Social and Economic Network Science

Paola Tubaro & Floriana Gargiulo

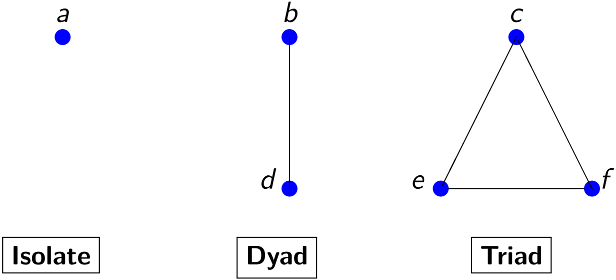
14 February 2024

Reminder: last week

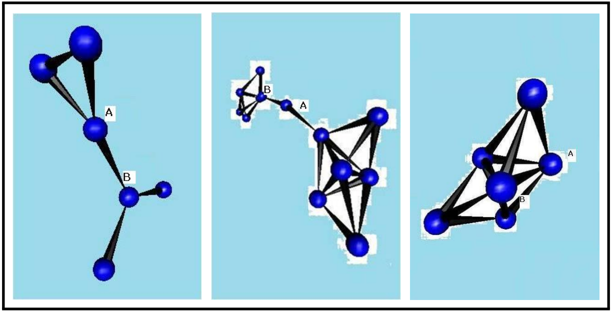
- Introduction: What is network science, and what it is used for; what is a network; introduction to graph theory
- Methods: Data formats (matrices, edgelists); directed and undirected graphs; binary and valued ties
- Exercises: Creating a graph/digraph with Networkx, adding/removing nodes and edges; importing edgelists; calculating degrees; plotting graphs

Outline for today

Isolates, dyads and triads

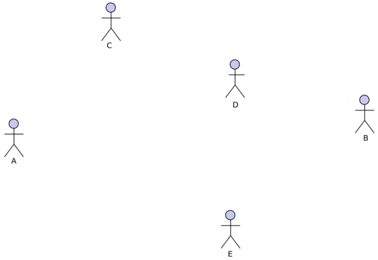


Networks are combinations of isolates, dyads and triads



Isolates

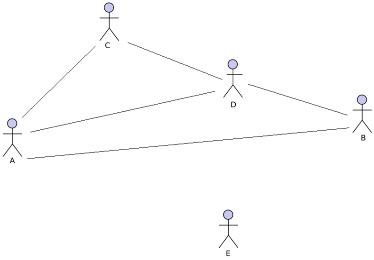
- Traditional economics is a world of isolates (Robinson)
- But in SNA, isolates are somewhat problematic to deal with
- Isolates reveal *absence* of relationships
- Sometimes, an important information to understand social systems (Phillips, 2011)



Case n. 1 : E is isolate and everyone else is isolate too

Isolates

- Traditional economics is a world of isolates (Robinson)
- But in SNA, isolates are somewhat problematic to deal with
- Isolates reveal *absence* of relationships
- Sometimes, an important information to understand social systems (Phillips, 2011)



Case n. 2 : E is isolate but all others are connected

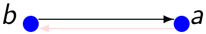
Accounting for isolates

- In practice, it is often difficult to calculate network metrics so isolates are often removed
- However, this may lead to loss of important information (Phillips, 2011)

Dyads

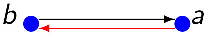
- The smallest unit allowing to observe a relationship
- Individuals can make relational decisions at this level (if a decision can be made at all)
- It is also the smallest unit allowing to observe similarities or differences in behaviour – possibly in connection with the relationship

An important dyadic mechanism: reciprocity



- A fundamental mechanism of formation, persistence or deletion of relationships
- Reduces the implicit cost of ties (in terms of time, emotional investment, self-disclosure...)
- The break of a reciprocal tie may be perceived as more 'costly' than the break of a non-reciprocal one
- Importance of reciprocity in socio-economic behaviours, eg gift exchange systems (Akerlof, 1982; Fehr & Gächter, 2000)

An important dyadic mechanism: reciprocity



- A fundamental mechanism of formation, persistence or deletion of relationships
- Reduces the implicit cost of ties (in terms of time, emotional investment, self-disclosure...)
- The break of a reciprocal tie may be perceived as more 'costly' than the break of a non-reciprocal one
- Importance of reciprocity in socio-economic behaviours, eg gift exchange systems (Akerlof, 1982; Fehr & Gächter, 2000)

Measuring reciprocity

Reciprocity index:

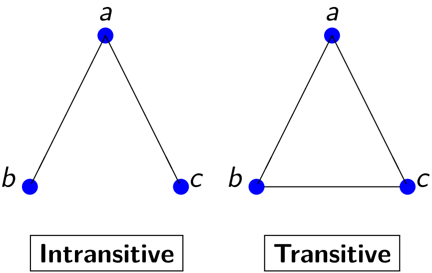
$$\frac{N\ reciprocated\ ties}{N\ ties}$$

- It is equal to 1 if for each $i - j$ tie, there is a $j - i$ tie in a directed graph.

Triads

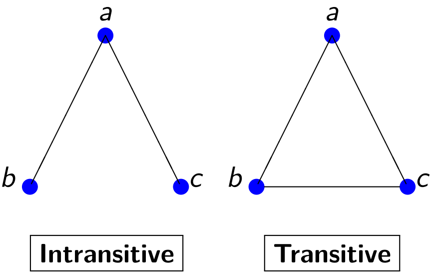
- Georg Simmel (1905): triad as a fundamental unit of sociological analysis
- Three actors in a triad may allow for social dynamics that are qualitatively different from what can be observed based on dyads or individuals
- Social behaviours and phenomena cannot be reduced to dyads or individuals
- The right starting point is the triad (and higher)

Triads and transitivity



- Intransitive: Only bilateral ties
- Transitive: A friend of my friend is my friend

Triads and transitivity



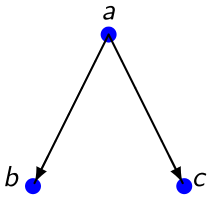
- Intransitive: Only bilateral ties
- Transitive: A friend of my friend is my friend

Transitivity in social networks

- Most empirically observed social networks exhibit a high degree of transitivity
- i.e. they have many transitive triads relative to intransitive ones
- They differ from random networks – where ties may form with the same probability between any two nodes
- Thus, transitivity is a key feature of how humans form ties: *we link to others through intermediaries.*

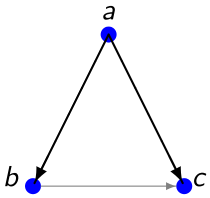
The 'forbidden triad'

- Granovetter (1973): The intransitive triad is unstable, prediction is that a tie will eventually be formed through transitivity
- It can be a 'weak' tie
- Hypothesis: if there is a strong tie between a and b, and between a and c, then c will link to b at least through weak ties

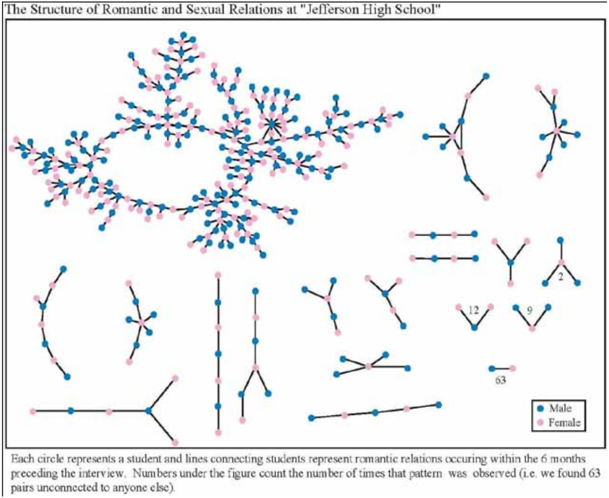


The 'forbidden triad'

- Granovetter (1973): The intransitive triad is unstable, prediction is that a tie will eventually be formed through transitivity
- It can be a 'weak' tie
- Hypothesis: if there is a strong tie between a and b, and between a and c, then c will link to b at least through weak ties



Beware: not all relations are transitive!



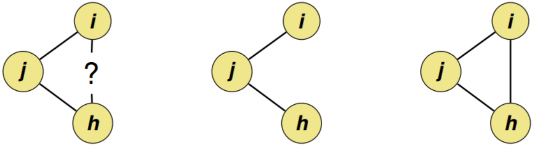
The network of romantic and sexual relationships between pupils at an American high school (Bearman et al., 2004)

Measuring transitivity

The global transitivity index of a network:

$$\frac{N \textit{ transitive triads}}{N \textit{ potentially transitive triads}}$$

- It is equal to 1 if all nodes have ties to all other nodes
- In random networks, it is often at or below 0.2
- In empirical social networks, typically between 0.3 and 0.6
- Also called **global** clustering coefficient



Potentially transitive

Intransitive

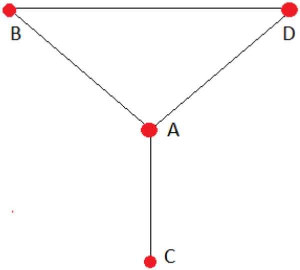
Transitive

The local clustering coefficient

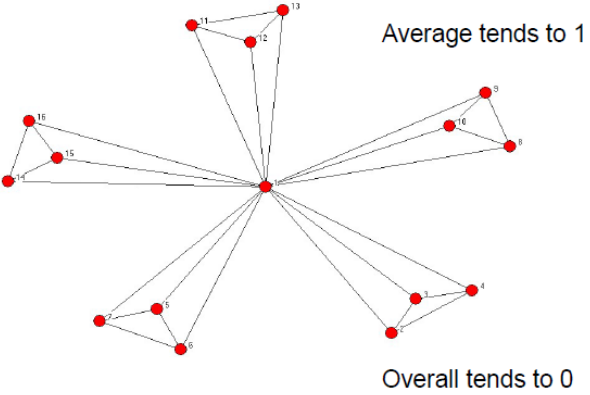
- The local clustering coefficient measures cohesion in the neighbourhood of a node (how many of i's friends are friends with each other)
- For each node i,
  $Cl_i = \dfrac{number\ existing\ ties\ between\ i's\ friends}{number\ possible\ ties\ between\ i's\ friends}$
- At the level of the whole graph, you can take the average $\sum_{i=1}^{n} \dfrac{Cl_i}{n}$

Example

- Total transitivity index (global clustering coefficient) $= 0.6$
- Local clustering coefficient:
  - Node A: 0.3333
  - Node B, node D: 1
  - Node C: 0
- Average of local CC: 0.5833



Differences in Clustering



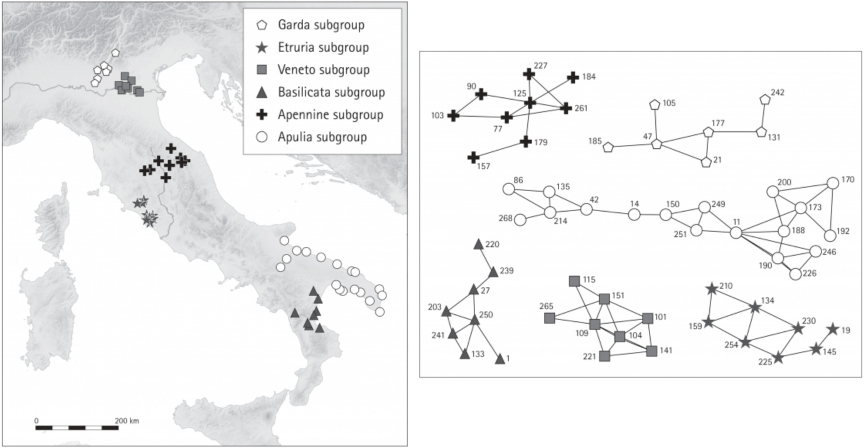
Source: (Jackson, 2008)

Why clustering matters: historical example

> *When many of the people a person knows interact with one another, [...] the members of his network tend to reach consensus on norms and they exert consistent informal pressure on one another to conform to the norms, to keep in touch with one another, and, if need be, to help one another. If both husband and wife come to marriage with such close knit networks, and if conditions are such that the previous pattern of relationships is continued, [...] both spouses will continue to be drawn into activities with people outside their own elementary family [...]. Each will get some emotional satisfaction from these external relationships and will likely demand correspondingly less of the spouse. Rigid segregation of roles will be possible because each spouse can get help from other people.*

Bott (1957)

The bronze age in Italy



(Blake, 2014)

The bronze age in Italy

- Premise: commercial relationships reveal other relationships
- Hypothesis: more relationships $\Rightarrow$ More transitivity
- Globally low transitivity (11% - 35%)
- But difference between two groups :
  - Apulia, Basilicata, Garda : 17,76% (external actors)
  - Apennines, Etruria, Veneto : 26,7 % (local actors)

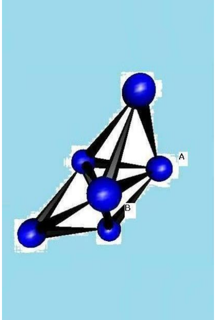
  $\Rightarrow$ Start of a collective regional identity

(Blake, 2014)

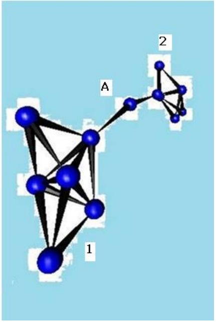
Now you know

- Networks can be seen as combinations of isolates, dyads and triads
- A key dyadic property: reciprocity
- Transitive and intransitive triads
- Transitivity and the clustering coefficient

Cohesion and connectivity



Cohesive structures: intense communication, trust, social sanctions



Connecting separate structures (Burt, 1992): control on flows of information, opportunities, power

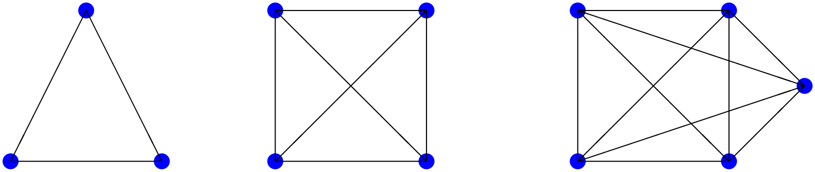
Measuring cohesion: Density

- The ratio of ties that actually exist and the ties that could exist in principle:
- Density =
  - $\frac{L}{(n*(n-1))}$ with directed ties
  - $\frac{L}{\frac{(n*(n-1))}{2}}$ with undirected ties
- where L = number of edges, n = number of nodes
- varies between 0 and 1

Density, in practice

- Density is typically low in empirical social networks, which are mostly sparse
- An indication that having ties has benefits, but also costs

Cliques



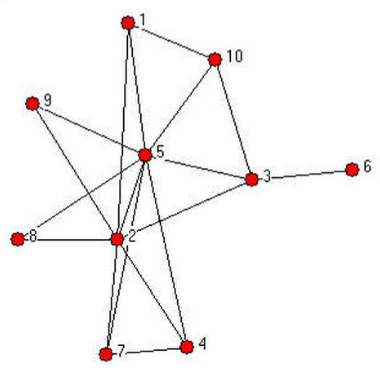
3-member clique **4-member clique** **5-member clique**

A clique is a sub-set of nodes where all possible pairs of nodes are directly connected

Cliques (cont.)

- Intuitively, the clique generalizes the triangle
- Idea is, starting from the triangle, to build outward from single ties to 'construct' larger structures
- Formally, a 'maximally complete sub-graph'

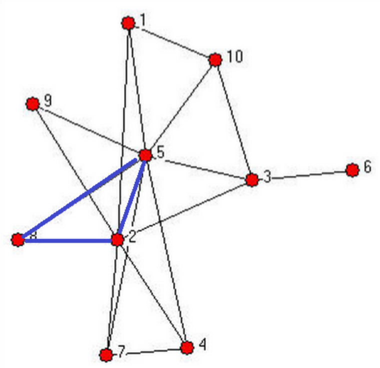
Clique analysis: overlaps



- Co-memberships and adjacency of actors
- Number of members in common

Alba & Kadushin (1976)

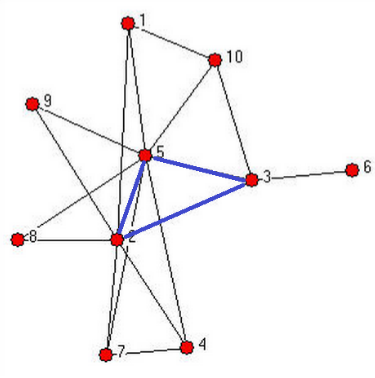
Clique analysis: overlaps



- Co-memberships and adjacency of actors
- Number of members in common

Alba & Kadushin (1976)

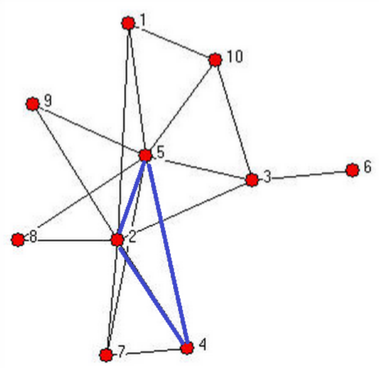
Clique analysis: overlaps



- Co-memberships and adjacency of actors
- Number of members in common

Alba & Kadushin (1976)

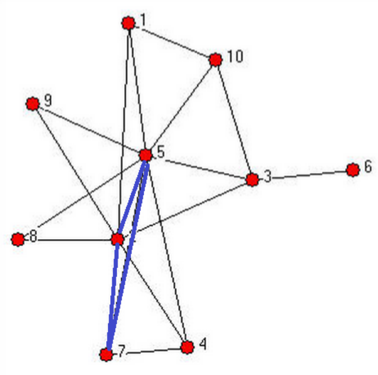
Clique analysis: overlaps



- Co-memberships and adjacency of actors
- Number of members in common

Alba & Kadushin (1976)

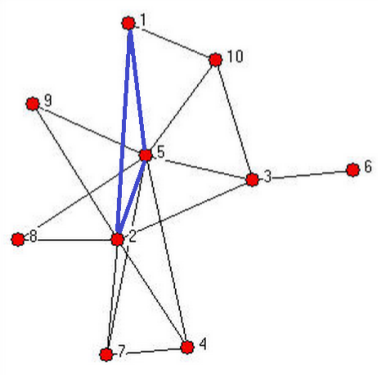
Clique analysis: overlaps



- Co-memberships and adjacency of actors
- Number of members in common

Alba & Kadushin (1976)

Clique analysis: overlaps



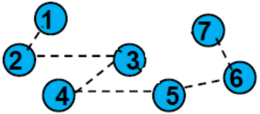
- Co-memberships and adjacency of actors
- Number of members in common

Alba & Kadushin (1976)

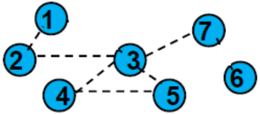
Measuring connectivity: preliminary definitions

- Walk: a sequence of vertices and edges, where each edge's endpoints are the preceding and following vertices in the sequence
- A path is a walk if each node appears at most once
- Geodesic: The shortest path between two nodes
- Cycle: a walk where the same node is the starting and ending point
- Connected graph: a walk/path exists between each pair of nodes

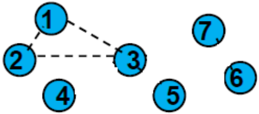
Paths, Walks, Cycles...



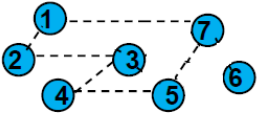
Path (and a walk) from 1 to 7:
1, 2, 3, 4, 5, 6, 7



Walk from 1 to 7 that is not a path:
1, 2, 3, 4, 5, 3, 7



Simple Cycle (and a walk) from 1 to 1: 1, 2, 3, 1



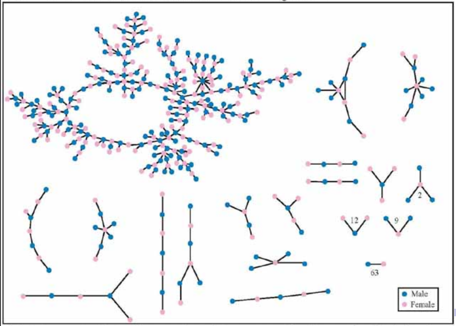
Cycle (and a walk) from 1 to 1:
1, 2, 3, 4, 5, 7, 1

Adapted from: (Jackson, 2008)

Measuring connectivity: components

- A graph is connected if a path or walk exists between any pair of nodes
- Component: a sub-graph in which any two vertices are connected by a path/walk, and are connected to no other vertices

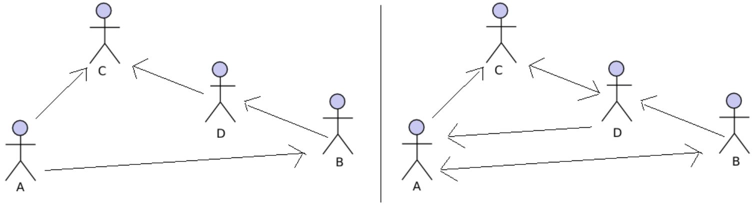
The Structure of Romantic and Sexual Relations at "Jefferson High School"



Each circle represents a student and lines connecting students represent romantic relations occuring within the 6 months preceding the interview. Numbers under the figure count the number of times that pattern was observed (i.e. we found 63

'Weakly' and 'strongly' connected components

- In a directed graph, a component is weakly connected if a path or walk exists between any two nodes (regardless of direction of ties)
- A component is strongly connected if a path or walk exists between any two nodes following the direction of ties



Left: weak connectivity; right: strong connectivity

Measuring connectivity: Distance and Diameter

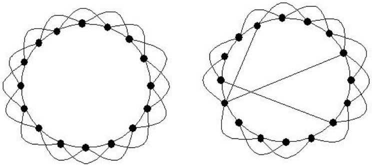
- (Geodesic) distance: number of steps from one member to another (taking shortest paths)
- Connected nodes have distance 1
- Diameter: longest geodesic distance between any two nodes in a network

Diameter and A.P.L.

- Diameter: longest distance between any two nodes in a network
- Average path length, A.P.L.: average distance between all pairs of nodes (less sensitive to outliers than Diameter)
- Shorter paths in a network involve quicker flows of information

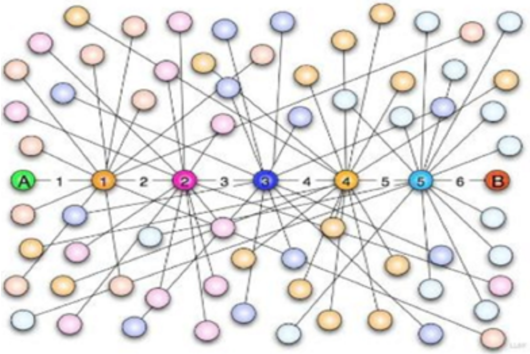
A way to use this: small worlds

- Most nodes are unconnected to each other
- Yet most nodes can be reached from every other in few steps
- E.g. two strangers connected through a mutual acquaintance



Left: Longer paths; Right: Shorter paths.

Milgram's "Six degrees of separation" ..

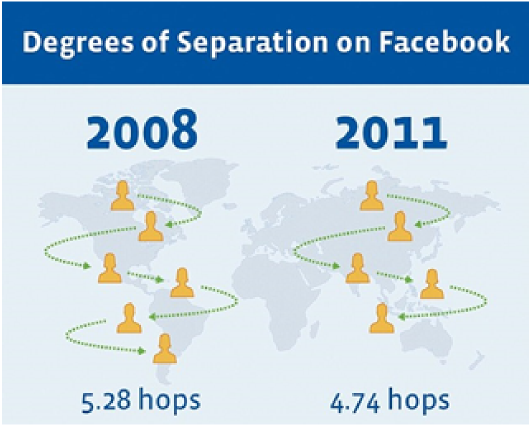


(Milgram, 1967)

Why is this surprising?

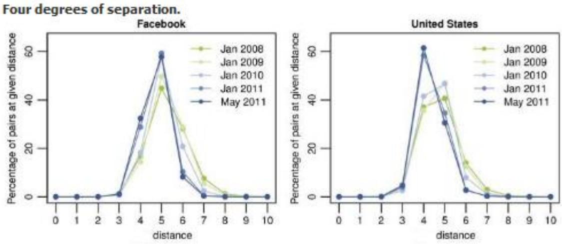
- How many friends does the average person have?
- How many friends of friends?
- How does (local) clustering play a role?

Facebook's "distance 4"...



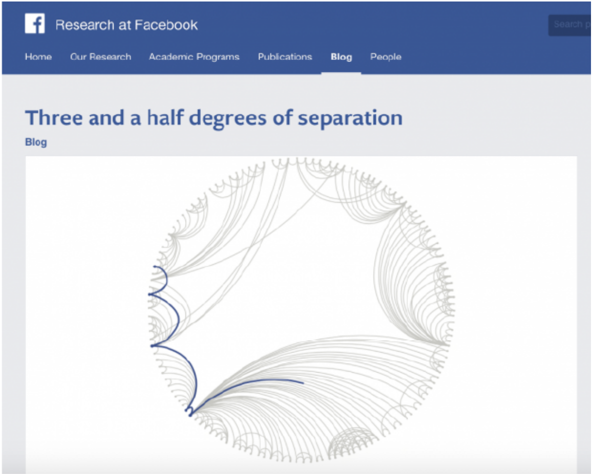
(Backstrom et al., 2011)

... More precisely, 4.74!



(Ugander et al., 2011)

… and 3.57 in 2016



Bagat et al., Three and a half degrees of separation, 2016

Small world network

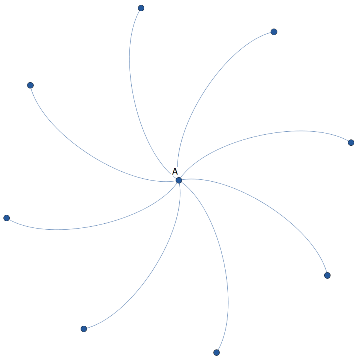
- High clustering coefficient
- Low average path length

Now you know

- Density measures global cohesion of a network
- Cliques generalize the triangle
- Different measures of global connectivity: Components, diameter, APL

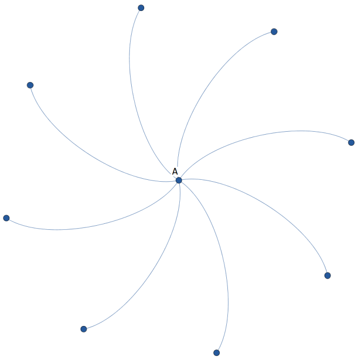
Activity, power and status through centrality concepts in networks

- Who are the most “important” actors in a network?
- Intuitively, A is the most important actor in a “star” network
- Is anyone important in a “circle” network?
- How to identify the most important actor(s), more generally?



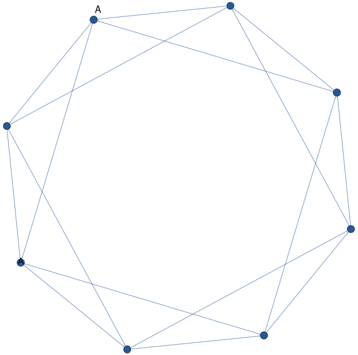
Activity, power and status through centrality concepts in networks

- Who are the most "important" actors in a network?
- Intuitively, A is the most important actor in a "star" network
- Is anyone important in a "circle" network?
- How to identify the most important actor(s), more generally?



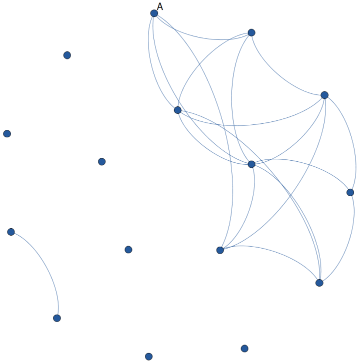
Activity, power and status through centrality concepts in networks

- Who are the most "important" actors in a network?
- Intuitively, A is the most important actor in a "star" network
- Is anyone important in a "circle" network?
- How to identify the most important actor(s), more generally?



Activity, power and status through centrality concepts in networks

- Who are the most "important" actors in a network?
- Intuitively, A is the most important actor in a "star" network
- Is anyone important in a "circle" network?
- How to identify the most important actor(s), more generally?



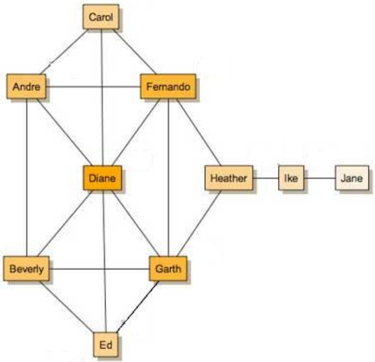
Four main answers

- Degree centrality
- Betweenness centrality
- Eigenvector centrality
- Closeness centrality

Degree centrality

- Who are the most "active" nodes?
- Diane has the highest number of direct connections (degree)
- A connector, or hub
- For actor i, it is equal to:

$$C_D(i) = \sum_{j=1}^{n} x_{ij} = \sum_{j=1}^{n} x_{ji}$$

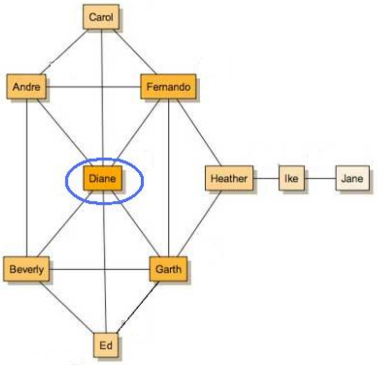


Krackhardt's kite network

Degree centrality

- Who are the most "active" nodes?
- Diane has the highest number of direct connections (degree)
- A connector, or hub
- For actor i, it is equal to:

$$C_D(i) = \sum_{j=1}^{n} x_{ij} = \sum_{j=1}^{n} x_{ji}$$

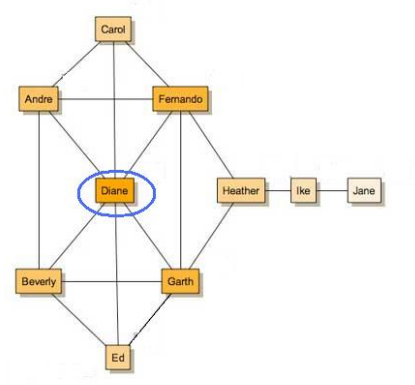


Krackhardt's kite network

Degree centrality

- Who are the most "active" nodes?
- Diane has the highest number of direct connections (degree)
- A connector, or hub
- For actor i, it is equal to:

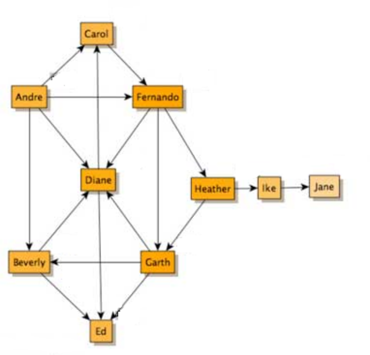
$$C_D(i) = \sum_{j=1}^{n} x_{ij} = \sum_{j=1}^{n} x_{ji}$$



Krackhardt's kite network

Indegree and outdegree centrality

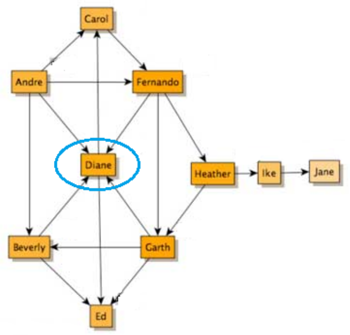
- If a network is directed, $\sum_{j=1}^{n} x_{ij}$ may differ from $\sum_{j=1}^{n} x_{ji}$
- Thus, distinguish:
  - Indegree centrality: the number of incoming ties $C_{In}(i) = \sum_{j=1}^{n} x_{ji}$ 1
  - Outdegree centrality: the number of outgoing ties $C_{Out}(i) = \sum_{j=1}^{n} x_{ij}$
- Here, Diane has indegree = 4, but outdegree = 2



Krackhardt's kite network

Indegree and outdegree centrality

- If a network is directed, $\sum_{j=1}^{n} x_{ij}$ may differ from $\sum_{j=1}^{n} x_{ji}$
- Thus, distinguish:
  - Indegree centrality: the number of incoming ties $C_{In}(i) = \sum_{j=1}^{n} x_{ji}$ 1
  - Outdegree centrality: the number of outgoing ties $C_{Out}(i) = \sum_{j=1}^{n} x_{ij}$
- Here, Diane has indegree $= 4$, but outdegree $= 2$



Krackhardt's kite network

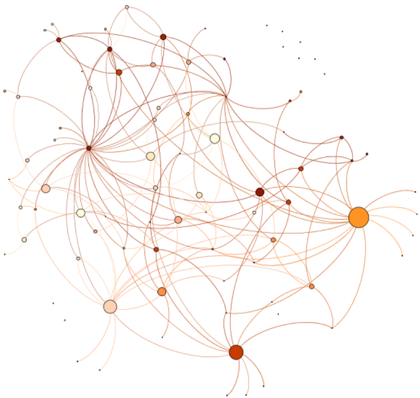
Example: a friendship network

- Who are your friends in this room?
- You will nominate some persons they may or may nor nominate you back
- In this case, indegree = number of nominations you receive: measure of *prestige* or *popularity*
- Instead, outdegree = number of friends that you nominate: measure of *expansiveness* or *activity*
- A very expansive person (high outdegree) may not be a very *popular* person (high indegree)!

Example: In-outdegree, power and hierarchy in an intra-organisational advice network

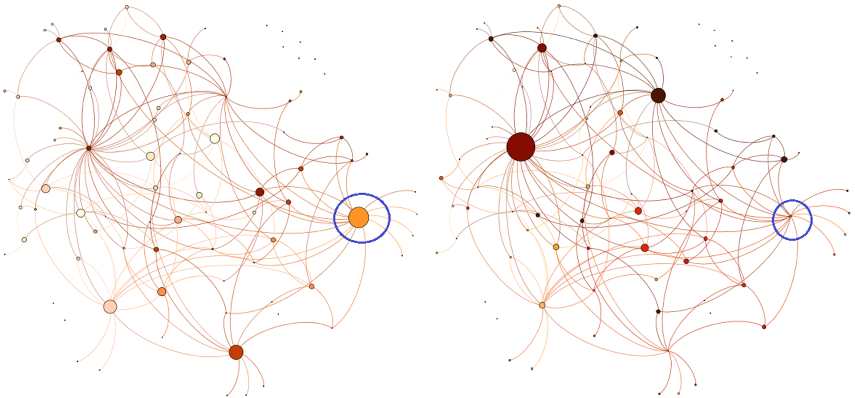
- Advice ties between judges in a court in Paris
- Advice network = information and knowledge circulation in an organisation
- High indegree = receiving many requests for advice
  - a sign of knowledge (asking the most knowledgeable)
  - a sign of tenure (ask the most senior)
  - a sign of power (ask those high up in the ladder)

Outdegree vs Indegree



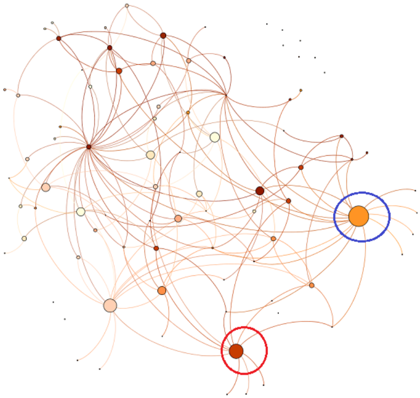
Wave 1: Judges with highest outdegree (left) and highest indegree (right): not the same! (Lazega et al., 2012)

Outdegree vs Indegree



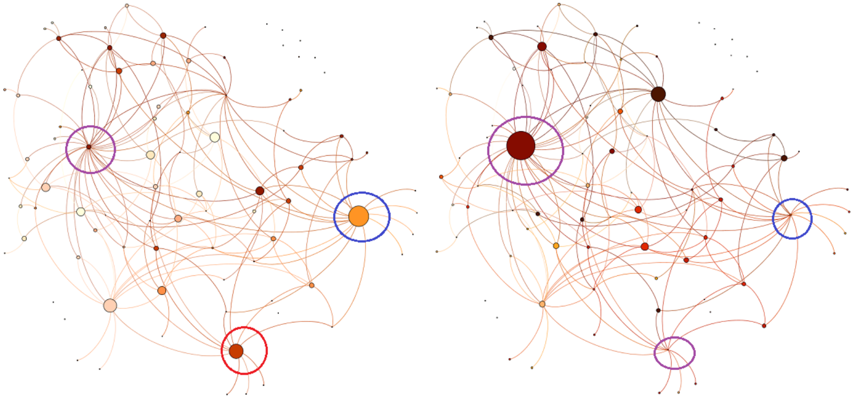
Wave 1: Judges with highest outdegree (left) and highest indegree (right): not the same! (Lazega et al., 2012)

Outdegree vs Indegree



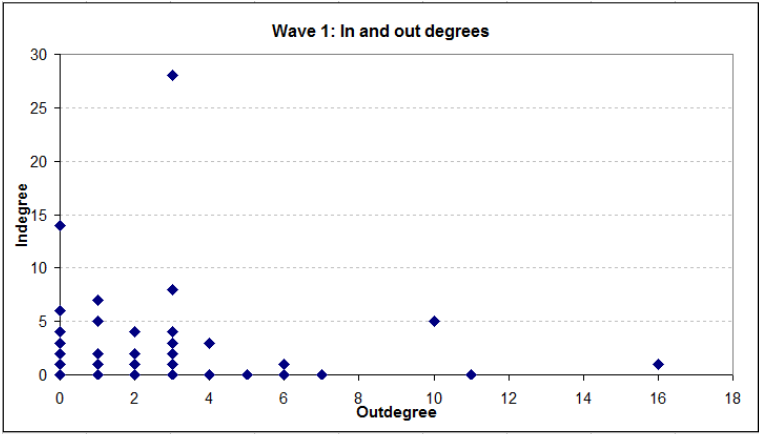
Wave 1: Judges with highest outdegree (left) and highest indegree (right): not the same! (Lazega et al., 2012)

Outdegree vs Indegree



Wave 1: Judges with highest outdegree (left) and highest indegree (right): not the same! (Lazega et al., 2012)

Outdegree vs Indegree



Wave 1: Judges with highest outdegree are not those with highest indegree (Lazega et al., 2012)

Indegree, seniority and hierarchy

- The more senior receive more requests for advice
- A sign of hierarchy in the institution
- Hierarchy largely based on tenure
- A way through which organisational structure affects knowledge-sharing and information circulation

Outdegree among judges

- High outdegree = seeking advice from many people
- It may mean:
  - a sign of lack of knowledge (asking in order to learn)
  - a sign of lack of tenure (seek guidance from others' experience)
  - perhaps, a sign of involvement and activity (engaging others in discussions)

Outdegree and lack of power

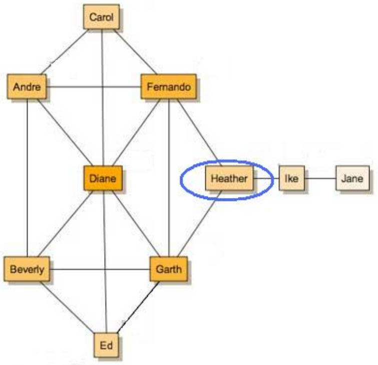
- The less senior send out more requests for advice
- Not much a sign of activity, but rather of lack of knowledge/power
- Confirms hierarchy in the institution, largely based on tenure
- Suggests that information circulates vertically in this organisation: from those higher up the ladder, to those below

A technical note: comparing different networks

- Centrality comparisons are meaningful only if networks are of the same size: indegree = 8 is different with 10 actors (you are tied to 80% of the network) or 40 actors (20% only)
- To ensure comparability across different networks, we adjust for network size (*normalization*)
- Normalized Degree Centrality: $C'_D(i) = \frac{\sum_{j=1}^{n} x_{ij}}{n-1}$
- Normalized Indegree Centrality: $C'_{In}(i) = \frac{\sum_{j=1}^{n} x_{ji}}{n-1}$
- Normalized Outdegree Centrality: $C'_{Out}(i) = \frac{\sum_{j=1}^{n} x_{ij}}{n-1}$

Betweenness centrality

- Heather has fewer connections than Diane
- Yet she occupies a strategic position, between different parts of the network
- She controls what flows in the network



Krackhardt's kite network

What we need betweenness centrality for

- Betweenness centrality identifies the gatekeepers within the network
- These actors have more paths running through them, allowing them to pass (or restrict or block) information to others in the network
- They are sometimes in the position of tertius gaudens, the third who benefits from priviledged access to multiple sources of information
- Thus, betweenness centrality is often a more reliable measure of “importance" in a network

Actors with higher betweenness centrality

- Connect different groups within a network, and are sometimes the only link between them
- Have control over communication flows in the network
- May act as intermediaries

Calculating betweenness centrality

- We take into account not just the direct connections of actor i, but the whole network
- We count the number of shortest paths between actors k and j, and isolate those that pass through actor i
- For actor i, betweenness centrality is equal to:

$$C_B(i) = \sum_{jk} \frac{s_{kij}}{s_{kj}}, i \neq j \neq k$$

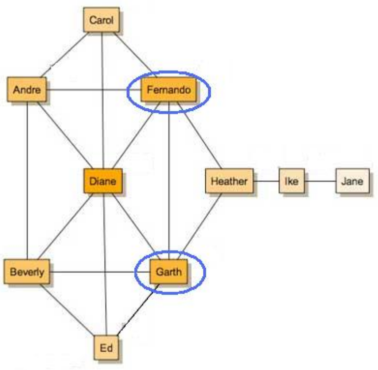
Normalization

- With undirected ties: $C'_B(i) = \dfrac{C_B(i)}{[(n-1)(n-2)]/2}$
- With directed ties: $C'_B(i) = \dfrac{C_B(i)}{[(n-1)(n-2)]}$

where: $(n-1)(n-2)$ is the maximum number of possible pairs of nodes, excluding node i itself; divided by 2 if ties are undirected

Closeness centrality

- Fernando and Garth have fewer connections than Diane
- But they are at a shorter distance from all other network members
- They can directly pass on and receive information flowing through the network
- They have quicker access to communication/information and other resources embedded in the network



Krackhardt's kite network

Actors with high closeness centrality

- A sign of an actor's independence:
  - if you are not central, you need to rely on others to relay messages through the network
  - if you are central, you can relay messages yourself –you are more independent
- May be correlated with organisational influence if the actor is a good communicator

Computing closeness centrality

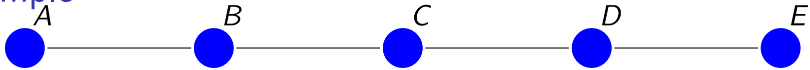
- We take into account the shortest paths linking actors together
- We take centrality as the inverse of *distance* between actors, where shortest distance = highest closeness centrality
- For actor i, it is calculated as:

  $$C_c(i) = \frac{1}{\sum_{j=1}^{n} d_{ij}}$$

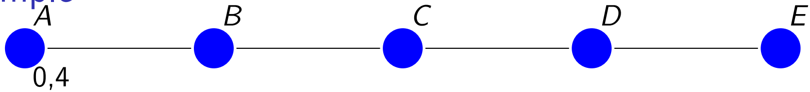
  where d_{ij} is the distance between actors i and j
- ie it is the inverse of *farness*
- Normalized closeness centrality:

  $$C'_c(i) = \frac{n-1}{C_c(i)}$$

Example

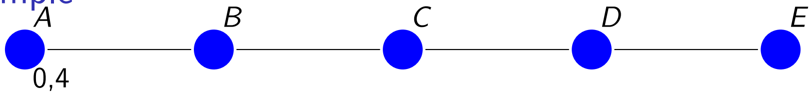


Example



$$C_c(A) = \frac{n-1}{\sum_{j=1}^{n} d_{ij}} = \frac{4}{1+2+3+4} = \frac{4}{10} = 0,4$$

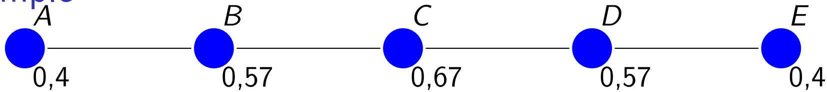
Example



$$C_c(A) = \frac{n-1}{\sum_{j=1}^{n} d_{ij}} = \frac{4}{1+2+3+4} = \frac{4}{10} = 0,4$$

Try to compute closeness centrality for nodes B, C, D and E!

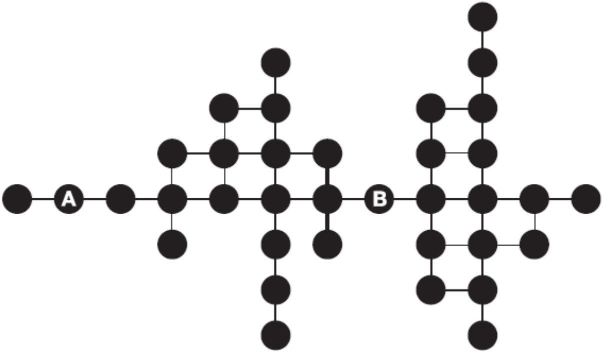
Example



$$C_c(A) = \frac{n-1}{\sum_{j=1}^{n} d_{ij}} = \frac{4}{1+2+3+4} = \frac{4}{10} = 0,4$$

Try to compute closeness centrality for nodes B, C, D and E!

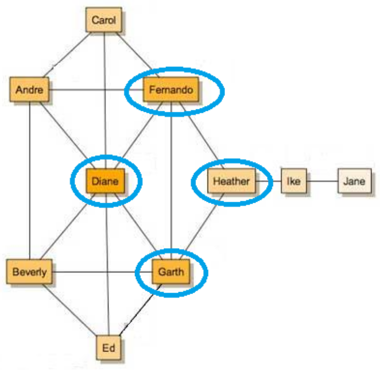
The limitations of degree centrality



A and B have the same degree but intuitively, they are not equally “important". How to weigh degree centrality to account for the fact that B’s contacts are better connected than A’s alters?

Eigenvector centrality

- Diane, Heather, Fernando and Garth are connected to the best-connected actors
- They may have direct influence on the most active, or most influential, people in the network
- They may derive influential power from their network position



Krackhardt's kite network

What eigenvector centrality is needed for

- Like degree centrality, eigenvector centrality gives high scores to nodes with connections with many other nodes
- But it specifically favours nodes that are connected to nodes that are themselves central within the network
- It does so by weighting contacts according to their centralities
- Thus, it takes into account the entire pattern of the network
- In passing, Google's PageRank algorithm is a variant of eigenvector centrality

Calculating eigenvector centrality

- Let A be the nxn adjacency matrix that describes the network
- Let λ be the largest eigenvalue, and x the corresponding eigenvector
- Then $C_E(i) = \dfrac{1}{\lambda} \displaystyle\sum_{j=1}^{n} a_{ij} x_j$
- In vector notation, $Ax = \lambda x$

Reminder: An eigenvector of a square matrix is a non-zero vector that, when multiplied by the matrix, yields a vector that differs from the original at most by a multiplicative scalar. Specifically, a non-zero column vector x is an eigenvector of a matrix A if (and only if) there exists a number λ such that $Ax = \lambda x$. The number λ is called the eigenvalue corresponding to that vector.

Calculating eigenvector centrality, in practice

Eigenvector centrality is computed as a recursive version of degree centrality:

1. Start by assigning centrality score of 1 to all nodes
2. Recompute scores of each node as weighted sum of centralities of all nodes in a node's neighbourhood
3. Normalize by dividing each value by the largest value
4. Repeat steps 2 and 3 until values stop changing

Now you know

- Degree centrality
- Betweenness centrality
- Eigenvector centrality
- Closeness centrality

References I

Akerlof, G. (1982). Labor contracts as partial gift exchange. *Quarterly Journal of Economics*, 97, 543–569.

Alba, R. & Kadushin, C. (1976). The intersection of social circles. A new measure of social proximity in networks. *Sociological Methods and Research*, 5(1), 77–102.

Backstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2011). Four degrees of separation. ArXiv:1111.4570, 19 Nov 2011.

Bearman, P., Moody, J., & Stovel, K. (2004). Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology*, 110(1), 44–91.

References II

Blake, E. (2014). Dyads and triads in community detection: a view from the Italian bronze agey. *Les nouvelles de l'archéologie*, (135), 28–32.

Bott, E. (1957). *Family and social networks (2e edition)*. The Free Press.

Burt, R. (1992). *Structural Holes : The Social Structure of Competition*. Harvard University Press.

Fehr, E. & Gächter, S. (2000). Fairness and retaliation: the economics of reciprocity. *Journal of Economic Perspectives*, 14(3), 159–181.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.

References III

Jackson, M. (2008). *Social and Economic Networks*. Princeton University Press.

Lazega, E., Mounier, L., Snijders, T., & Tubaro, P. (2012). Norms, status and the dynamics of advice networks: A case study. *Social Networks*, 34(3), 323–332.

Milgram, S. (1967). The small world problem. *Psychology Today*, 2, 60–67.

Phillips, D. (2011). Jazz and the disconnected: City structural disconnectedness and the emergence of a jazz. *American Journal of Sociology*, 117(2), 420–483.

Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The anatomy of the Facebook social graph. ArXiv:1111.4503, 23 Nov 2011.

Thank you!

Paola Tubaro, paola.tubaro@ensae.fr