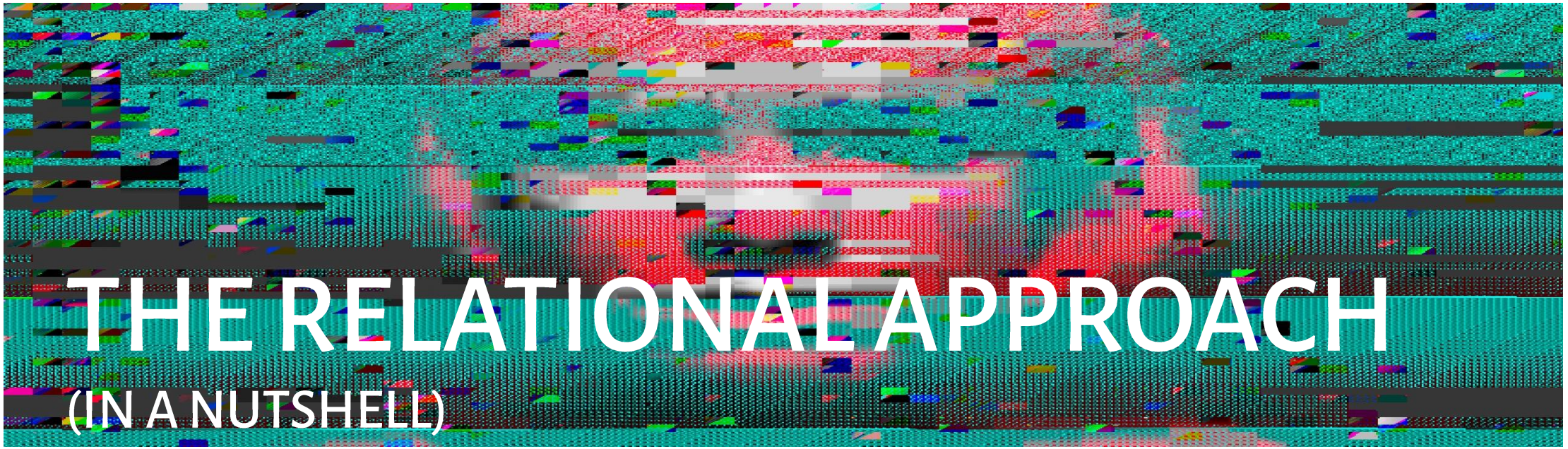
Introduction to Network Analysis & Online Network Analysis

DMI Winter School 2016 Amsterdam

Elena Pavan, ISUS SNS – elena.pavan@sns.it

CC BY NC ND


THE RELATIONAL APPROACH
(IN A NUTSHELL)

Who Said… ?

A collection of human beings does not become a society because each of them has an objectively determined of subjectively impelling life-content. It becomes a society only when the vitality of these contents attains to the form of reciprocal influence; only when an individual has an effect, immediate or mediate, upon another, is mere spatial aggregation or temporal succession transformed into society.

Georg Simmel (1908[1971]:23)

Who Said… ?

A collection of human beings does not become a society because each of them has an objectively determined of subjectively impelling life-content. It becomes a society only when the vitality of these contents attains to the form of reciprocal influence; only when an individual has an effect, immediate or mediate, upon another, is mere spatial aggregation or temporal succession transformed into society.

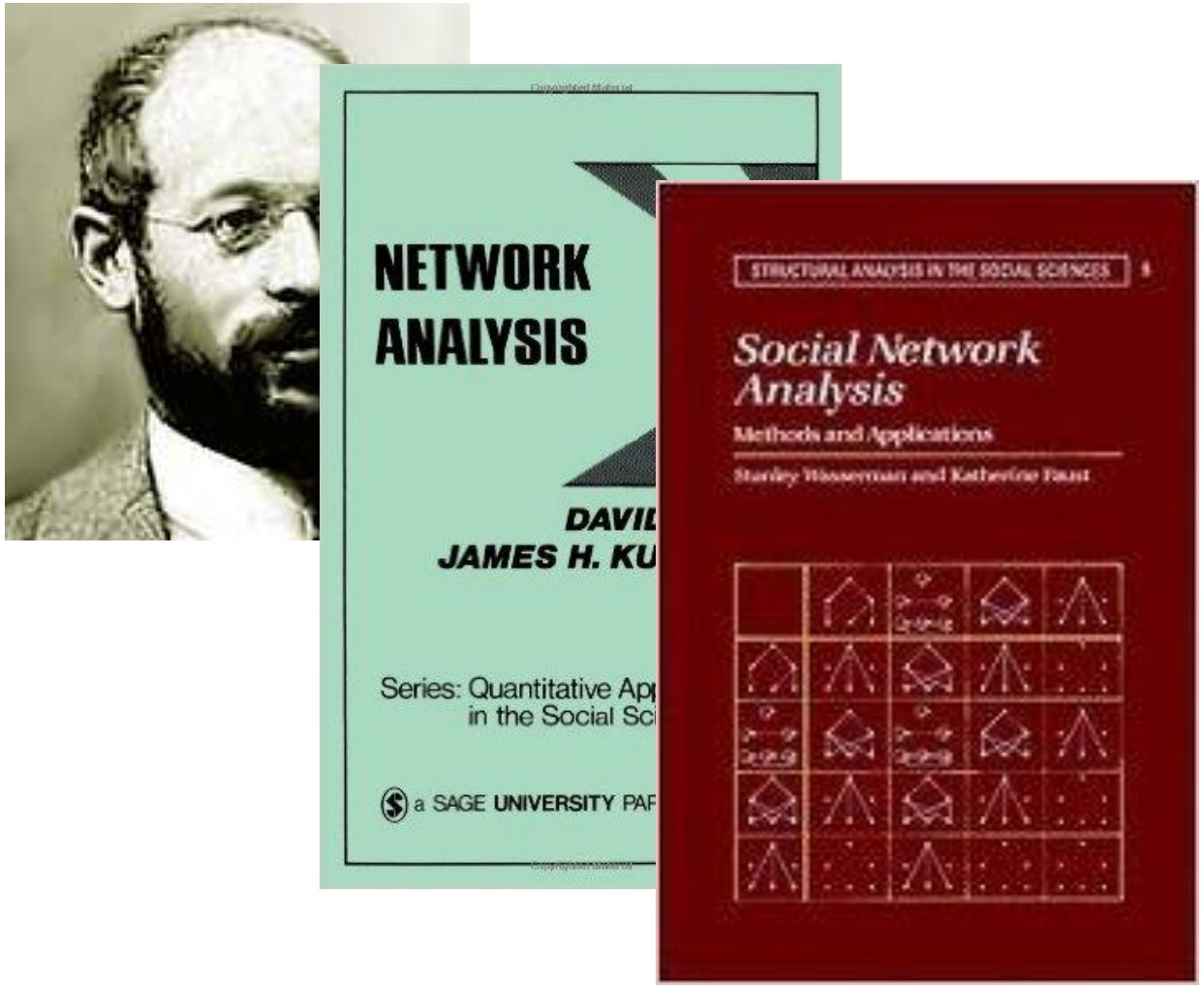
Georg Simmel (1908[1971]:23)

Simmelian Roots

1. Social ties are primary
2. There is no society without interactions
3. Society is not an aggregate of single individuals
4. Sociologists must engage in the study of relational patterns (i.e., the forms) rather than on that of individual motivations, beliefs, emotions (i.e., contents)
5. "Social geometry":
   - Isolated individual
   - Dyad
   - Triad

Networks Ideas and Definitions

- Simmel (1908): Society as an ensemble of relations amongst individuals
- Knoke and Kuklinski (1982): a specific kind of relation linking a defined set of persons, objects or events
- Wasserman and Faust (1994): a set of nodes (or network members) that are tied by one or more types of relations



Network: a concept and a method

SNA is neither a theory nor a methodology. Rather, it is a perspective or a paradigm (Marin and Wellman 2011).

- Concept: social context and processes deploying within it can be seen as the bulk of patterns and regularities characterizing the structure of relationships that are established amongst interacting units
- Method: assessment and evaluation of models and theories.
  - Formalist approach: describing the mathematical form of social network, their causes, their effects.

  Ex: networks with dense clusters with few connection amongst clusters imply the existence of short paths amongst the majority of nodes
  - Structuralist approach: studying how patterns of relation can shed light on substantive matters within a discipline.

  Ex: networks with dense clusters with few connection amongst clusters in a multistakeholder negotiation entails a limited interplay between different constituencies

Network elements

<table>
<tr><th>Nodes</th><th>Ties</th></tr>
<tr><td rowspan="2">Type: people, organizations, groups, institutions, technological artifacts, Web resources, events...</td><td>Content: friendship, resource exchange, collaboration, conflict, following, being followed, being a fan, posting, tagging ...</td></tr>
<tr><td>Frequency: sustained over time, occasional, frequent...</td></tr>
<tr><td>Number: one, two, three... N</td><td rowspan="2">Direction: symmetrical or asymmetrical</td></tr>
<tr><td>Attributes: qualities of nodes (categorical, continuous...)</td></tr>
<tr><td></td><td>Intensity: strong vs. weak</td></tr>
</table>

Network Boundaries

Laumann, Knoke and Prensky (1983) outlined 2 approaches:

- Realist: boundaries are set up directly by actors depending on their perceptions
- Nominalist: boundaries are set up by the researcher following her conceptualization of the object of study
  - Position-based: only actors in a certain position within a network are included
  - Event-based: only actors participating to a certain event
  - Relation-based: begins with a small group of nodes and includes only others holding with these a particular relationship

Network Words

- Actor: social entity whose relations we trace
- Relational tie: what connects actors
- Dyad: a group of two nodes and ties amongst them (→ reciprocity; likelihood of ties co-occurrence)
- Triad: a subset of three actors and possible ties amongst them (→ transitivity; balance)
- Subgroup: any subset of actors and ties amongst them
- Group: collection of all actors on which ties are to be measured (→ conceptual, theoretical or empirical reasons should ground the selection of actors)
- Relation: collection of ties of a specific kind amongst members of a group (e.g., the set of friendships amongst classmates)
- Social Network: a finite set or sets of actors and the relation or relations defined on them

(Wasserman and Faust 1994)

Guiding principles

- Relations, not attributes: causation is not located *in* the individual but *in the social structure* → while people with same attributes often behave in a certain way they also often occupy the same position within social structure, i.e., same opportunities and constraints to action
- Networks, not groups: network embeddedness is not binary, there are different levels of group membership, multiple memberships, cross-cutting ties between groups
- Relations in a relational context: not only relational ties but study of relational patterns, i.e., opportunities and constraints depend from a node's position but depends also from other nodes' position

(Marin and Wellman 2011)

Goals of analysis

- Applied/Descriptive NA: calculation of network metrics to describe the structure of the network or capture aspects of individuals' position in the network
- Basic/Explanatory NA**:** describe the variance of a variable as a function of others. Network variables can treated as:
  - Independent variables: we consider nodes' position as affecting their behavior (we elaborate a "network theory of ….")
  - Dependent variables: we consider presence/absence of ties as resulting from attributes (we elaborate a "…. theory of networks")

(Borgatti et. al 2013)

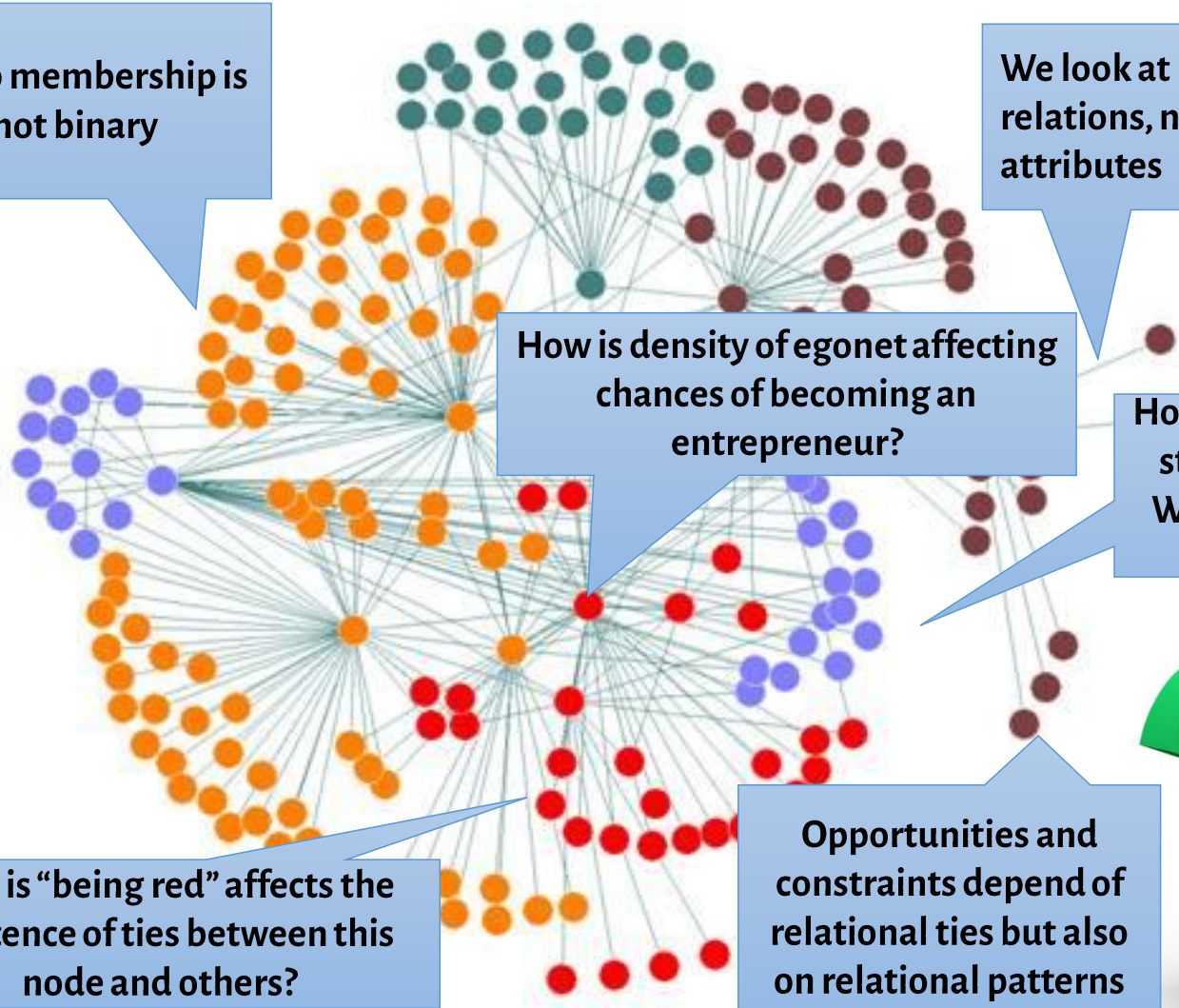
(Keep calm and) Think networks

Group membership is not binary

We look at relations, not at attributes

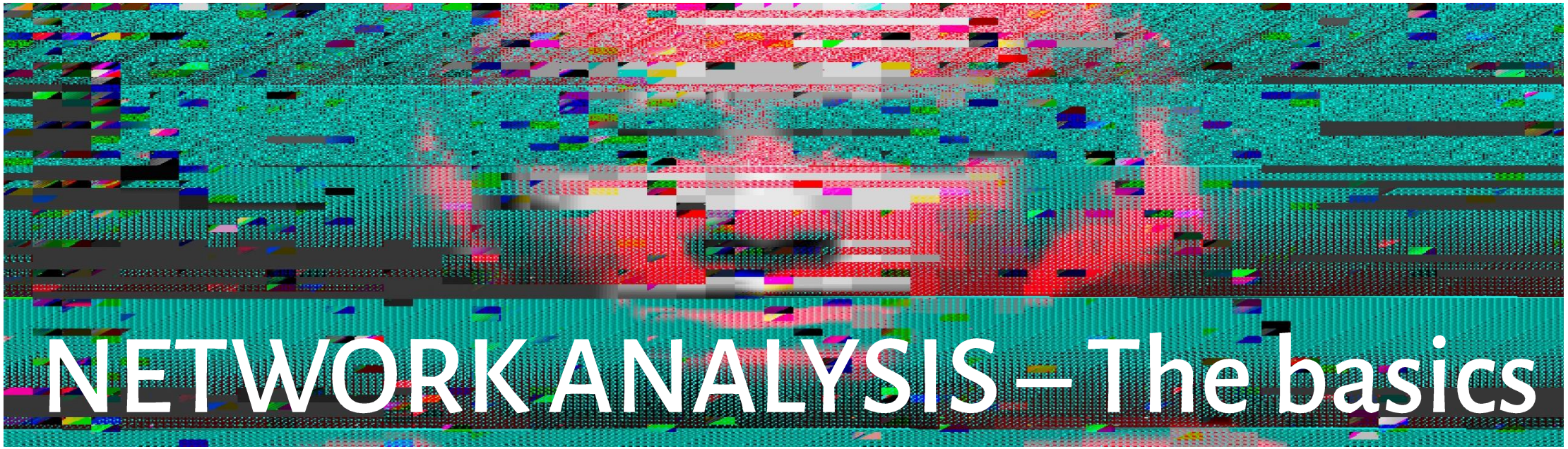
How is density of egonet affecting chances of becoming an entrepreneur?

How is this relational structure shaped? What does it mean for...



How is “being red” affects the existence of ties between this node and others?

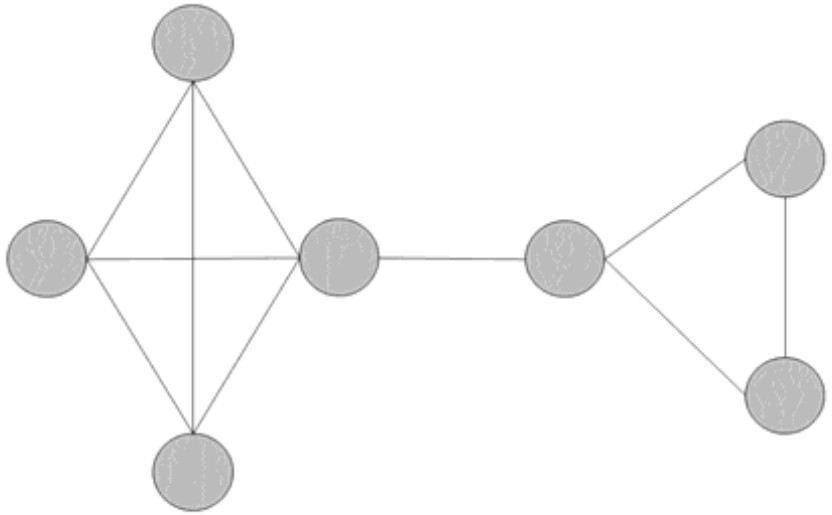
Opportunities and constraints depend of relational ties but also on relational patterns


NETWORK ANALYSIS – The basics

Representing Networks

Graphs

Mathematical object formed by nodes and ties whose properties can be studied from a formal point of view through “graph theory”



Matrices

Mathematical object representing sources and destination of ties as well as the presence/absence of ties.

- Ways : dimensions of a matrix (usually 2)
- Modes: kind of entities represented

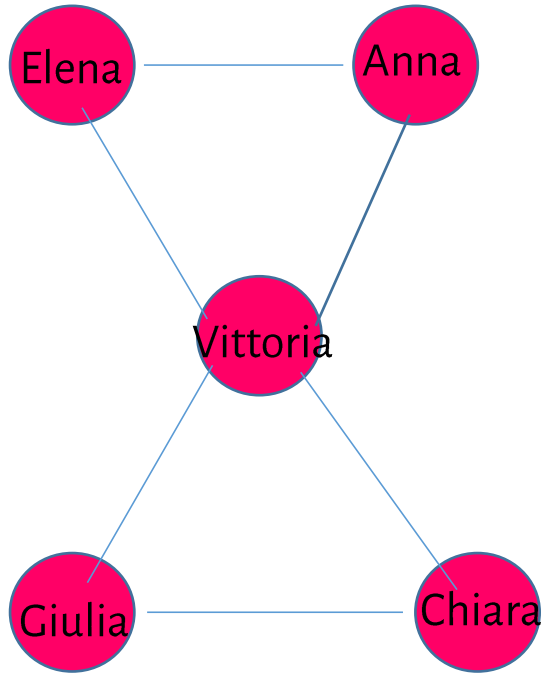
	A	B	C	D	E	F	G
A	-	1	0	1	0	0	0
B	1	-	0	1	0	0	0
C	0	0	-	0	0	0	0
D	1	1	0	-	0	1	0
E	0	0	0	0	-	0	1
F	0	0	0	1	0	-	0
G	0	0	0	0	1	0	-

One way of coding data: Edgelist

- Sequence of lines: every line represent a tie where the first name is the source of the tie, the second the receiver

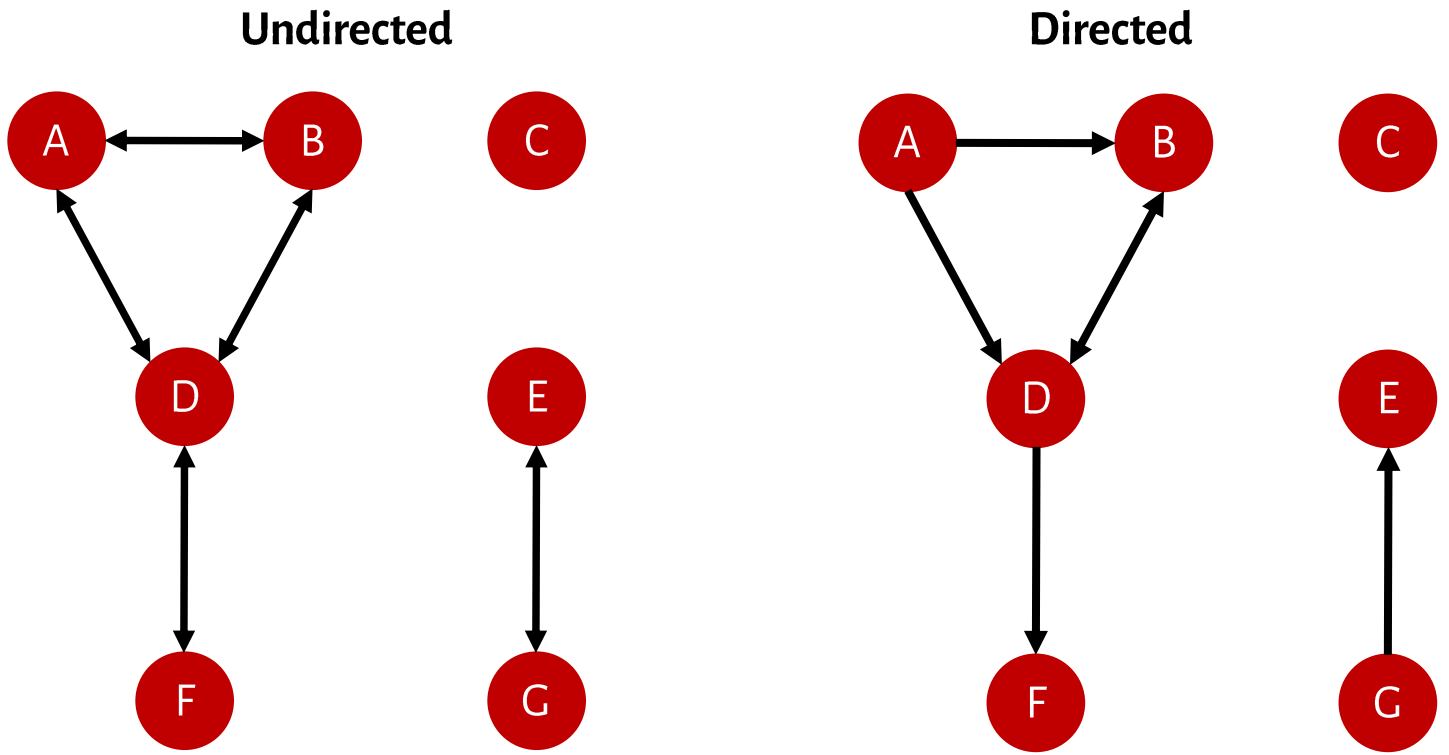
Node 1	Node 2
Elena	Anna
Elena	Vittoria
Anna	Vittoria
Vittoria	Giulia
Giulia	Chiara
Chiara	Vittoria

	E	A	V	G	C
E	-	1	1	0	0
A		-	1	0	0
V	1	1	-	1	1
G	0	0	1	-	1
C	0	0	1	1	-



Direction

- Networks can be directed or undirected (symmetrical or asymmetrical) – matrices must be (coded and) read from rows to columns



Symmetrical adjacency matrix

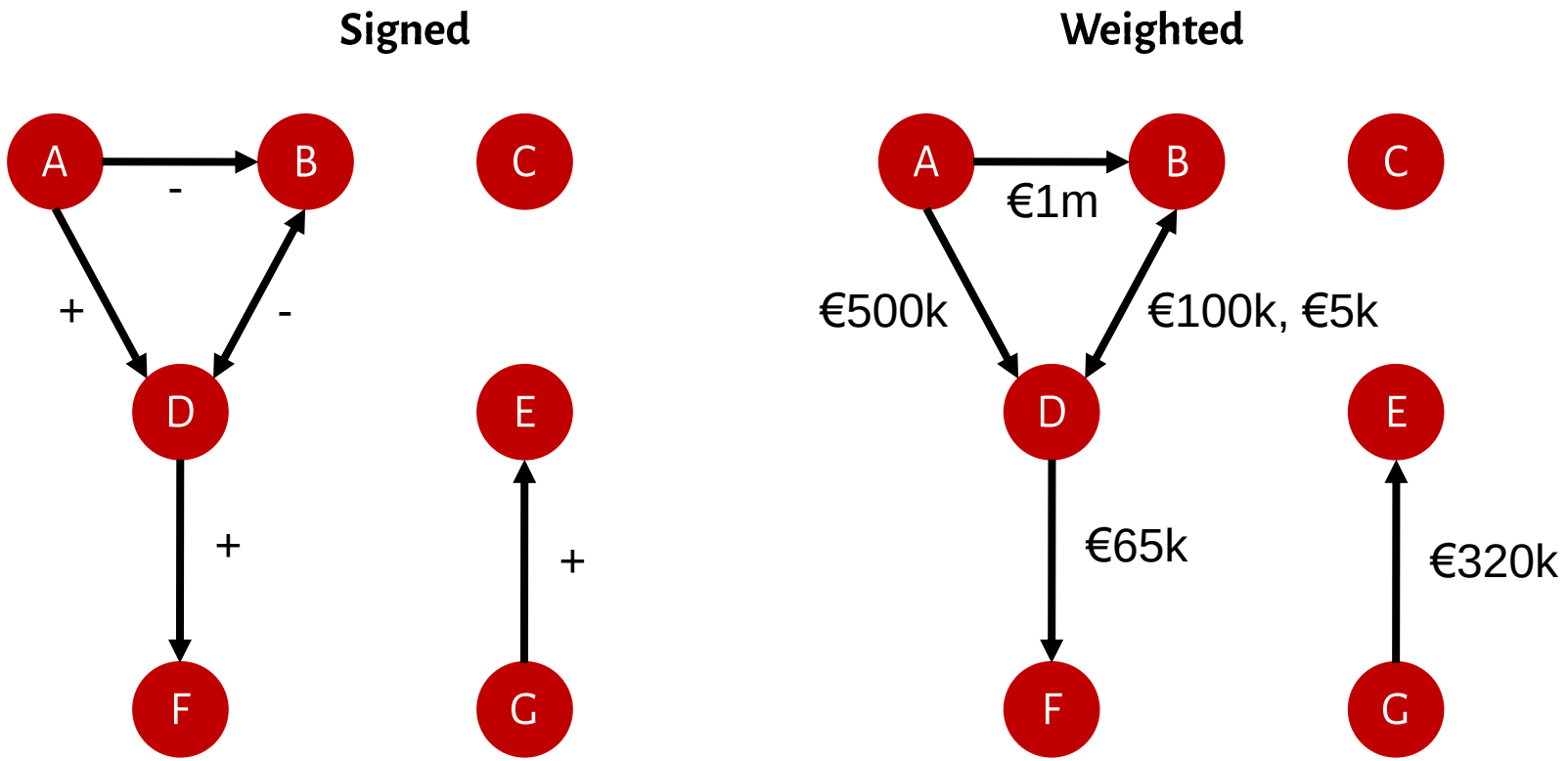
	A	B	C	D	E	F	G
A	-	1	0	1	0	0	0
B	1	-	0	1	0	0	0
C	0	0	-	0	0	0	0
D	1	1	0	-	0	1	0
E	0	0	0	0	-	0	1
F	0	0	0	1	0	-	0
G	0	0	0	0	1	0	-

Asymmetrical adjacency matrix

	A	B	C	D	E	F	G
A	-	1	0	1	0	0	0
B	0	-	0	1	0	0	0
C	0	0	-	0	0	0	0
D	0	1	0	-	0	1	0
E	0	0	0	0	-	0	0
F	0	0	0	0	0	-	0
G	0	0	0	0	1	0	-

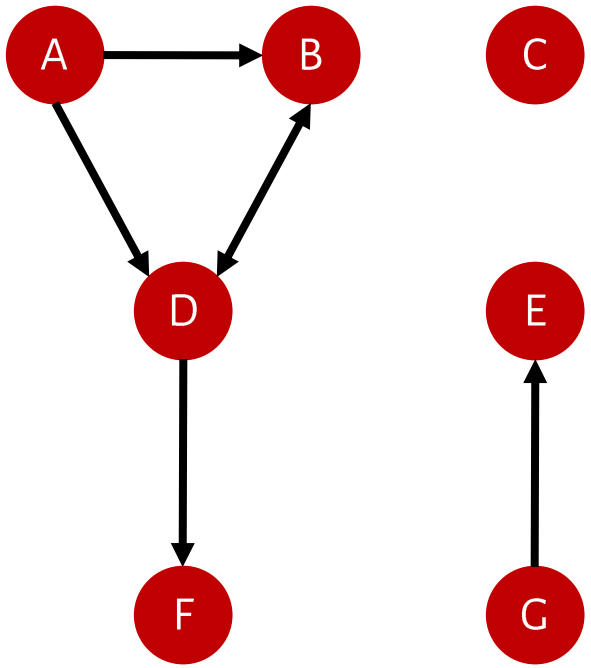
Weights

- Networks can be signed or weighted, i.e., matrices can have values ≠ from 0 and 1



Unidimensional networks

- Unimode, uniplex – single type of nodes and single relationship



Perhaps the **most used** network type so far

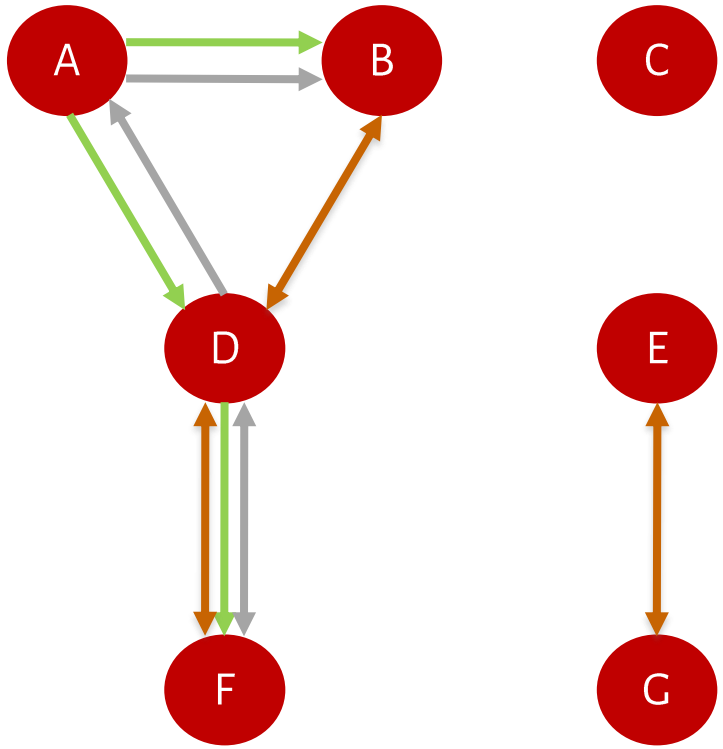
Adjacency matrix is **squared** (NxN)

Nodes can be social actors but also any kind of other entities (as long as they are the same in rows and columns) – see Faust and Skvoretz comparing 42 networks of social species (human, non-human primates etc.)

Relations: any content bonding nodes, can be signed, weighted, directed

Unimodal multiplex networks

- Unimodal, multiplex – one set of nodes, multiple relations amongst them



EXERTING AUTHORITY ON

TRUST

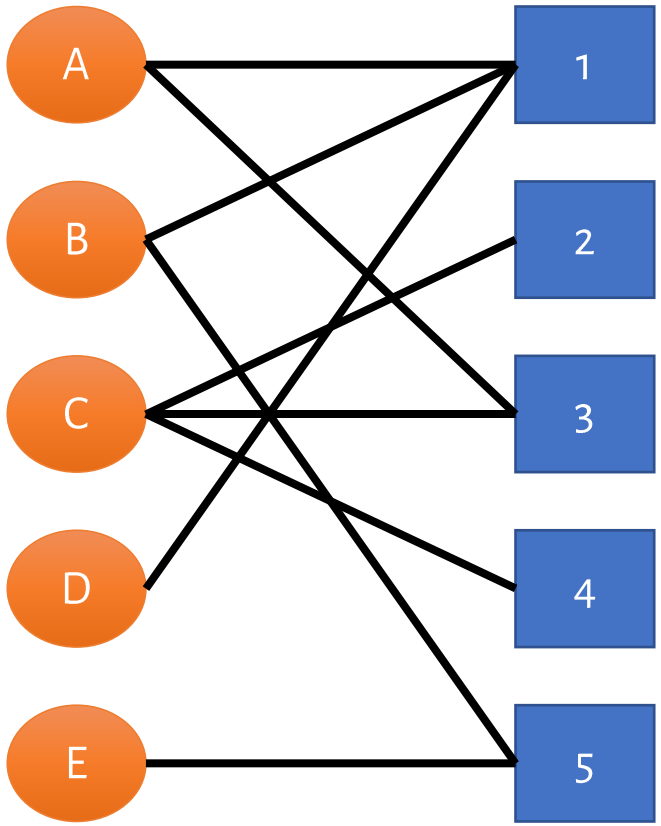
FRIENDSHIP

Set of **squared matrices**

Examination of different structures and **influence** of a relation over the other

Multimodal uniplex networks

- Multimodal, uniplex – more sets of nodes (usually 2) and one type of relation



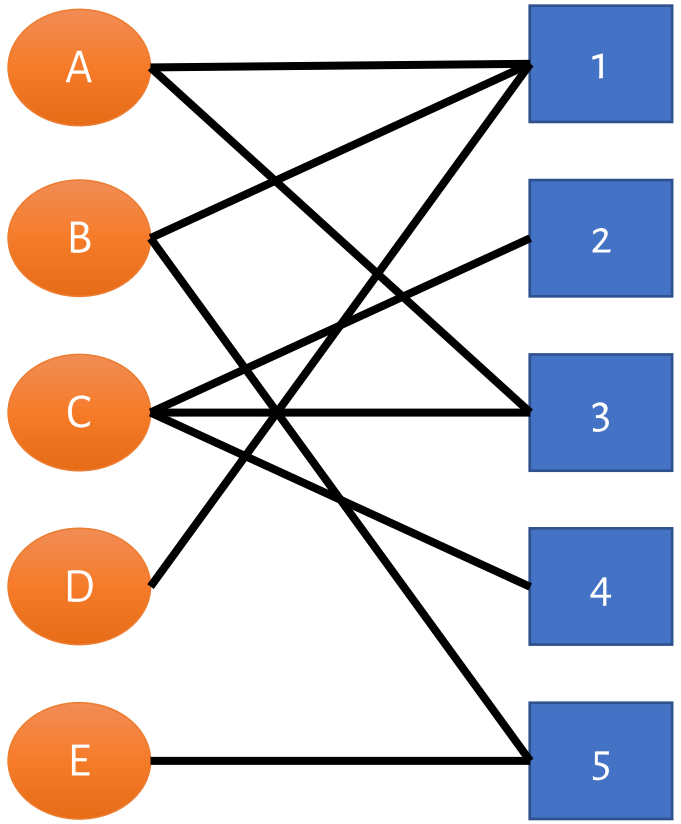
Relationship are established between **nodes in the first set** and **nodes in the second**

Matrices are **rectangular** (NxM) – the number of rows and columns can differ

Two-mode networks are the most popular version of multimode uniplex

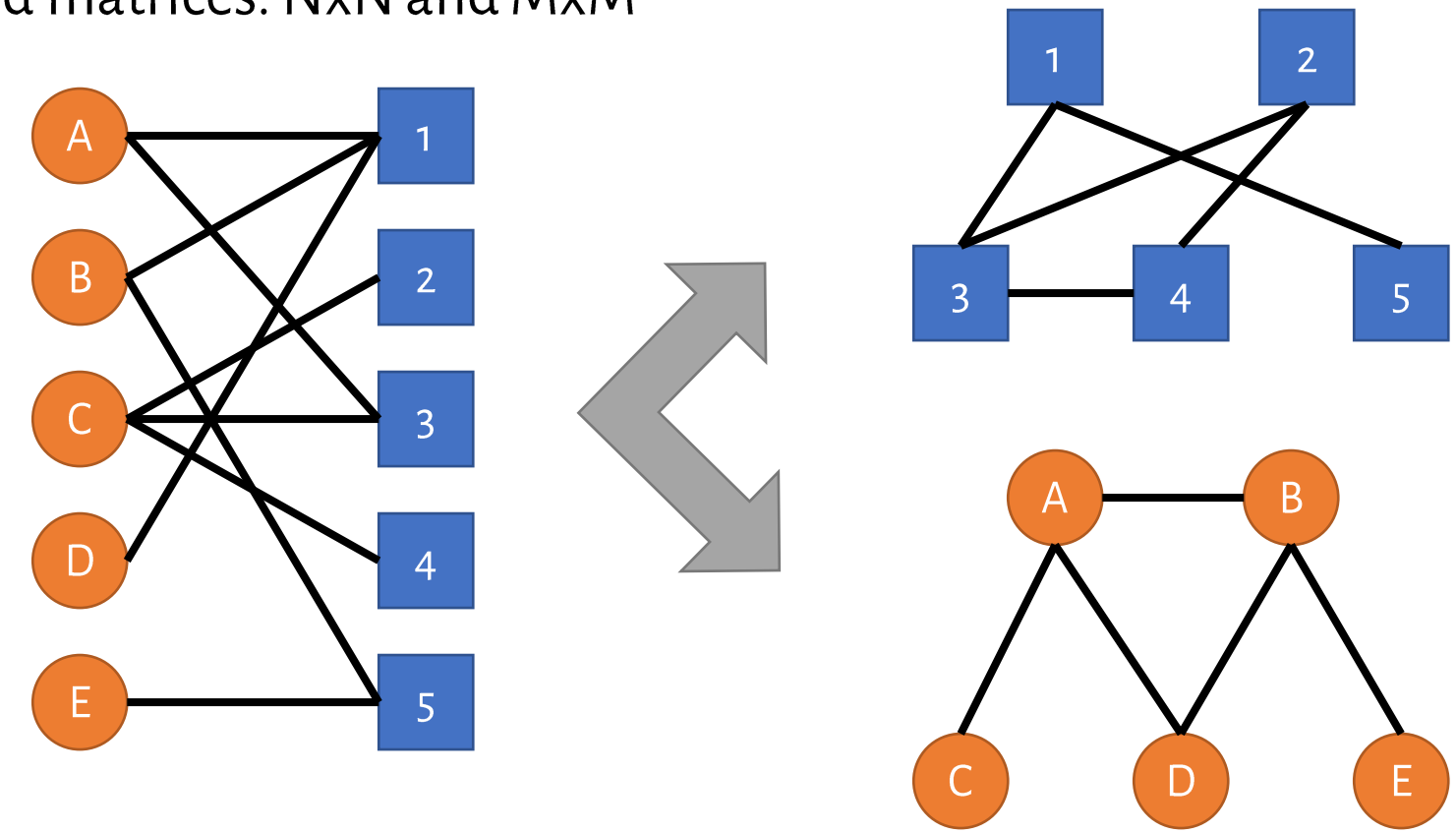
Two-mode networks

	1	2	3	4	5
A	1	0	1	0	0
B	1	0	0	0	1
C	0	1	1	1	0
D	1	0	0	0	0
E	0	0	0	0	1



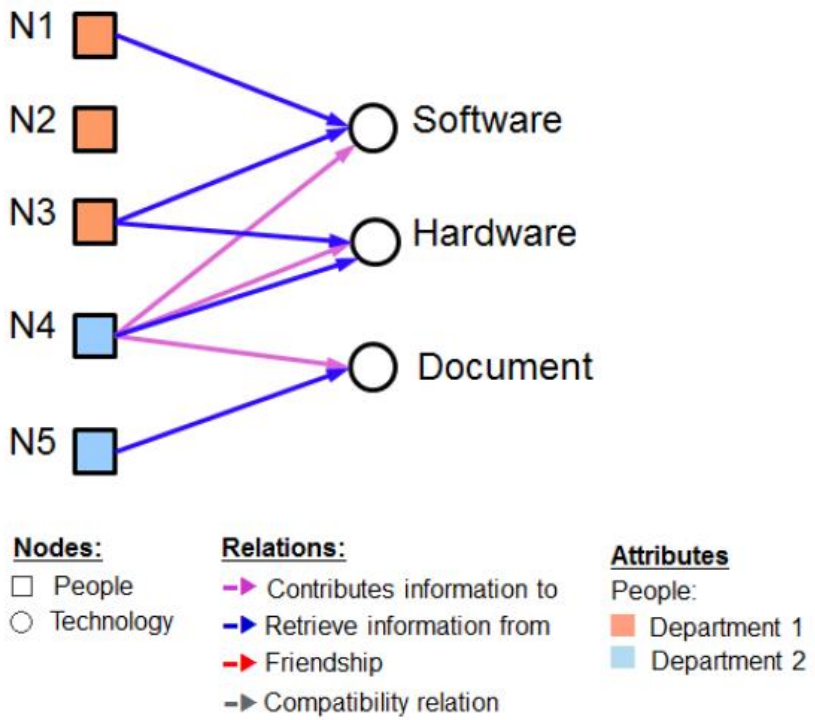
Combining perspectives

- Breiger (1972): the duality of persons and groups – from a NxM matrix we can derive two weighted matrices: NxN and MxM



Multidimensional networks

- Multimode, multiplex – different sets of nodes and different relations amongst them



Render an increased level of complexity

Usually disaggregated

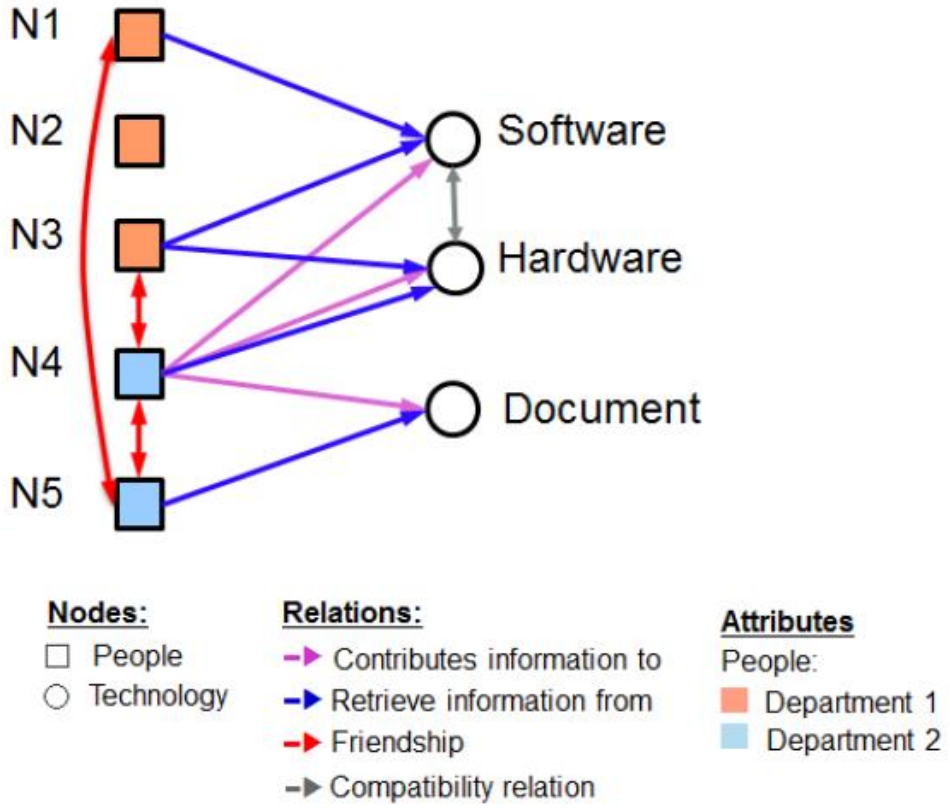
Analytical techniques in progress

Scholarly research in this field is rare – exception: Powell et al. (2011) on the evolution of biotechnology industry (heximodal and quadruplex)

Source: Contractor et al. 2011

Fully multidimensional networks

- Multimode, multiplex and ties established amongst and within groups



Higher level of complexity overcoming limitations derived from the study of two-mode networks

Preferably **longitudinal research design t**o see if presence/absence of ties at Time 1 influences ties configuration in Time 2

Source: Contractor et al. 2011

Summary

Network type	Modes	Relations	Place of relations	Example
Unidimensional	One	One	Within group	Friendship on Facebook
Unimodal Multiplex	One	More than one	Within group	Conversations on Twitter
Multimodal Uniplex	More than one	One	Between groups	Page editing in Wikipedia
Multimode Multiplex	More than one	More than one	Between groups	Comment/Like on Facebook
Fully Multidimensional	More than one	More than one	Within and between groups	Reply/Comment/Like Facebook

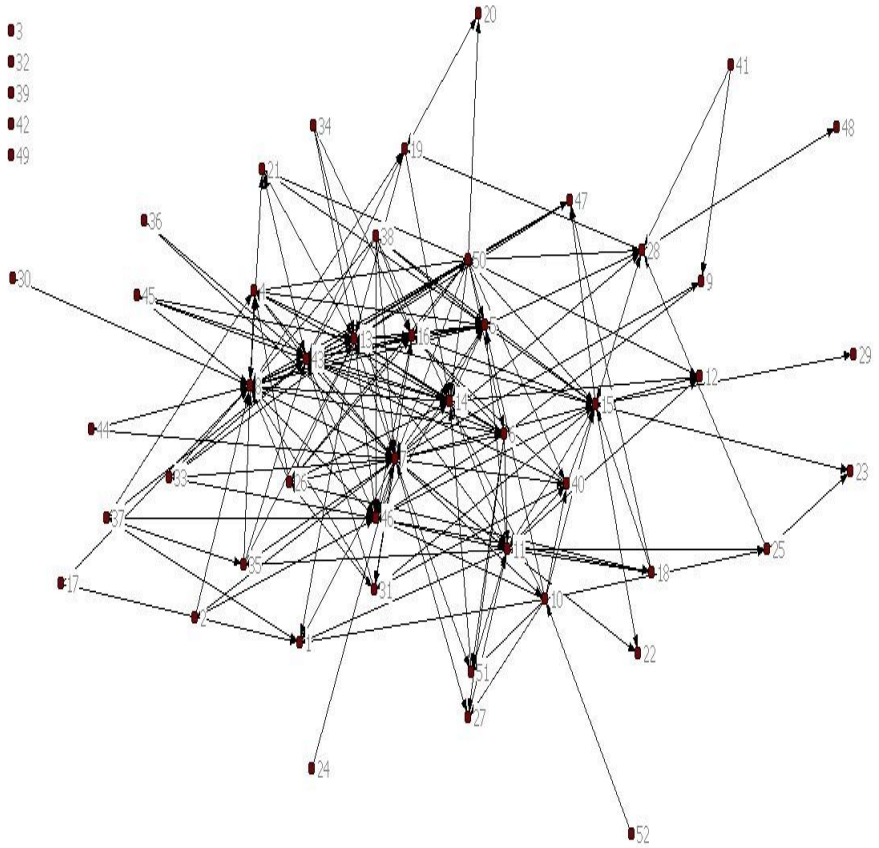
Cohesion

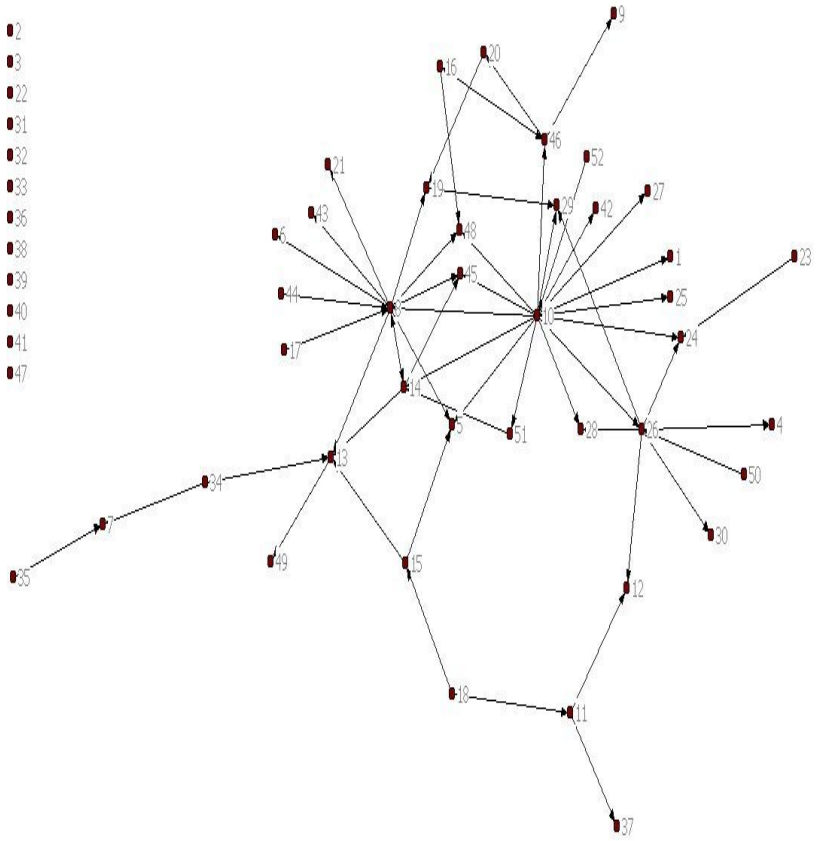
- How much a system of relations is connected
- Not necessarily "positive" – think of a network of "who hates whom": greater cohesion means less social cohesion
- Simplest measure: density = the proportion of ties activated on the possible number of ties
- Varies from 0 to 1
- In an ordinary, undirected, non-reflexive and unweighted network the number of ties= $n(n-1)/2$ and density= $2L/n(n-1)$
- In an ordinary, directed, non-reflexive and unweighted network the number of ties = $n(n-1)$ and density= $L/n(n-1)$
- In a weighted network density expresses the average strenght of ties in the relation system

Interpreting density

- Metric depending on the number of nodes in a network
- Content of ties must guide interpretation – in a close group (e.g., a class of students) where you map "who is friend with whom" a density of 0.264 is low; in a Twitter network a density of 0.004 is farily regular
- Often used in comparative terms to explore variations over time or differences amongst relations established in the same actor set
- Sometimes substituted by average degree (average number of ties of nodes in a network, easier to interpret)

Ex. Migrant organizations in Lyon

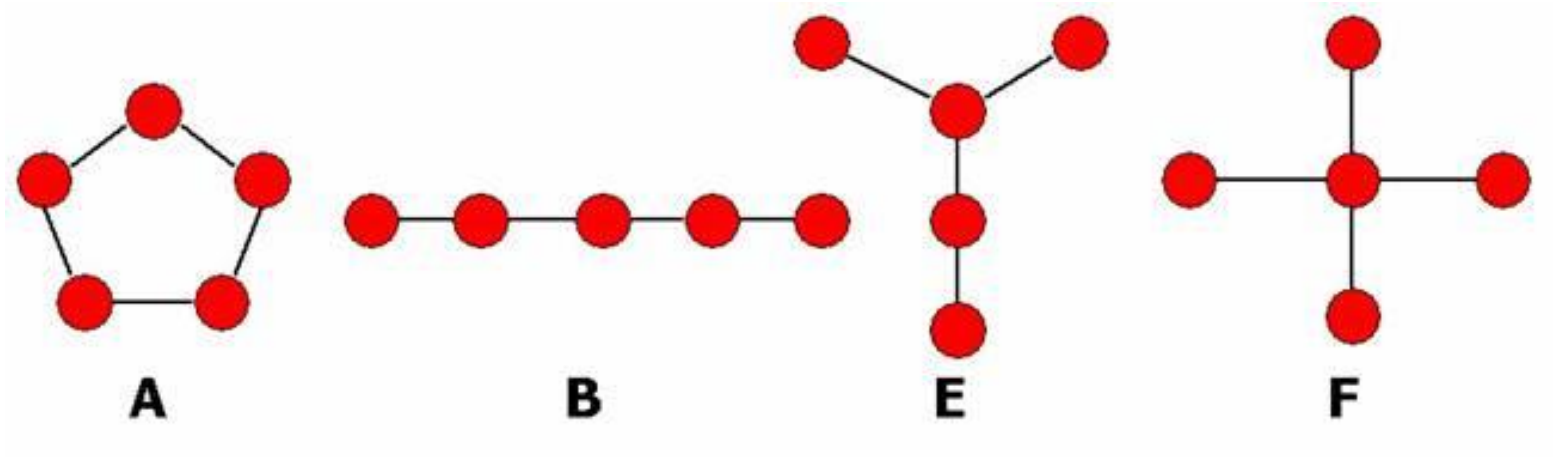




Source: Eggert (2011, 2014)

Centralization

The extent to which a network is “dominated” by a single node. The higher the centralization, the fewer actors stand at the core of a network



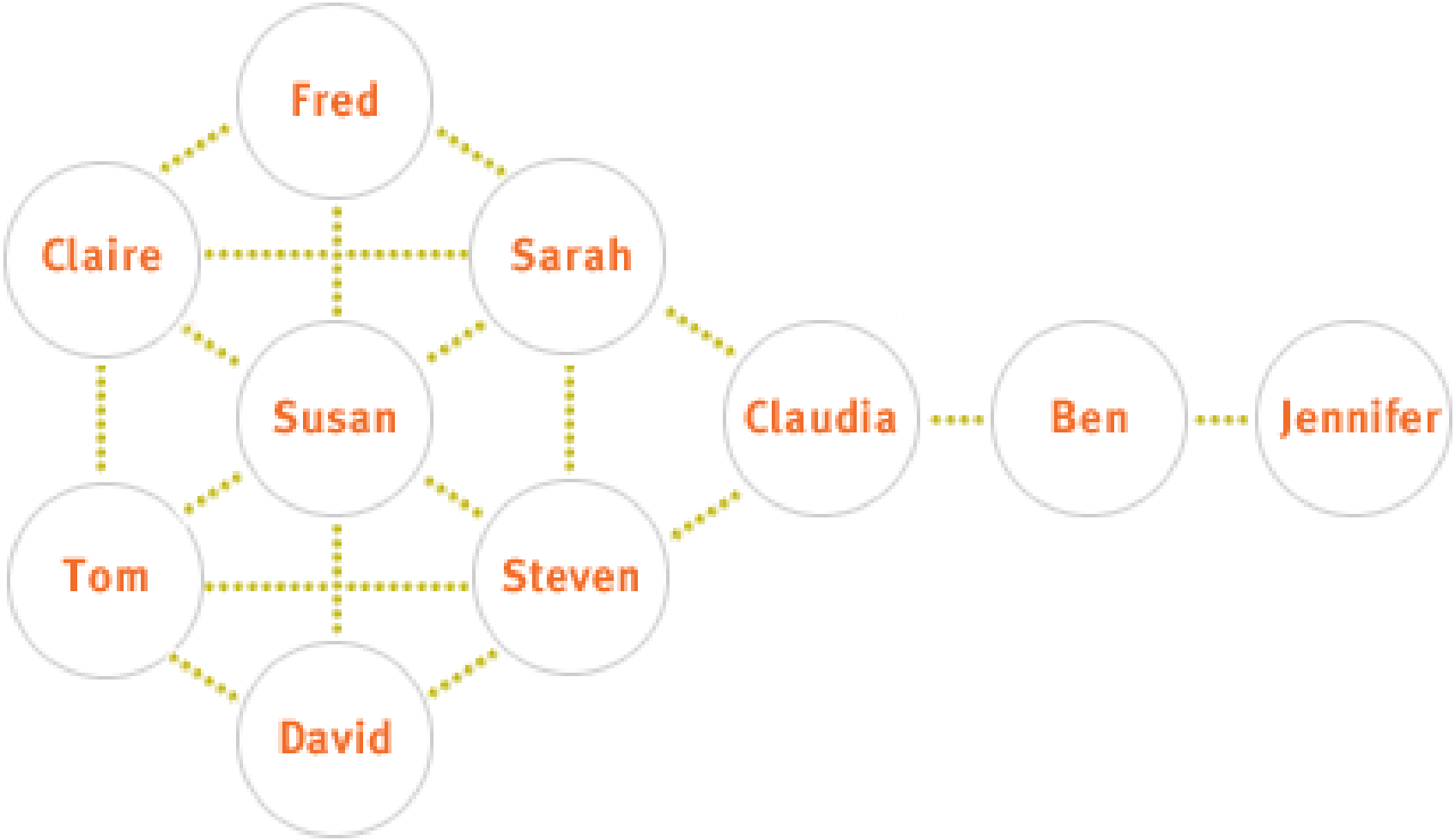
Clustering coefficient

- Proposed by Watts and Strogatz (1998)
- Identifies the extent to which a network has areas of high and low density
- Calculates the density of each egonet (individual clustering coefficient) and average this quantity
- Used to find “small-world” structures (human networks are, at the same time, clumpy but there is a rather short path amongst any two nodes)
- Becomes particularly useful in large scale networks

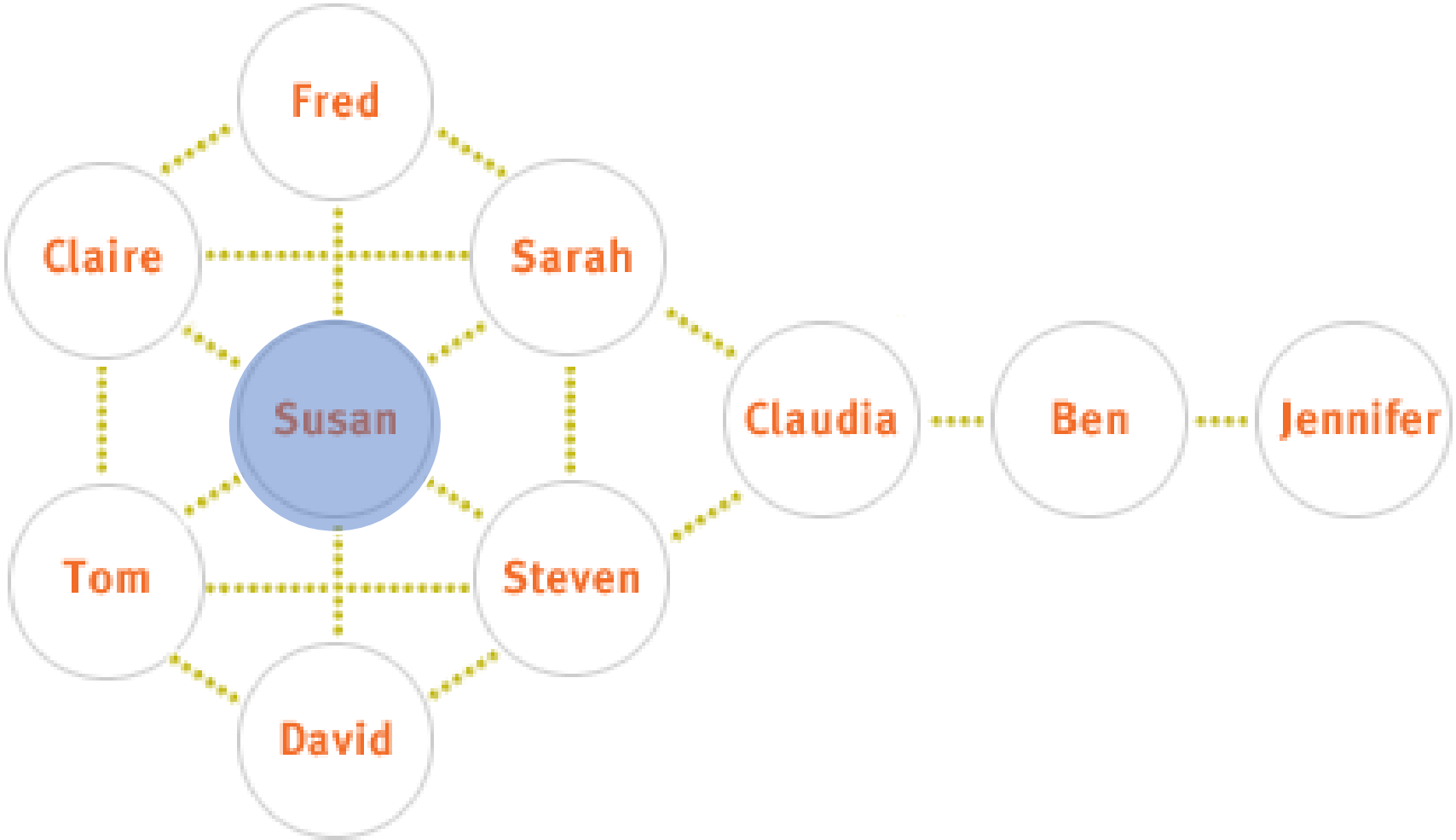
Centrality

- It is a **property of a node's position** in the network; stands for the "contribution" a node gives to the network it belongs to but also for the "advantages" she may derive from being in a certain position
- It's not connected to who a node is (her attributes) but, more likely, where she stands
- Connects to **prominence**, i.e., the extent to which a node is "extensively involved in relationships with others" (Wasserman and Faust 1994:173)
- Knoke and Burt (1983) distinguished prominence into **prestige** (directed relations) and **centrality** (undirected relations)
- It's a **family of measurements**, whose calculation and interpretation vary according to the type of network and, more broadly, according to the relational context/theoretical argument adopted (power, prestige, authority…)

Who is central?

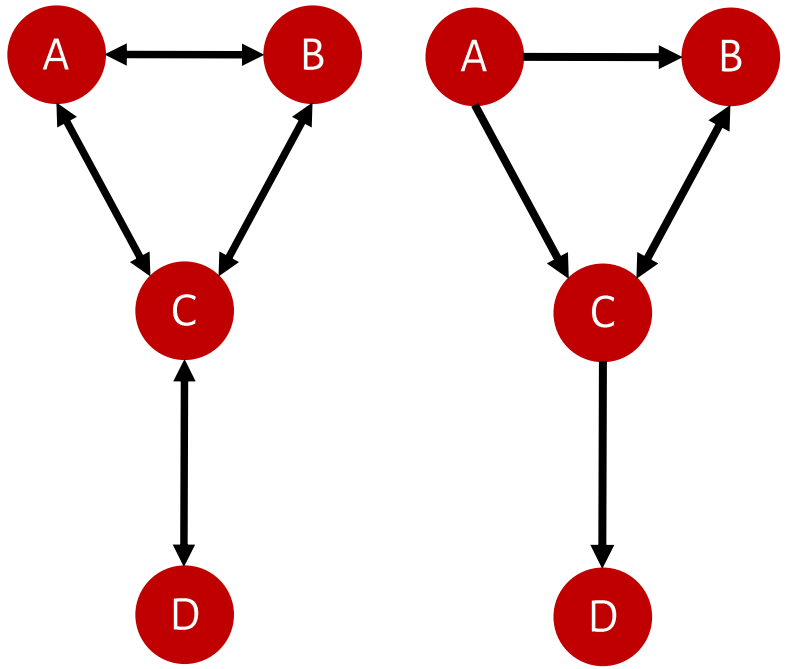


Degree


Fred
Claire
Sarah
Susan
Claudia
Ben
Jennifer
Tom
Steven
David

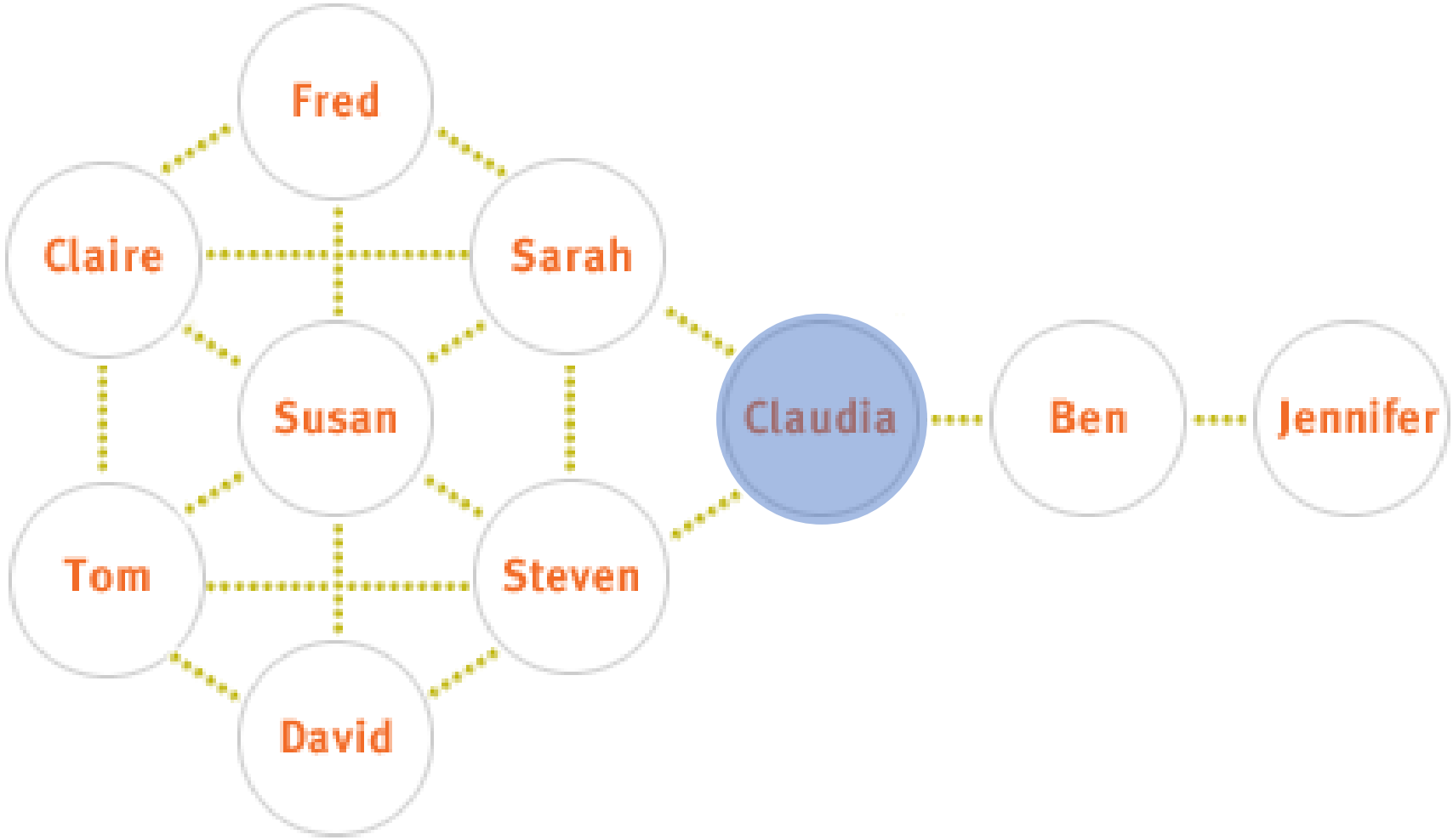
Degree

- Number of ties of a given type a node has, i.e., popularity
- In directed networks we distinguish between:
  - Indegree: number of ties received by a node (column marginal)
  - Outdegree: number of ties sent by a node (row marginal)



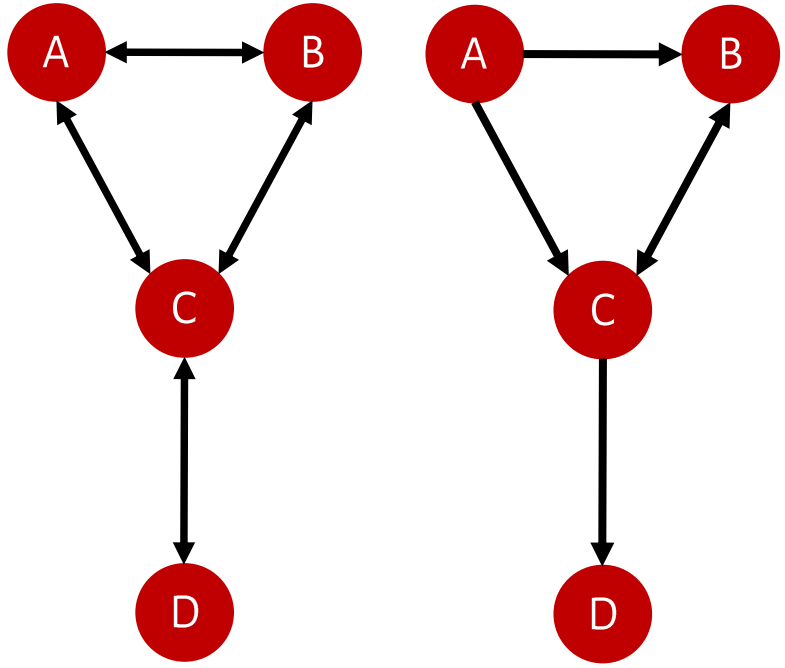
Node	D	I_d	O_d
A	2	0	2
B	2	2	1
C	3	2	2
D	1	1	0

Betweenness



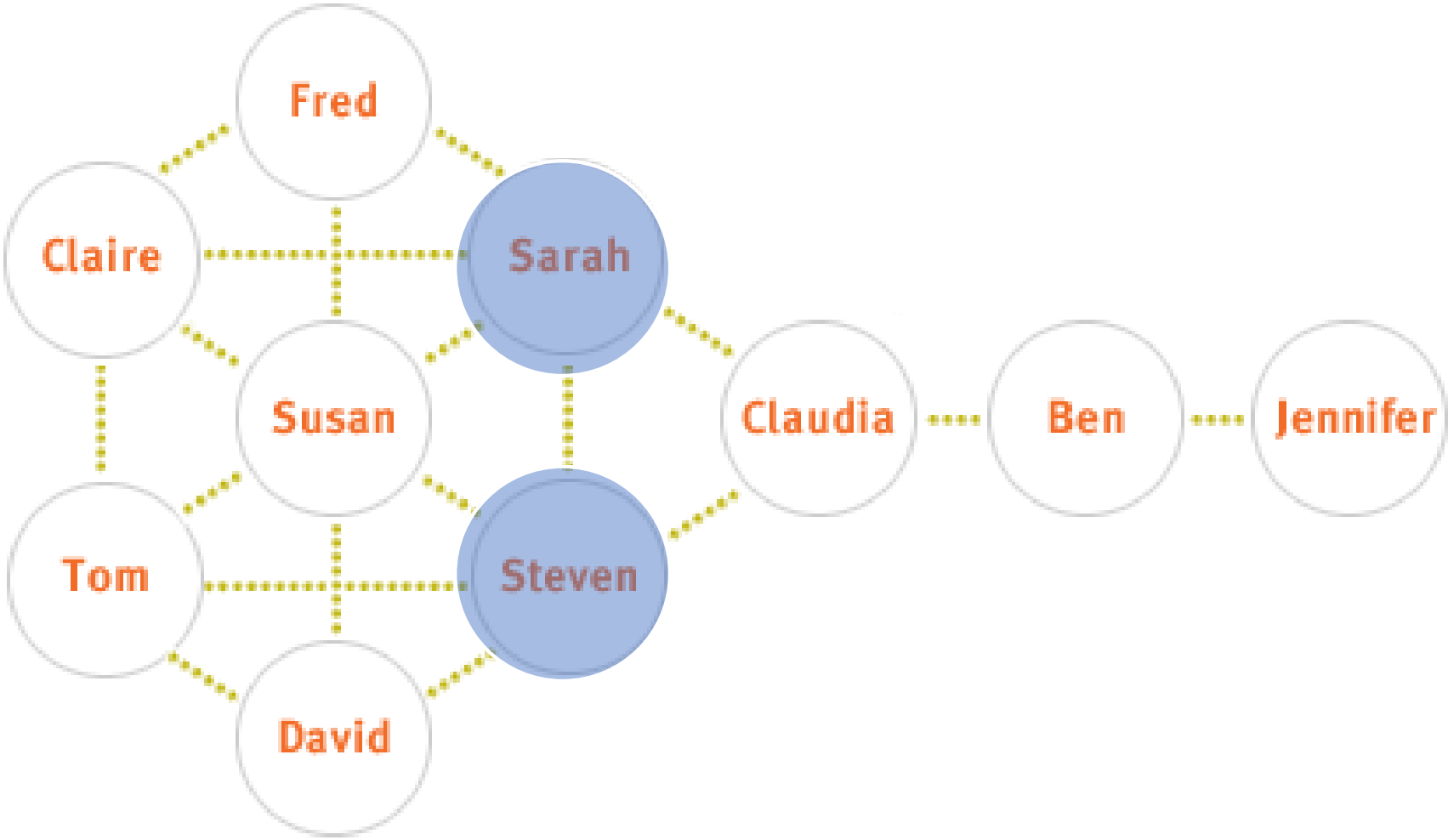
Betweenness

- How often a given node falls along the shortest path between two nodes, i.e., capacity to control, mediate the flows in the network



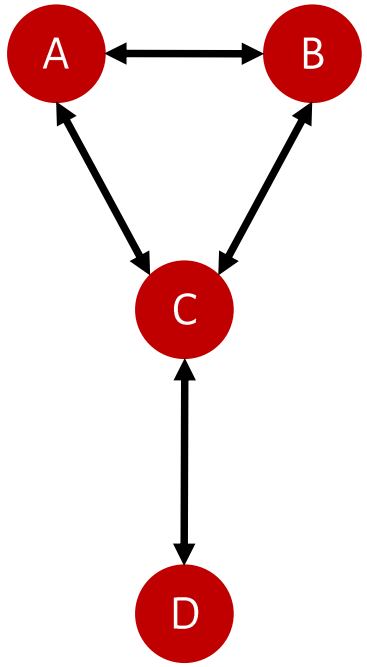
Node	B	B_D
A	0	0
B	0	0
C	2	2
D	0	0

Closeness


Fred
Claire
Sarah
Susan
Claudia
Ben
Jennifer
Tom
Steven
David

Closeness

- Sum of geodesic distances from a nodes to all others
- Inversed measure of centrality (larger number = more peripheral)
- Not suited for directed data and problematic for disconnected graphs (directed graphs tend to be disconnected)

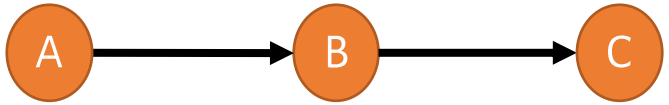


Calculate farness (how distant a node is from all others) and make the reciprocal of that measure

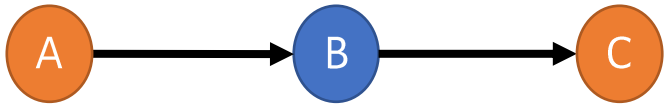
Node	F	NC
A	4	75
B	4	75
C	3	100
D	5	60

Brokerage

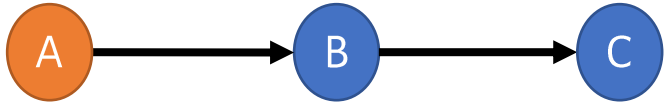
1. **COORDINATOR**
   - Mediation within a group



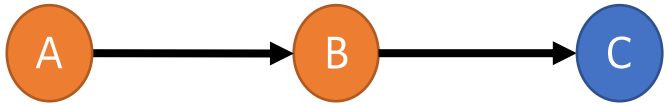
2. **CONSULTANT**
   - Outsider intervening in the group



3. **GATEKEEPER**
   - Outsider blocking access to another group



4. **REPRESENTATIVE**
   - Insider connecting to another group



5. **LIAISON**
   - Outsider linking two groups to which she does not belong to