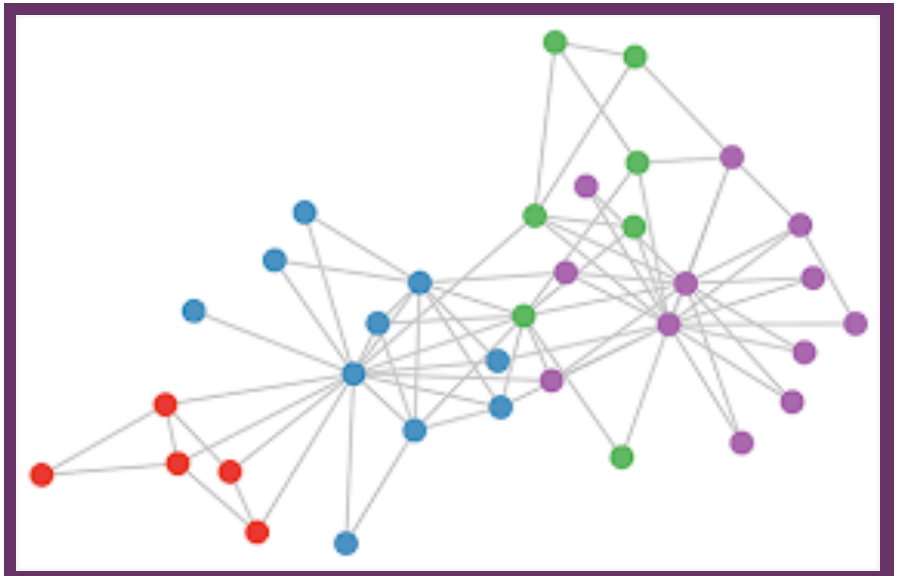


Network analysis in epidemiology

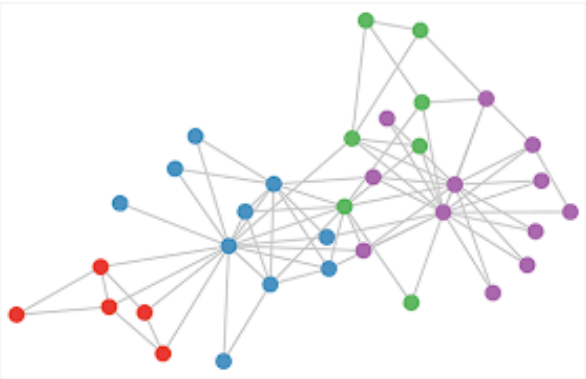
Fidisoa Rasambainarivo, DVM MSc
University of Missouri St Louis

Outline

- What is a network? Why use Network?
- How can networks be used to study disease transmission?
  - What can network analysis tell us?
    - Basic concepts
  - Further analysis from a network
  - MRQAP
- Building a network in R: Interactions between researchers at Valbio

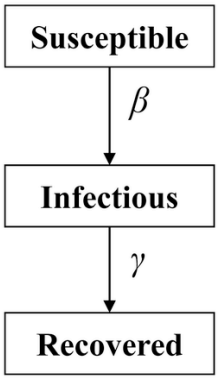
Why use a network in epidemiological studies?

- In epidemiology we want to understand:
  - Dynamics in the spreading of a disease
    - How do contagions spread in populations?
    - Will a disease become an epidemic?
    - Who to vaccinate?
    - …



Why use a network in epidemiological studies?

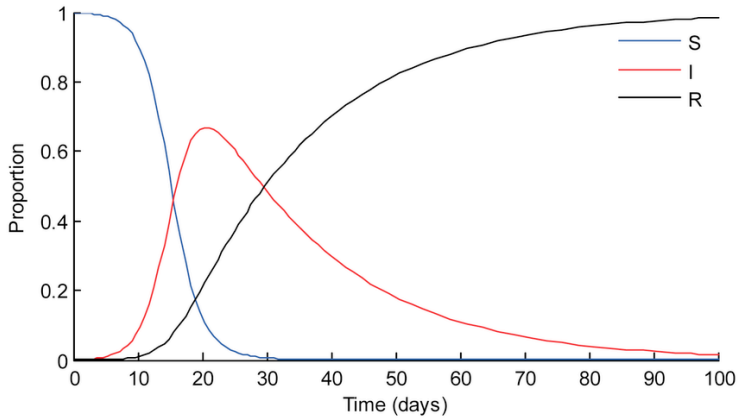
- Classic epidemiological models



$$\frac{dS}{dt} = -\beta SI$$

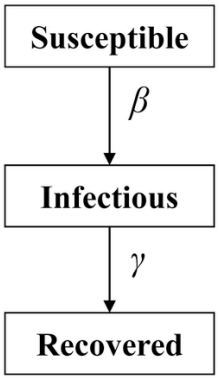
$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$



Why use a network in epidemiological studies?

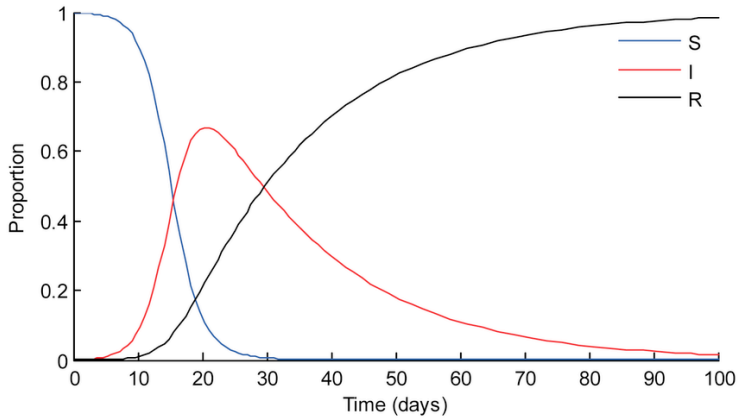
- Classic epidemiological models.



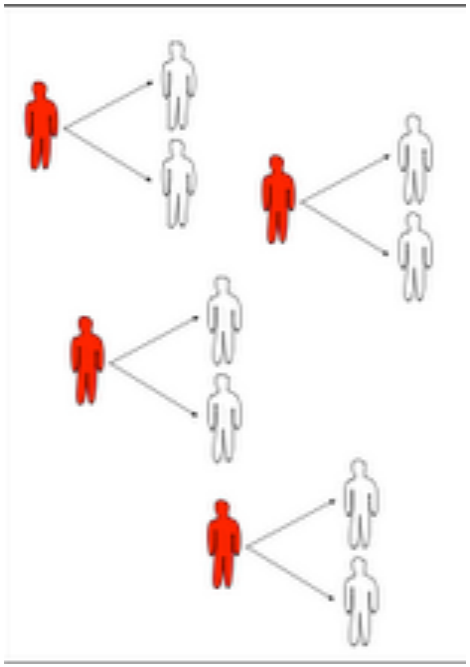
$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

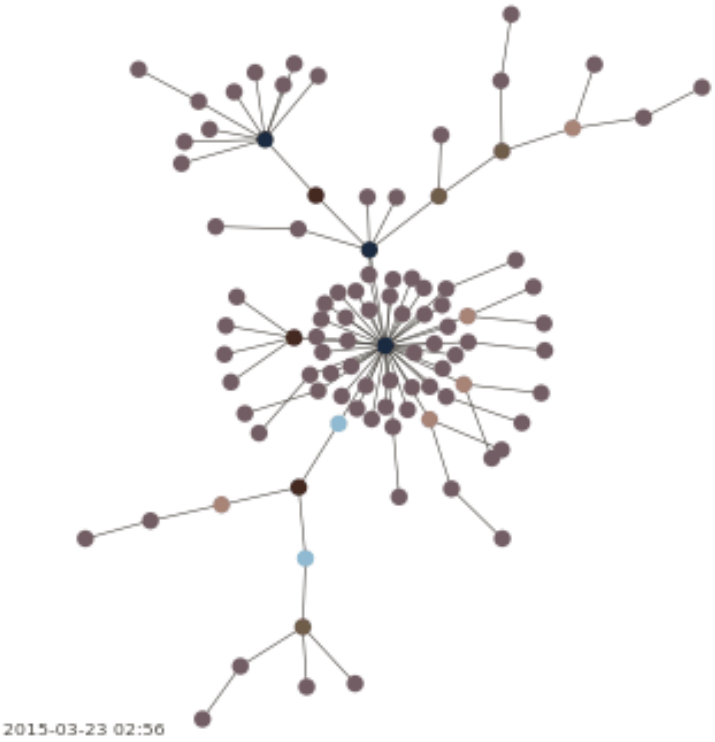
$$\frac{dR}{dt} = \gamma I$$



- Calculate R0
- Major assumption:
  - Full mixing

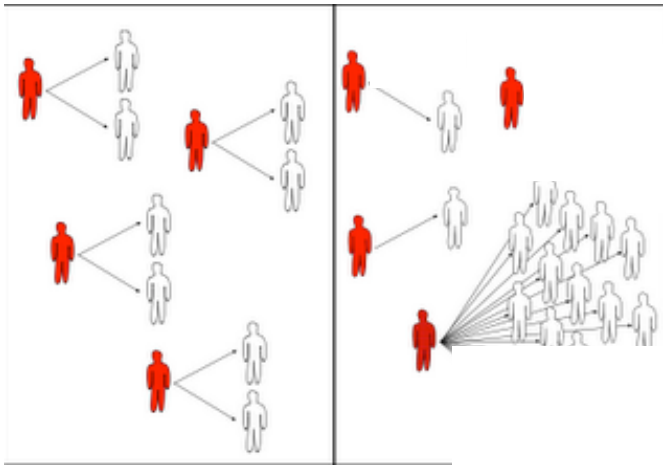


Why use a network in epidemiological studies?



Why use a network in epidemiological studies?

- In reality, heterogeneities…
  - Pareto principle: 20/80



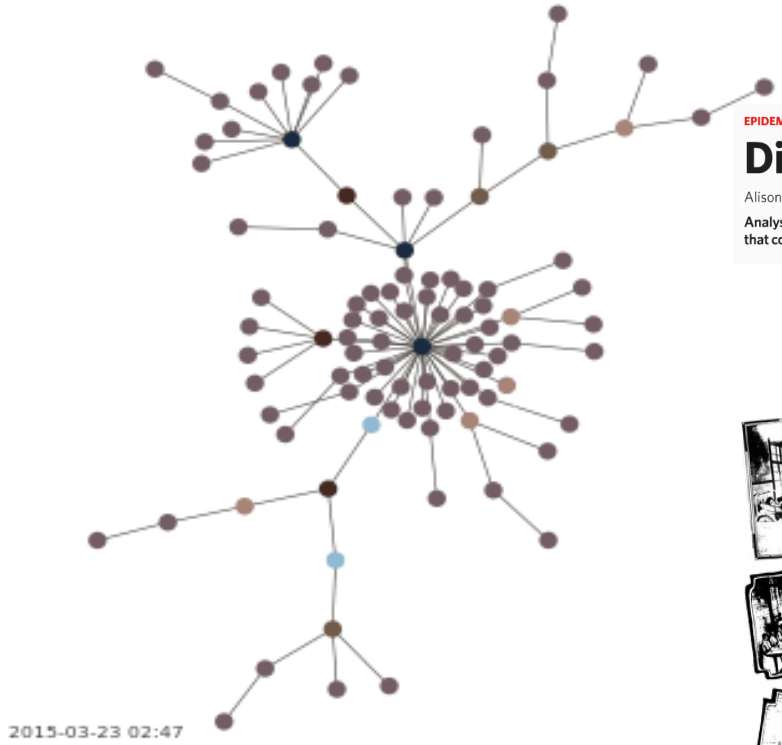
EPIDEMIOLOGY

Dimensions of superspreading

Alison P. Galvani and Robert M. May

Analyses of contact-tracing data on the spread of infectious disease, combined with mathematical models, show that control measures require better knowledge of variability in individual infectiousness.

Why use a network in epidemiological studies?



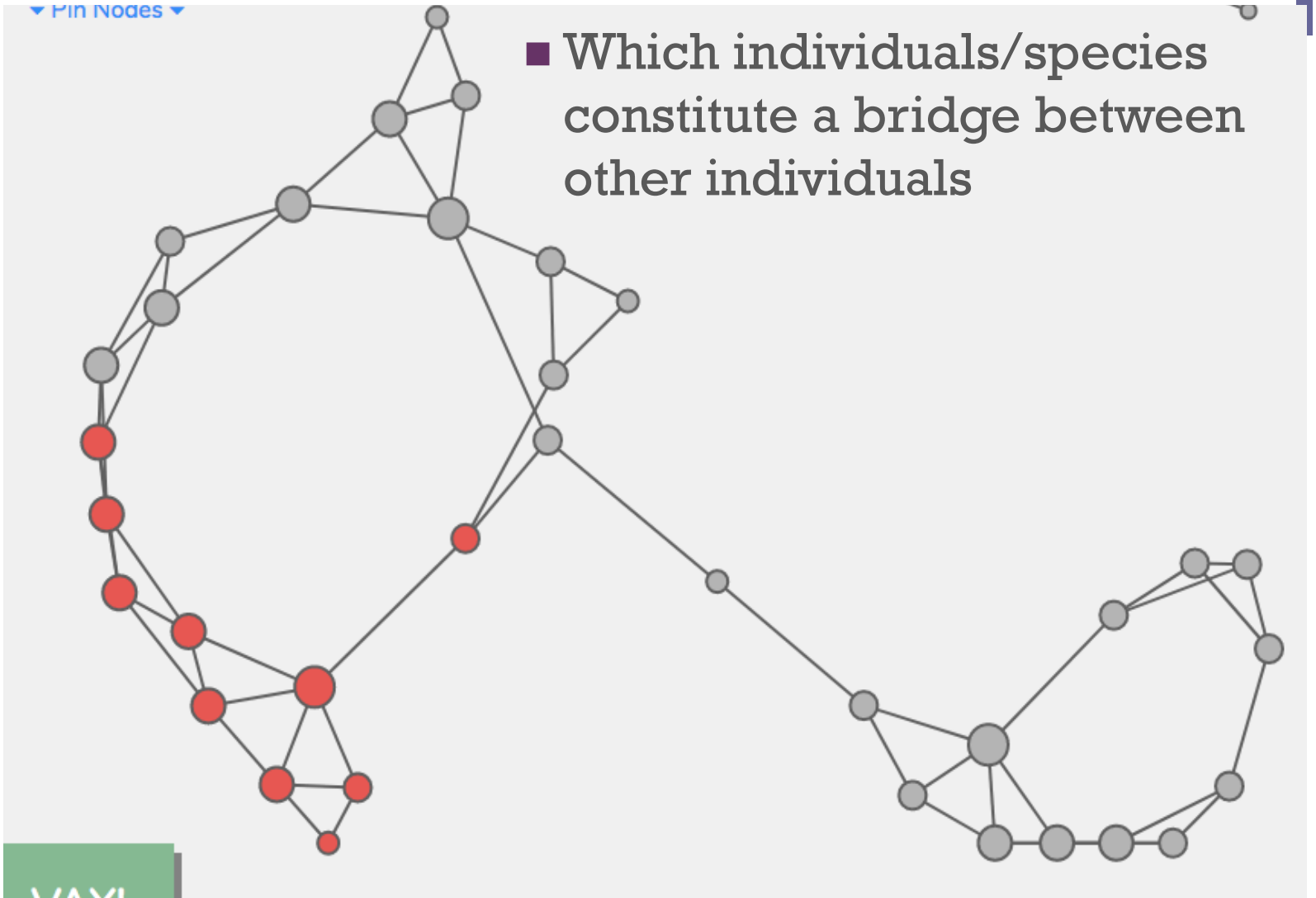
EPIDEMIOLOGY

Dimensions of superspreading

Alison P. Galvani and Robert M. May

Analyses of contact-tracing data on the spread of infectious disease, combined with mathematical models, show that control measures require better knowledge of variability in individual infectiousness.

Why use a network in epidemiological studies?



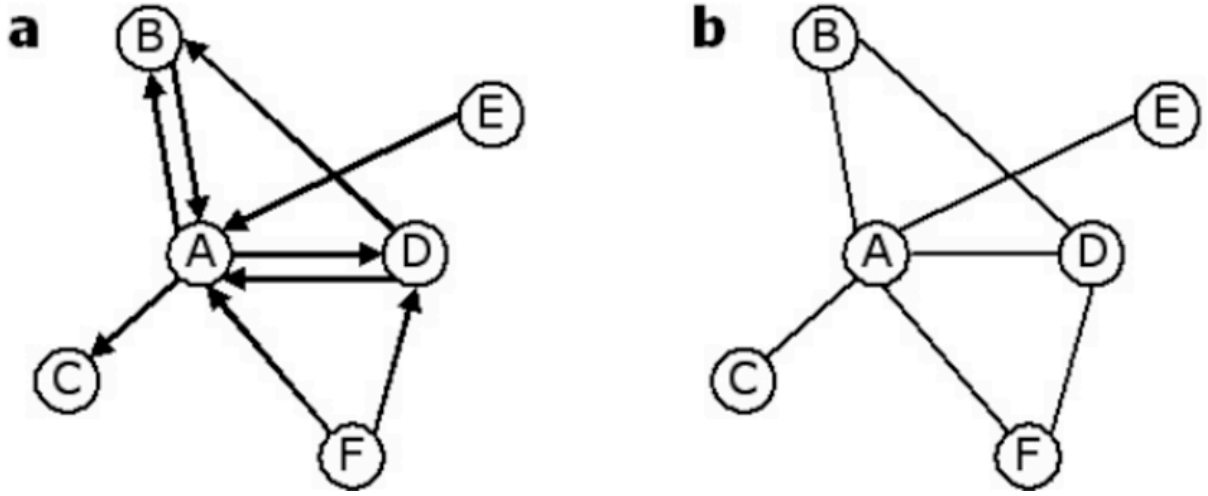
What is a network

- A (social) network: social structure made up of actors that are **interacting**
- Each actor (individual/village) is called a **vertex** (plural: vertices)

- Ties or link between two (dyad) vertices is called an **edge** and may represent sharing information, photographs, resources, space, pathogen… whatever…

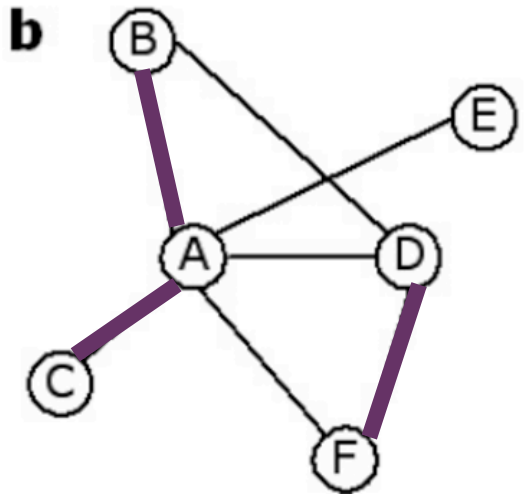
Network: basic terms and concepts

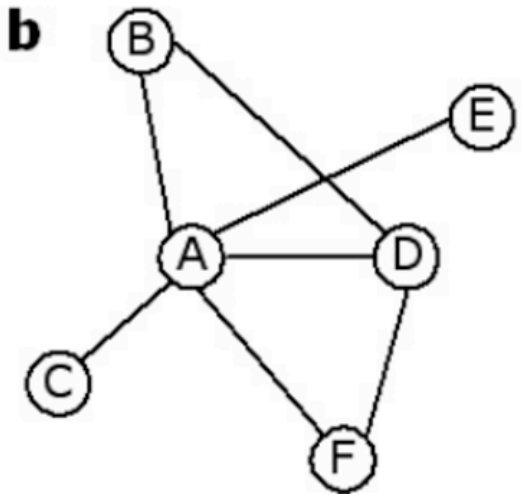
- Directed, undirected,



Network: basic terms and concepts

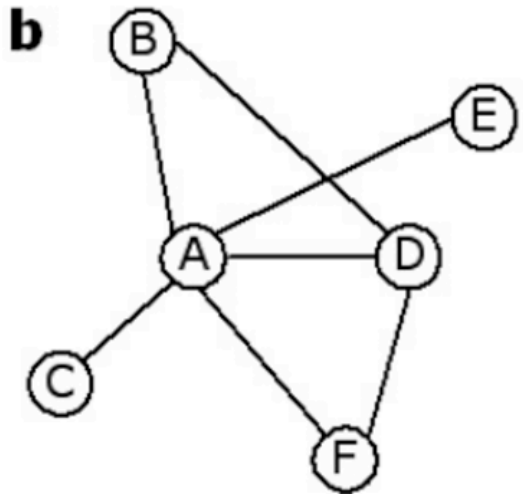
- Weighted, unweighted

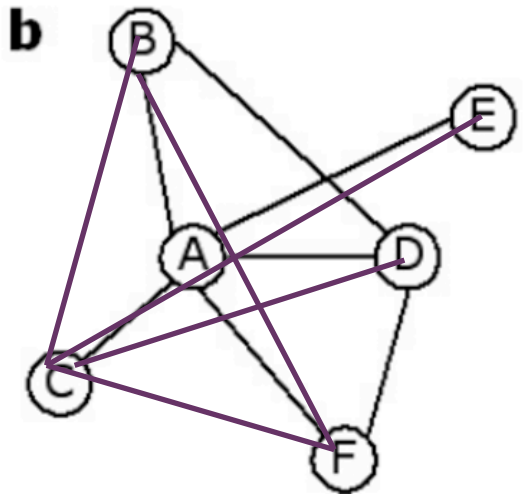




Network: basic metrics

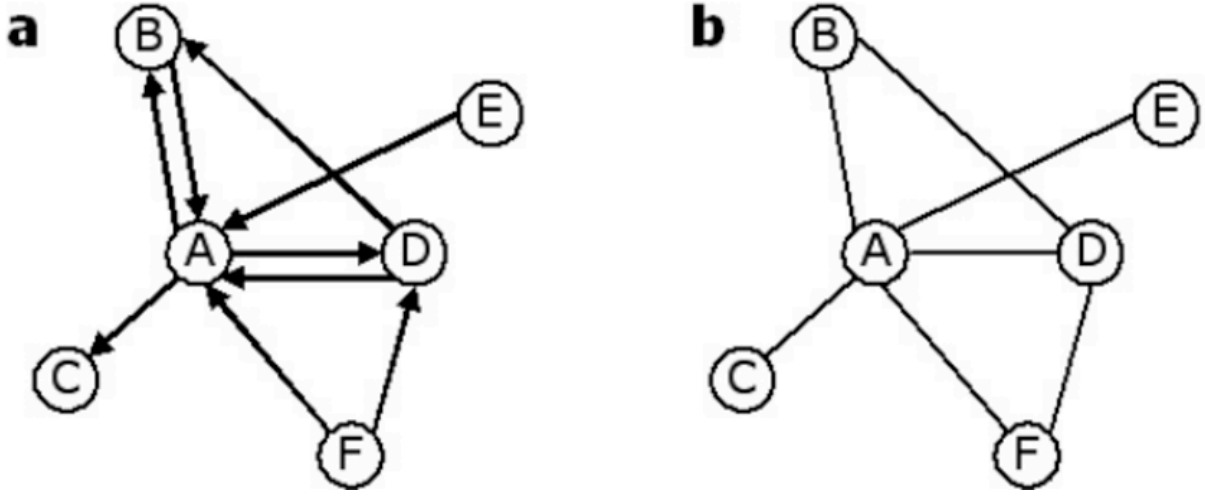
- Network density





Network: basic metrics

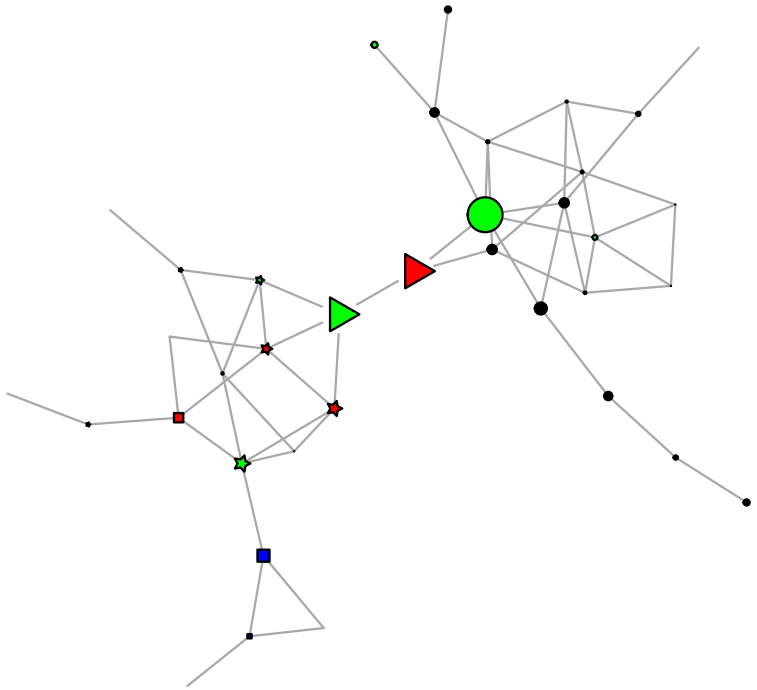
- Degree (k)



Individual	Degree
A	4
B	2
C	1
D	3
E	1
F	2

Network: basic metrics

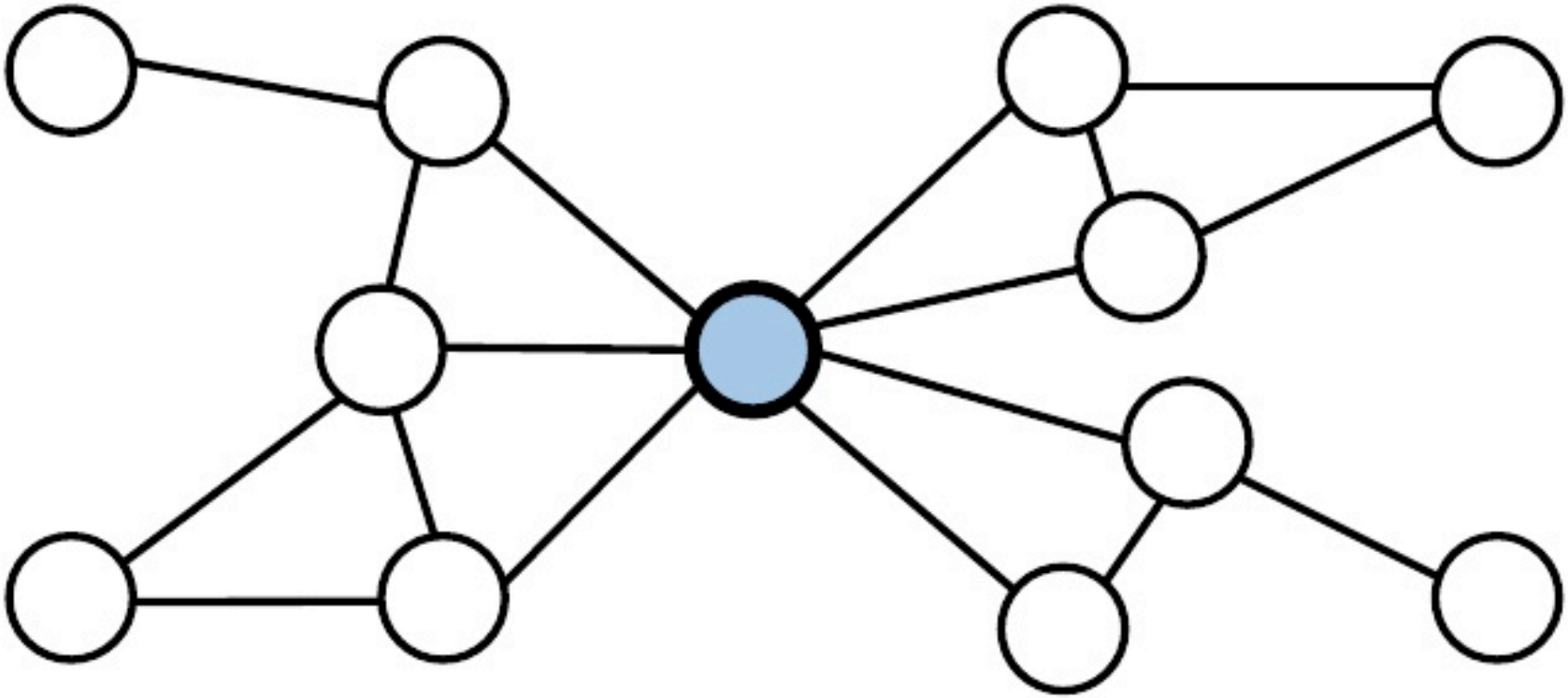
- Betweenness
  - Number of shortest paths that go through a node



$$C_B(p_i) = \sum_{j=i}^{N} \sum_{k=1}^{j-1} \frac{g_{jk}(p_i)}{g_{jk}}$$

Network basic metrics

- Closeness



To analyze and compare multiple networks…

- For example, are two individuals of the same group more likely to share another characteristic?
-

To analyze and compare multiple networks…

- For example, are two individuals of the same group more likely to share another characteristic?

- A very simple approach to find out would be to use GLM/GLMM (or, if the data were binary, to use logistic regression—or if the data were a count, to use a negative binomial, etc.)
  - E.g: If two individuals both have a or both do not have a disease, we note it 1, then measure geographic distance two individuals and we could run GLM…

To analyze and compare multiple networks…

- For example, are two individuals of the same group more likely to share another characteristic?

- A very simple approach to Network Analysis would be to use GLM/GLMM (or, if the data were binary, to use logistic regression—or if the data were a count, to use a negative binomial, etc.)

- But there is a problem…

To analyze and compare multiple networks…

- Network inherently examines relationship: Violating assumptions of independence.

To analyze and compare multiple networks…

- Network inherently examines relationship: Violating assumptions of independence.

- Many statistical social network models
  MRQAP (both static and dynamic network)

MRQAP (Quadratic Assignment Procedure)

- Multiple Regression Quadratic Assignment Procedure
  - Basically logistic regression analysis applied to matrix data.
  - Is your response variable linked to explanatory variable 1 while controlling for all other variables?

MRQAP (Quadratic Assignment Procedure)

- Multiple Regression Quadratic Assignment Procedure
  - Basically logistic regression analysis applied to matrix data.
  - Is your response variable linked to explanatory variable 1 while controlling for all other variables?

MRQAP (Quadratic Assignment Procedure)

- Essentially, what the QAP does is to "scramble" the dependent variable data through several permutations. By taking the data, and "scrambling" it repeatedly, resulting in multiple random datasets with the dependent variable—and then multiple analyses can be performed.

Working example:

- Cases of TB are threatening the Centre Valbio (cases detected in a nearby community). Public health officials launch a research project to evaluate the risks of spread of the disease in the community and help mitigate the risks.

- Specifically:
  - Dr Antso would like to know is there a group of individuals in this community that have more (or less) interactions with other individuals?
  - Because of the apparent social structure in the community, Dr Rado is interested in knowing, are individuals of the same sex, individuals from the same institution (IPM, MISA…), and individuals with similar status regarding TB vaccination more likely to interact with each other?

- 51 E2M2 participants record the amount of time they spend with another individual on a specific day.

Thank you!