Descriptive Network Analysis C

Dr Daniele Rotolo

SPRU (Science Policy Research Unit)
Business School
University of Sussex

Network Analysis
& Infographics

Week 6

Learning Outcomes

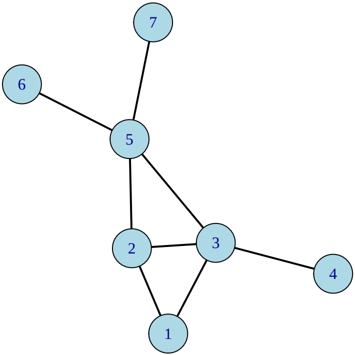
	Learning outcome	Assessment mode
1	Explain the concept of network and list the main network indicators	ESS
2	Describe and apply the major techniques for the collection of network data and their statistical analysis	ESS, GPN + GWS
3	Identify the main characteristics of networks by means of network measures	ESS, GPN + GWS
4	Employ network analysis techniques to produce network data-based infographics	GPN + GWS

Note: ESS: Essay; GPN: Group Presentation; GWS: Group Written Submission

Node-level measures

1. Centrality
   - Degree
   - Closeness
   - Betweenness
   - Centralisation
   - Bonacich's centrality
   - Weighted centrality

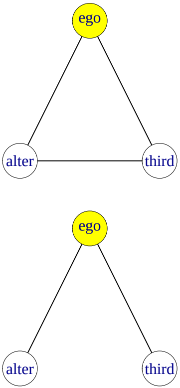
2. Brokerage
   - Brokerage roles
   - Effective network size/efficiency
   - Constraint



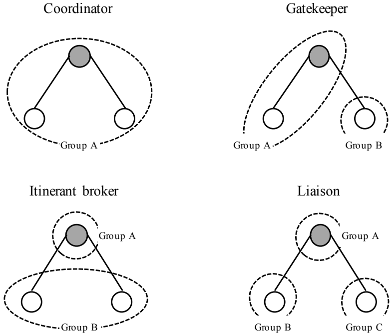
Note: We will mostly focus on undirected and unweighted networks

Node-level measures
Brokerage

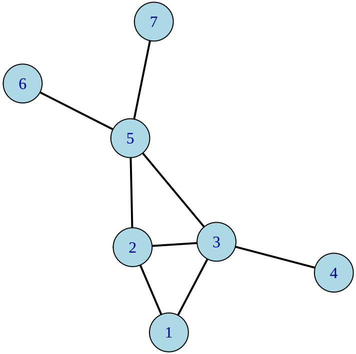
- The concept of brokerage is related to the absence of ties between the nodes of an ego network
- The absence of a tie between an *alter* and a *third* party is called structural hole
- Structural holes create tertius gaudens conditions: the *ego* has control of the spread of knowledge and resources between the *alter* and *third*



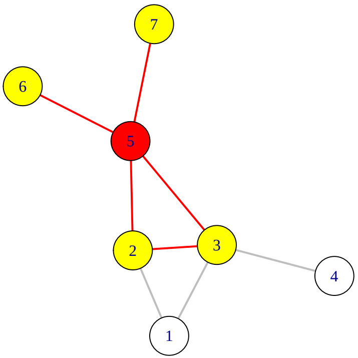
- At the level of an open triad, we can distinguish different types of brokerage roles [Gould and Fernandez, 1989]
- Roles are defined on the basis of groups to which the *ego*, *alter*, and *third* belong



- The ego-network of a node includes all open and closed triads that contain the *ego*
- For each triad we can assess the presence of structural holes
- To identify the number of triads we can count the number of possible ties, which is $(N-1)(N-2)/2$ in an undirected network
- Example: 6 triads and 5 structural holes



- The ego-network of a node includes all open and closed triads that contain the *ego*
- For each triad we can assess the presence of structural holes
- To identify the number of triads we can count the number of possible ties, which is $(N-1)(N-2)/2$ in an undirected network
- Example: 6 triads and 5 structural holes



Ronald Burt introduced two main measures of structural holes. These are based on a node's ego-network [Burt, 1992]

- Effective network size/efficiency
- Constraint

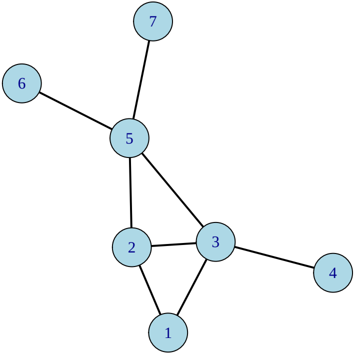
To define a node's effective network size and efficiency, we need to define the node's dyadic redundancy with its primary contacts:

$$dr_{ij} = \sum_q p_{iq} m_{jq} \qquad q \neq i, j$$

- let's define z_{ij} as the strength of the relation between node i and node j ($z_{ij} = z_{ji}$ if the network is undirected)
- p_{iq} is the proportional strength of node i's relation with q: $\frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})}$
- m_{jq} is the marginal strength of node j's relation with q: $\frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})}$
- **Interpretation**
  - dr_{ij} provides indication of how many of the actors in the ego-network of node i are also tied to node j
  - the higher is dr_{ij}, the more redundant is the tie of node i with node j

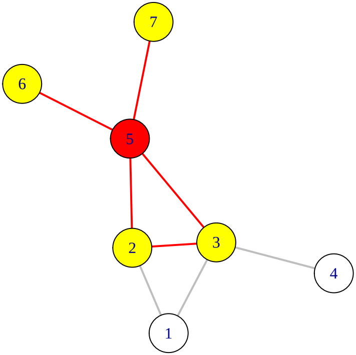
Node-level measures

Effective network size/efficiency

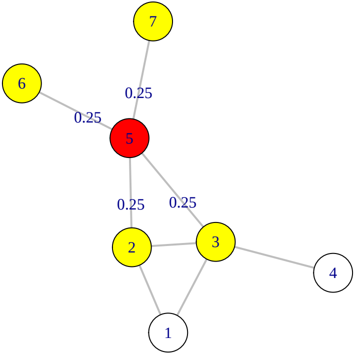


Node-level measures

Effective network size/efficiency



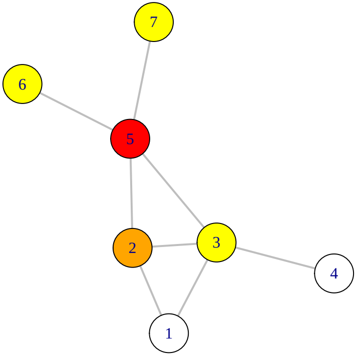
$$p_{52} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)+(1+1)+(1+1)+(1+1)]} = 0.25$$



$$p_{52} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)+(1+1)+(1+1)+(1+1)]} = 0.25$$
$$p_{53} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)+(1+1)+(1+1)+(1+1)]} = 0.25$$
$$p_{56} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)+(1+1)+(1+1)+(1+1)]} = 0.25$$
$$p_{57} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)+(1+1)+(1+1)+(1+1)]} = 0.25$$

Node-level measures

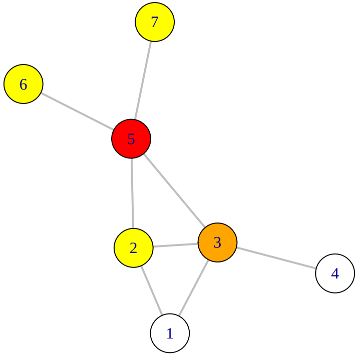
Effective network size/efficiency



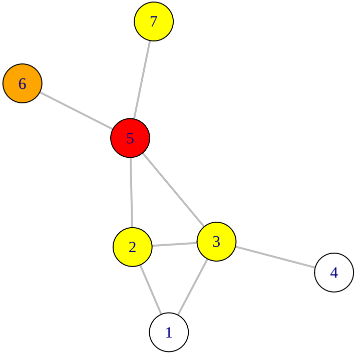
$$m_{23} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{(1+1)}{max[(1+1),(1+1),(1+1)]} = 1$$
$$m_{26} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1),(1+1),(1+1)]} = 0$$
$$m_{27} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1),(1+1),(1+1)]} = 0$$

Node-level measures

Effective network size/efficiency



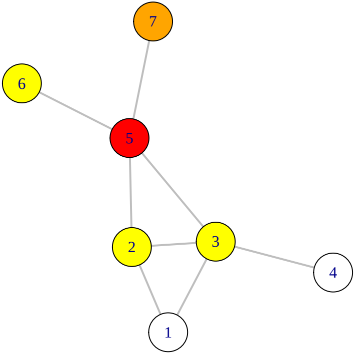
$$m_{32} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{(1+1)}{max[(1+1),(1+1),(1+1),(1+1)]} = 1$$
$$m_{36} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1),(1+1),(1+1),(1+1)]} = 0$$
$$m_{37} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1),(1+1),(1+1),(1+1)]} = 0$$



$$m_{62} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1)]} = 0$$
$$m_{63} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1)]} = 0$$
$$m_{67} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1)]} = 0$$

Node-level measures

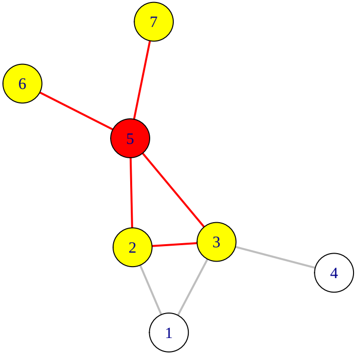
Effective network size/efficiency



$$m_{72} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1)]} = 0$$
$$m_{73} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1)]} = 0$$
$$m_{76} = \frac{z_{jq}+z_{qj}}{max(z_{jk}+z_{kj})} = \frac{0}{max[(1+1)]} = 0$$

Node-level measures

Effective network size/efficiency



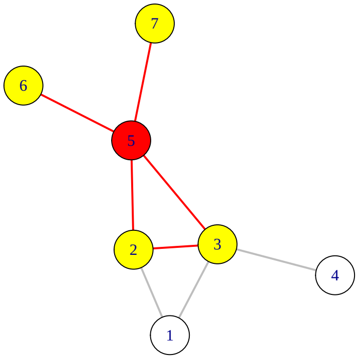
$$dr_{52} = \sum_q p_{iq} m_{jq} = p_{53} m_{23} + p_{56} m_{26} + p_{57} m_{27} = 0.25 * 1 + 0.25 * 0 + 0.25 * 0 = 0.25$$
$$dr_{53} = \sum_q p_{iq} m_{jq} = p_{52} m_{32} + p_{56} m_{36} + p_{57} m_{37} = 0.25 * 1 + 0.25 * 0 + 0.25 * 0 = 0.25$$
$$dr_{56} = \sum_q p_{iq} m_{jq} = p_{52} m_{62} + p_{53} m_{63} + p_{57} m_{67} = 0.25 * 0 + 0.25 * 0 + 0.25 * 0 = 0$$
$$dr_{57} = \sum_q p_{iq} m_{jq} = p_{52} m_{72} + p_{53} m_{73} + p_{56} m_{76} = 0.25 * 0 + 0.25 * 0 + 0.25 * 0 = 0$$

A node's effective network size is the sum of one minus all the dyadic redundancy values of the node with the other nodes in its ego-network [Burt, 1992]

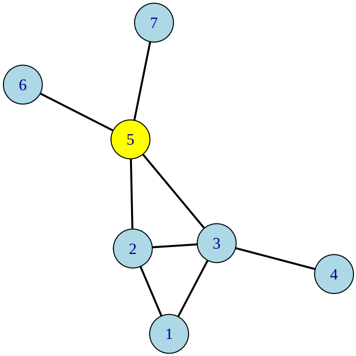
$$EN_i = \sum_j [1 - \sum_q p_{iq} m_{qj}] = \sum_j [1 - dr_{ij}] \qquad q \neq i, j$$

- EN_i ranges from 0 to $N - 1$
- EN_i / N is defined efficiency (EFF_i)

Node-level measures
Effective network size/efficiency



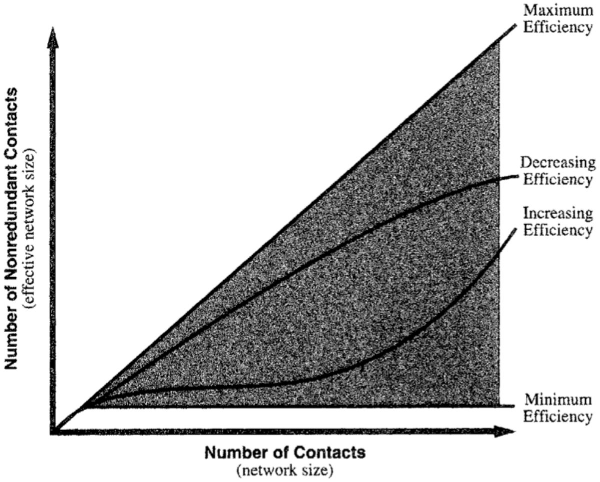
$$EN_5 = [1 - dr_{52}] + [1 - dr_{53}] + [1 - dr_{56}] + [1 - dr_{57}] = [1 - 0.25] + [1 - 0.25] + [1 - 0] + [1 - 0] = 3.5$$



n_i	$EN(n_i)$	$EFF(n_i)$
1	1.00	0.14
2	1.67	0.24
3	3.00	0.43
4	1.00	0.14
5	3.50	0.50
6	1.00	0.14
7	1.00	0.14

Node-level measures

Effective network size/efficiency



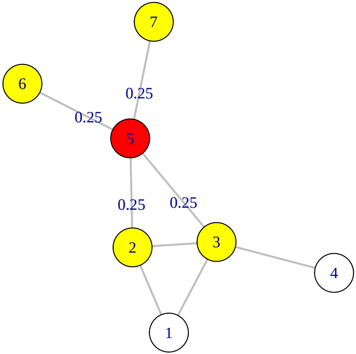
Source: [Burt, 1992]

To define a node's constraint we need to define the node's dyadic constraint with its primary contacts:

$$dc_{ij} = \left(p_{ij} + \sum_q p_{iq} p_{jq} \right)^2 \qquad q \neq i, j$$

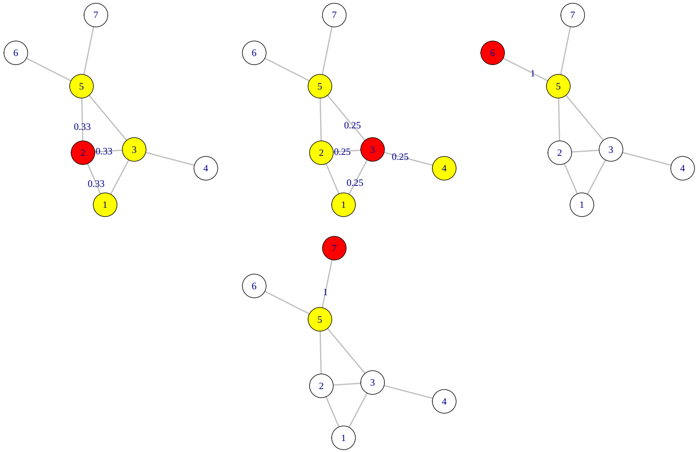
- let's define z_{ij} as the strength of the relation between node i and node j ($z_{ij} = z_{ji}$ if the network is undirected)
- p_{iq} is the proportional strength of node i's relation with q: $\frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})}$
- p_{jq} is the proportional strength of node j's relation with q: $\frac{z_{jq}+z_{qj}}{\sum_k (z_{jk}+z_{kj})}$
- dc_{ij} provides indication on the extent to which j is more or less connected to nodes to which i is strongly connected
- **Interpretation**
  - dc_{ij} provides indication of the extent to which the tie between node i and node j "constraints" node i
  - the higher is dc_{ij}, the more constrained is node i by node j
  - $dc_{ij} = 1$ when node j is the only connection of i

Node-level measures
Constraint



$$p_{5q} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)+(1+1)+(1+1)+(1+1)]} = 0.25 \qquad q = 2, 3, 6, 7$$

Node-level measures
Constraint



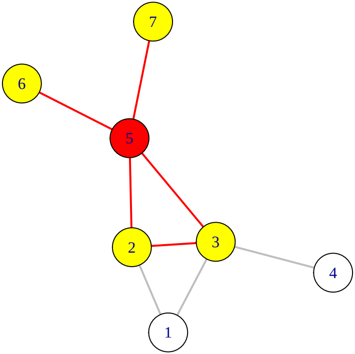
$$p_{2q} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)+(1+1)+(1+1)]} = 0.33 \qquad q = 1, 3, 5$$

$$p_{3q} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)+(1+1)+(1+1)+(1+1)]} = 0.25 \qquad q = 1, 2, 4, 5$$

$$p_{6q} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)]} = 1.00 \qquad q = 5$$

$$p_{7q} = \frac{z_{iq}+z_{qi}}{\sum_j (z_{ij}+z_{ji})} = \frac{(1+1)}{[(1+1)]} = 1.00 \qquad q = 5$$

Node-level measures
Constraint

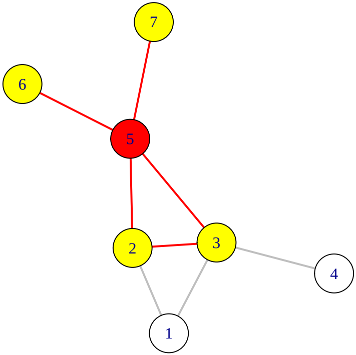


$$dc_{52} = (p_{52} + p_{53}p_{23} + p_{56}p_{26} + p_{57}p_{27})^2 = (0.25 + 0.25 * 0.33 + 0.25 * 0 + 0.25 * 0)^2 = 0.1111$$
$$dc_{53} = (p_{53} + p_{52}p_{32} + p_{56}p_{36} + p_{57}p_{37})^2 = (0.25 + 0.25 * 0.25 + 0.25 * 0 + 0.25 * 0)^2 = 0.0976$$
$$dc_{56} = (p_{56} + p_{52}p_{62} + p_{53}p_{63} + p_{57}p_{67})^2 = (0.25 + 0.25 * 0 + 0.25 * 0 + 0.25 * 0)^2 = 0.0625$$
$$dc_{57} = (p_{57} + p_{52}p_{72} + p_{53}p_{73} + p_{56}p_{76})^2 = (0.25 + 0.25 * 0 + 0.25 * 0 + 0.25 * 0)^2 = 0.0625$$

A node's constraint is the sum of all dyadic constraint values of the node with the other nodes in the ego-network [Burt, 1992]

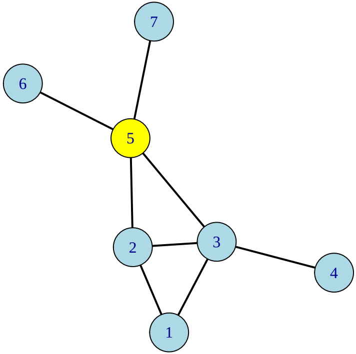
$$C_i = \sum_j dc_{ij}$$

- C_i depend on the size and density of the ego network (it tends to be higher in small and dense networks)
- The density of the ego-network is also used to assess a node's constrain



$$C_5 = \sum_j dc_{ij} = dc_{52} + dc_{53} + dc_{56} + dc_{57} = 0.1111 + 0.0976 + 0.0625 + 0.0625 = 0.3337$$

Node-level measures
Constraint



n_i	$C(n_i)$
1	0.83
2	0.69
3	0.48
4	1.00
5	0.33
6	1.00
7	1.00

Node-level measures

Constraint (example)

Country alliances

- 1816-2012 period
- 3,222 alliances
- Nodes' size proportional to nodes' constraint

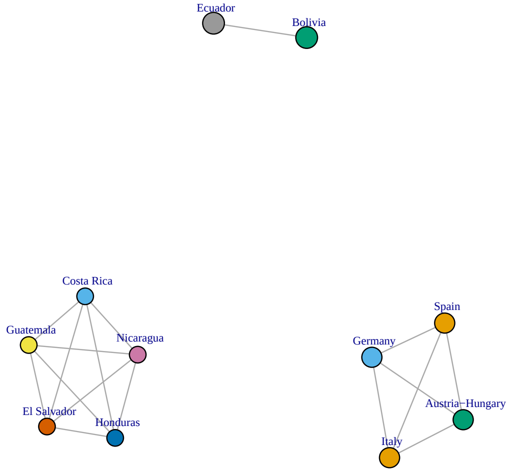
state_name1	ccode2	state_name2	dyad_st_day	dyad_st_month	dyad_st_year	dyad_end_day	dyad_end_month	dyad_end_year	left_censor	right_censor	defense	neutrality	nonaggression	entente	asymmetric
United Kingdom	235	Portugal	1	1	1816				1	1	1	0	1	0	0
United Kingdom	380	Sweden	1	1	1816	15	2	1911	1	0	0	0	0	1	0
Hanover	245	Bavaria	1	1	1838	15	3	1848	0	0	1	0	1	1	0
Hanover	245	Bavaria	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Hanover	255	Germany	1	1	1838	15	3	1848	0	0	1	0	1	1	0
Hanover	255	Germany	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Hanover	267	Baden	1	1	1838	15	3	1848	0	0	1	0	1	1	0
Hanover	267	Baden	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Hanover	269	Saxony	1	1	1838	15	3	1848	0	0	1	0	1	1	0
Hanover	269	Saxony	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Hanover	271	Wuerttemburg	1	1	1838	15	3	1848	0	0	1	0	1	1	0
Hanover	271	Wuerttemburg	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Hanover	273	Hesse Electoral	1	1	1838	15	3	1848	0	0	1	0	1	1	0
Hanover	273	Hesse Electoral	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Hanover	275	Hesse Grand Ducal	1	1	1838	15	3	1848	0	0	1	0	1	1	0
Hanover	275	Hesse Grand Ducal	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Hanover	280	Mecklenburg Schwerin	1	1	1843	15	3	1848	0	0	1	0	1	1	0
Hanover	280	Mecklenburg Schwerin	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Hanover	300	Austria-Hungary	1	1	1838	15	3	1848	0	0	1	0	1	1	0
Hanover	300	Austria-Hungary	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Bavaria	255	Germany	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Bavaria	255	Germany	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Bavaria	267	Baden	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Bavaria	267	Baden	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Bavaria	269	Saxony	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Bavaria	269	Saxony	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Bavaria	271	Wuerttemburg	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Bavaria	271	Wuerttemburg	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Bavaria	273	Hesse Electoral	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Bavaria	273	Hesse Electoral	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Bavaria	275	Hesse Grand Ducal	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Bavaria	275	Hesse Grand Ducal	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Bavaria	280	Mecklenburg Schwerin	1	1	1843	15	3	1848	0	0	1	0	1	1	0
Bavaria	280	Mecklenburg Schwerin	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Bavaria	300	Austria-Hungary	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Bavaria	300	Austria-Hungary	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Germany	267	Baden	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Germany	267	Baden	29	11	1850	15	6	1866	0	0	1	0	1	1	0
Germany	269	Saxony	1	1	1816	15	3	1848	1	0	1	0	1	1	0
Germany	269	Saxony	29	11	1850	15	6	1866	0	0	1	0	1	1	0

Source: http://www.correlatesofwar.org/

Country alliances

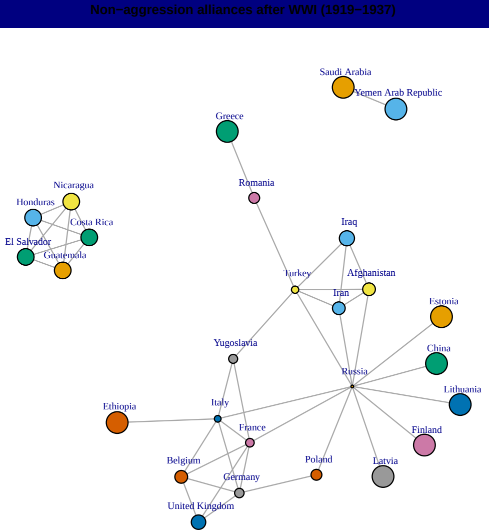
- 1816-2012 period
- 3,222 alliances
- Nodes' size proportional to nodes' constraint

Non−aggression alliances before WWI (1885−1913)



Country alliances

- 1816-2012 period
- 3,222 alliances
- Nodes' size proportional to nodes' constraint



Node-level measures

Constraint (example)

```
#1.Loading data  ----------
df <- read_csv("alliance_v4.1_by_dyad.csv") %>%
      filter(nonaggression == 1)

#2.Network Before WWI  ----------
g <-  df %>%
      filter(dyad_st_year > 1884 & dyad_st_year < 1914) %>%
      distinct(state_name1, state_name2) %>%
      graph_from_data_frame(., directed = F)

V(g)$size <- constraint(g) * 10
V(g)$label.cex <- 0.6
V(g)$color <- 1:vcount(g)
palette(gray.colors(vcount(g)))

pdf('alliances_before_wwi.pdf', width = 6, height = 4)
plot(g, layout = layout_nicely(g), vertex.label.dist = 1, vertex.label.degree = 30)
title(main = "Non-aggression alliances before WWI (1885-1913)", cex.main = 0.7)
dev.off()

#3.Network After WWI (but before WWII)  ----------
g <-  df %>%
      filter(dyad_st_year > 1918 & dyad_st_year < 1938) %>%
      distinct(state_name1, state_name2) %>%
      graph_from_data_frame(., directed = F)

V(g)$size <- constraint(g) * 10
V(g)$label.cex <- 0.5
V(g)$color <- 1:vcount(g)
palette(gray.colors(vcount(g)))

pdf('alliances_after_wwi.pdf', width = 6, height = 6)
plot(g, layout = layout_nicely(g), vertex.label.dist = 1, vertex.label.degree = 30)
title(main = "Non-aggression alliances after WWI (1919-1937)", cex.main = 0.7)
dev.off()
```

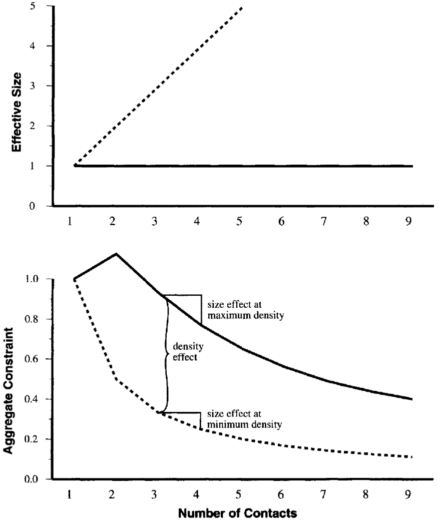


Figure 2.3 Comparisons across network size and density. (Each point on the horizontal axis refers to a different network. Solid lines describe maximum density networks. Dashed lines describe minimum density networks.)

Source: [Burt, 1992]

Measure	Interpretation
Brokerage roles	Coordinator, gatekeeper, itinerant broker, liaison
Effective network size	To what extent a node share ties with other nodes?
Constraint	To what extent a node's action is constrained by other nodes in the ego-network?

Questions

Next time ...

Next time ...

- **Seminar: Descriptive network analysis C**
  - Assessment of node-level measures (brokerage measures)
- **Lecture: Principles of infographics**
  - Principles and good practices to generate infographics
  - Network layout algorithms

References I

Burt, R. (1992).
Structural holes: The social structure of competition.
Harvard Business Press, London.

Gould, R. V. and Fernandez, R. M. (1989).
Structures of Mediation: A Formal Approach to Brokerage in Transaction Networks.
Sociological Methodology, 19(1989):89–126.