
- Exponential random graph models (ERGMs)

Research Design

Network as IV

- Friendship -> Behavior
- Centrality -> Promotion

Network as DV

- Homophily -> Friendship
- Personality -> Central actors

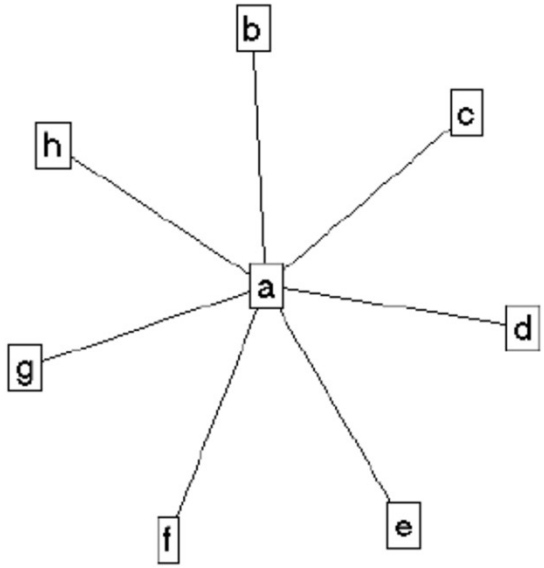
Centrality

- What does it mean for an actor to be central?

Centralities

- Degree
  - Total number of ties
  - In-degree
  - Out-degree
- Eigenvector
  - If one's alters have high degree centrality, then the focal actor has high eigenvector centrality.

Centralities



- Betweenness
  - How often an actor rests between two other actors
  - $C_B(k) = \sum \frac{\partial_{ikj}}{\partial_{ij}}, i \neq j \neq k,$

  Where ∂_{ikj} is the number of geodesics linking actors i and j that pass through node k, and ∂_{ij} is the number of geodesics linking actors k and j
    - Notes: geodesic means the shortest path in the network.

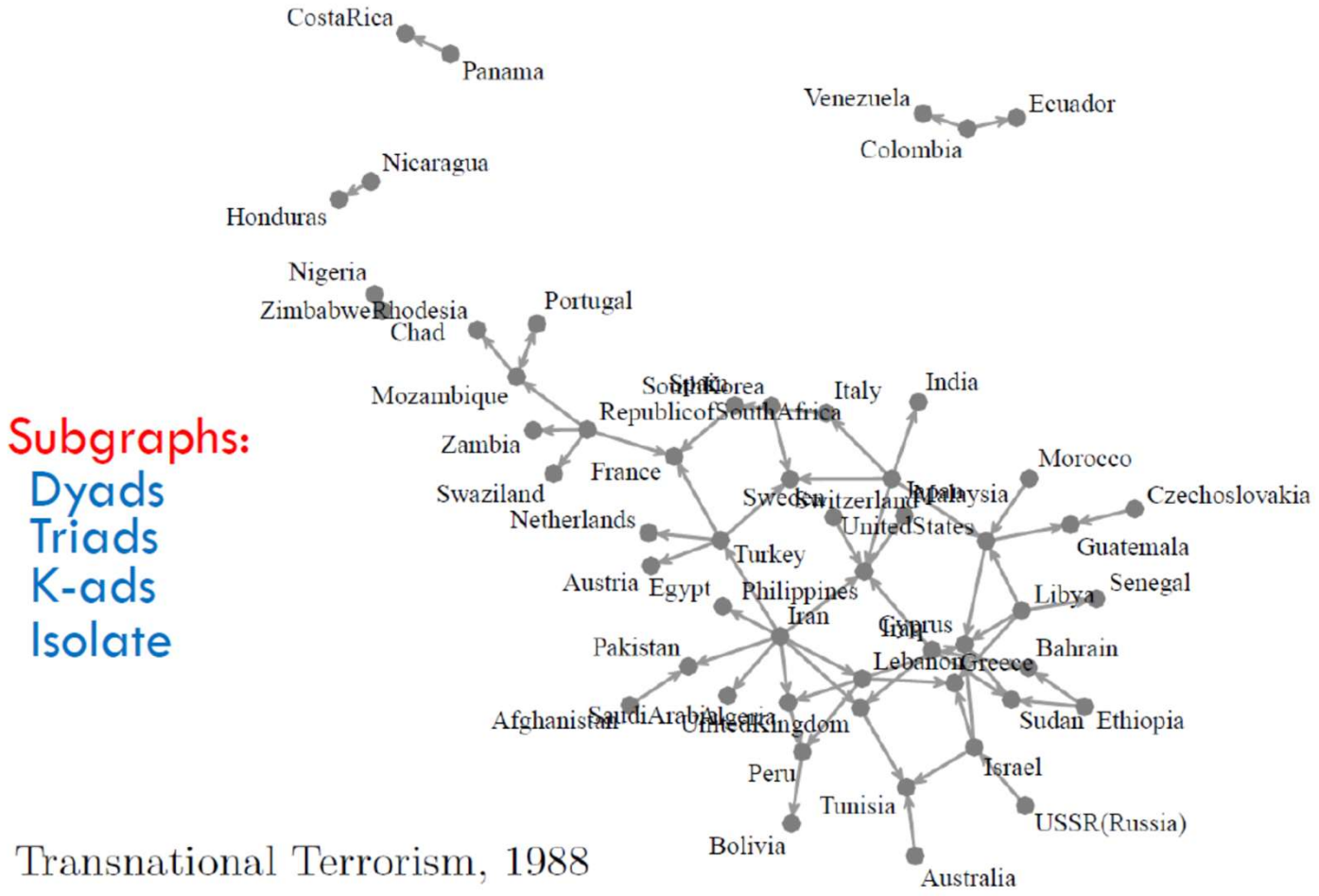
Centralities

- Closeness
  - Emphasizes an actor’s independence.
  - If an actor is not central, the actor generally needs to rely on others to relay messages through the network. Thus, an actor who is close to many other actors is a very independent actor.

R – Centralities

Subgraphs:

- Dyads
- Triads
- K-ads
- Isolate



Transnational Terrorism, 1988

Exponential random graph models (ERGMs)

- Exponential random graph models (ERGMs) are statistical models for network structure, permitting inferences about how network ties are patterned. ERGMs are tie-based models for understanding how and why social network ties arise.
- Principal goal: to understand a given observed network structure and so to obtain insight into the underlying processes that create and sustain the network-based social system.

Lusher et al. (2013)

Exponential random graph models (ERGMs)

$$P_\theta(G) = ce^{\theta_1 z_1(G) + \theta_2 z_2(G) + \cdots + \theta_p z_p(G)}$$

- The Probability of a given network G is given by a sum of network statistics (the zs) weighted by parameters (the θs) inside an exponential (and where c is a normalizing constant).
- The network statistics are counts of the number of network configurations in the given network G, local subgraphs in the network. In short, the probability of the network depends on how many of those configurations are present, and the parameters inform us of the importance of each configuration.
- A researcher specifies an ERGM by choosing a set of configurations of theoretical interest.

Lusher et al. (2013)

Exponential random graph models (ERGMs)

$$P_\theta(G) = ce^{\theta_1 z_1(G) + \theta_2 z_2(G) + \cdots + \theta_p z_p(G)}$$

- Three types of tie formation processes
  - Self-organizing network (endogenous/structural)
  - Attribute-based
  - Exogenous dyadic

Lusher et al. (2013)