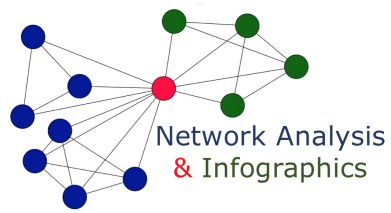
Network Definition

Dr Daniele Rotolo

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



Week 2

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

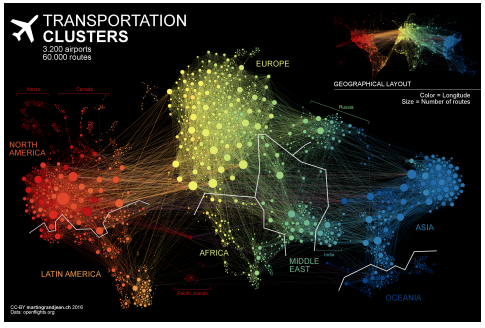
Overview

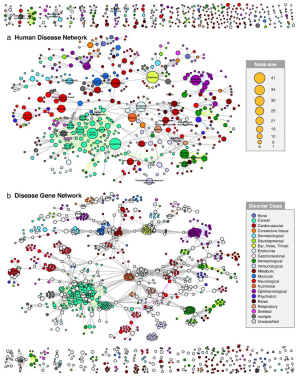
1. Defining a network
2. 'Social' Network Analysis and its origins
3. Branches of network analysis
4. Key definitions

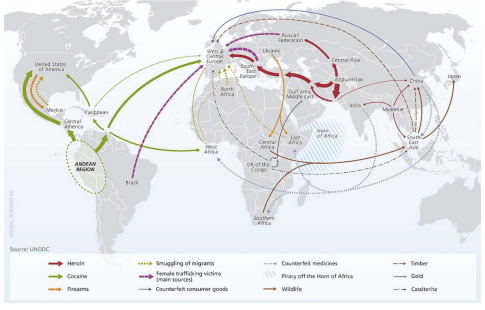
Defining a network

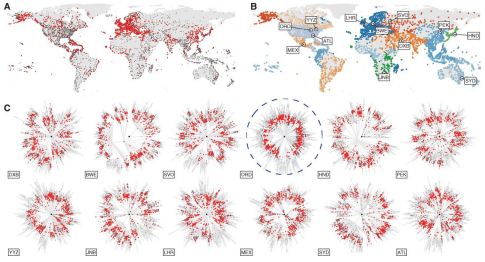
Defining a network

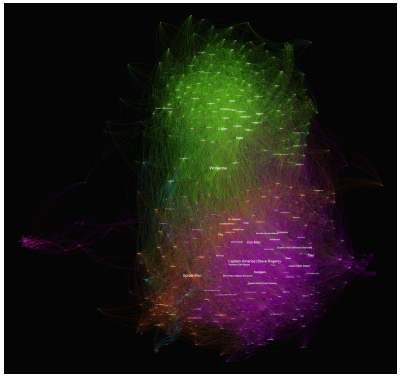
From natural to social phenomena

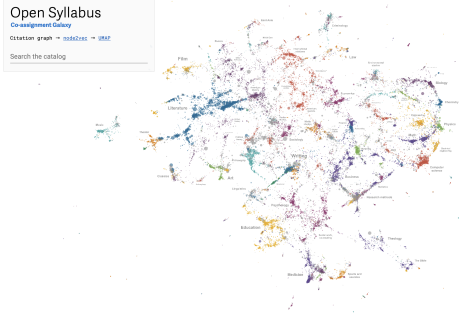










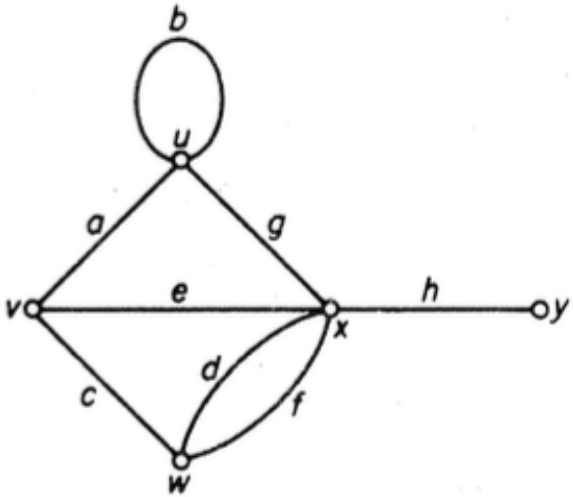


Defining a network

Graph theory

We rely on graph theory to define (social) networks

- Graphs model relations between pairs of objects
- Terms and vocabulary to denote structural properties
- Formulas to quantify structural properties

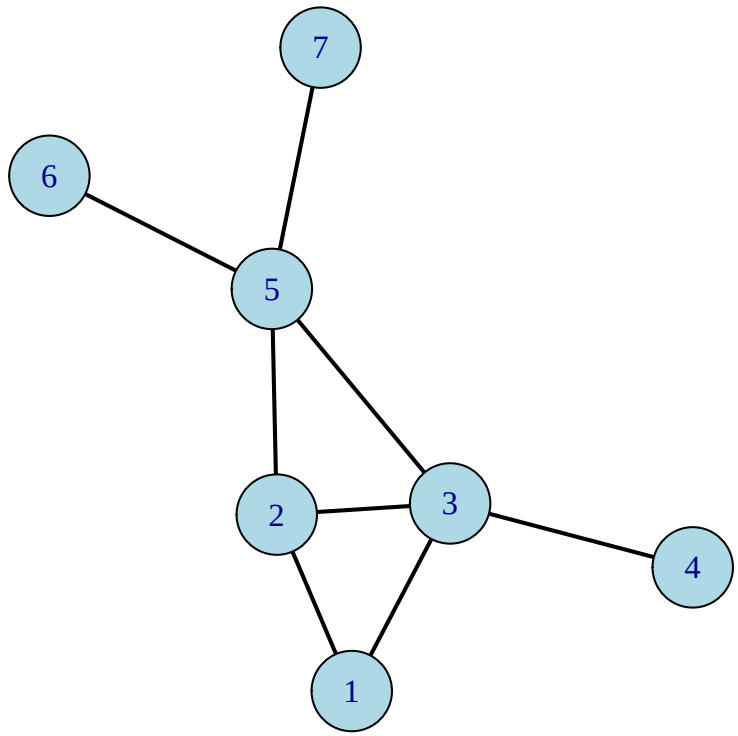


Source: [Bondy and Murty, 1976]

Defining a network

Graph theory

- A graph is defined as: $G(N, E)$
- N vertices, $N = n_1, n_2, ..., n_N$
- E edges, $E = e_1, e_2, ..., e_E$
- Example: $G(7, 8)$

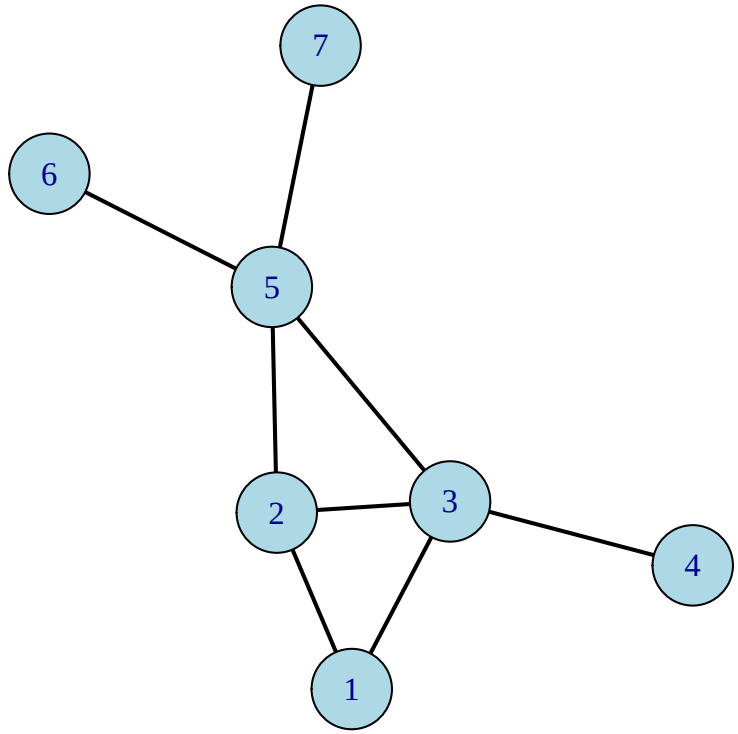


Graph theory	**Network science**
Graph	Network
Vertex	Node
Edge	Link

Defining a network

Graph theory

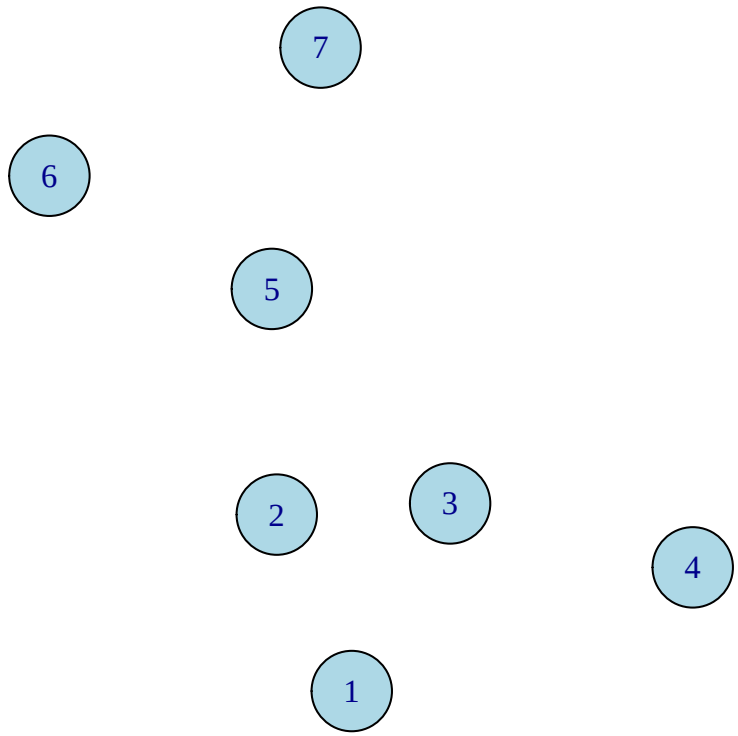
- "A network consists of a graph and additional information on the vertices or the lines of the graphs"
  [de Nooy et al., 2005]
- "Networks can be conceived as metaphors to visualise a (finite) set of edges (links, ties) among a (finite) set of nodes (vertices)"
  [Wasserman and Faust, 1994]



Defining a network
Definition

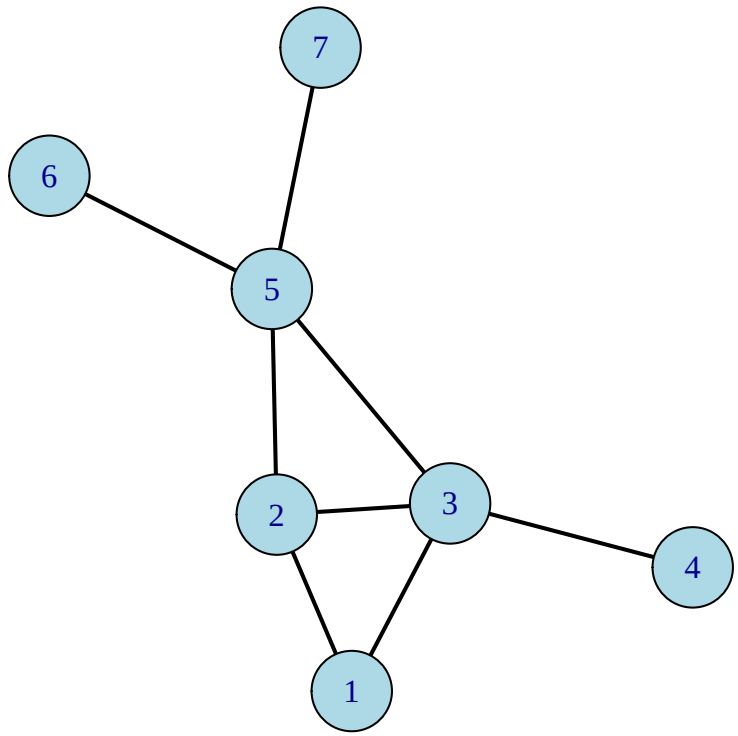
Nodes: entities on which the analysis focuses

- individuals (e.g. family members, employees, inventors)
- organisations (e.g. firms, universities, business units)
- documents (e.g. publications, patents)
- geographical areas (e.g. cities, regions, countries)
- ...

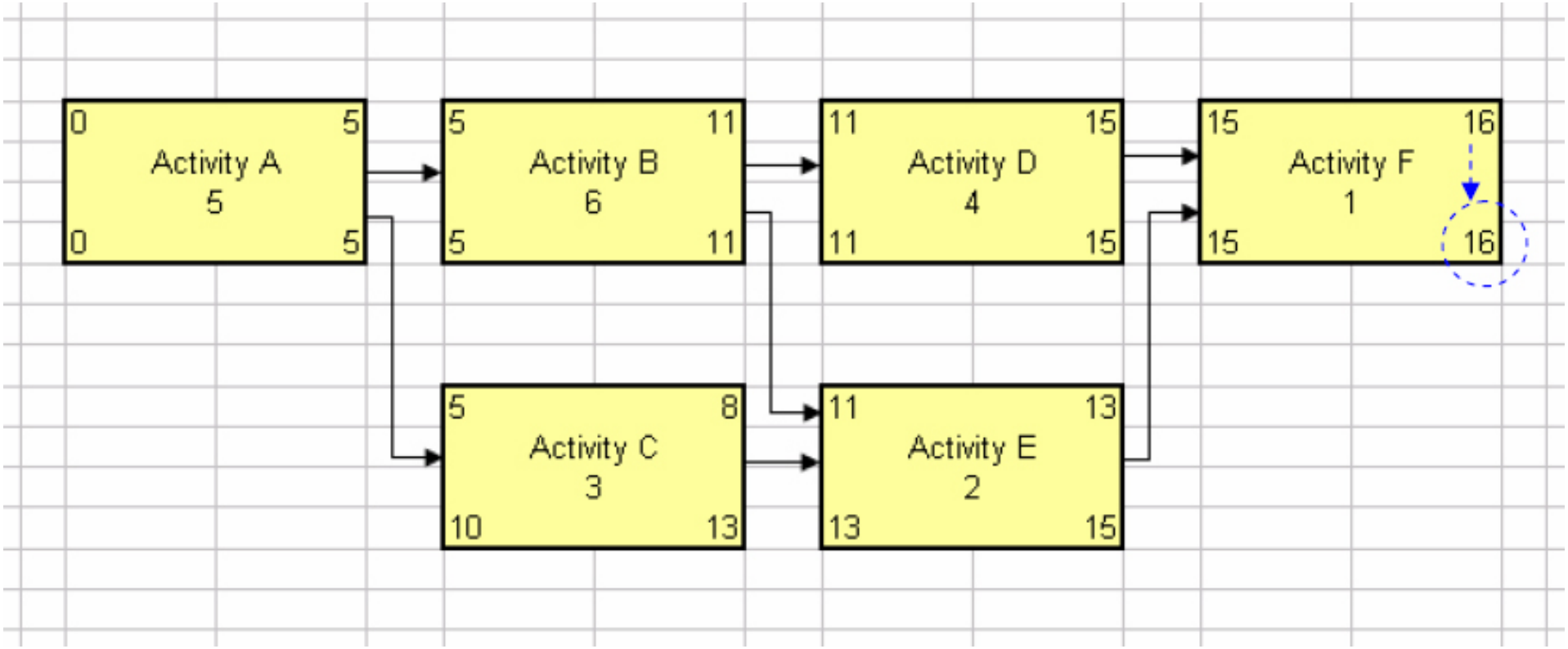


Links (ties): ties linking the entities (modes of interactions)

- formal and informal relations (e.g. collaboration, friendship)
- biological relations (e.g. kinship)
- transfer of knowledge and resources (e.g. licensing, alliances)
- association or affiliation (e.g. social club, employer)
- physical connections (e.g. bridges)
- events (e.g. sending e-mails, transactions)
- ...

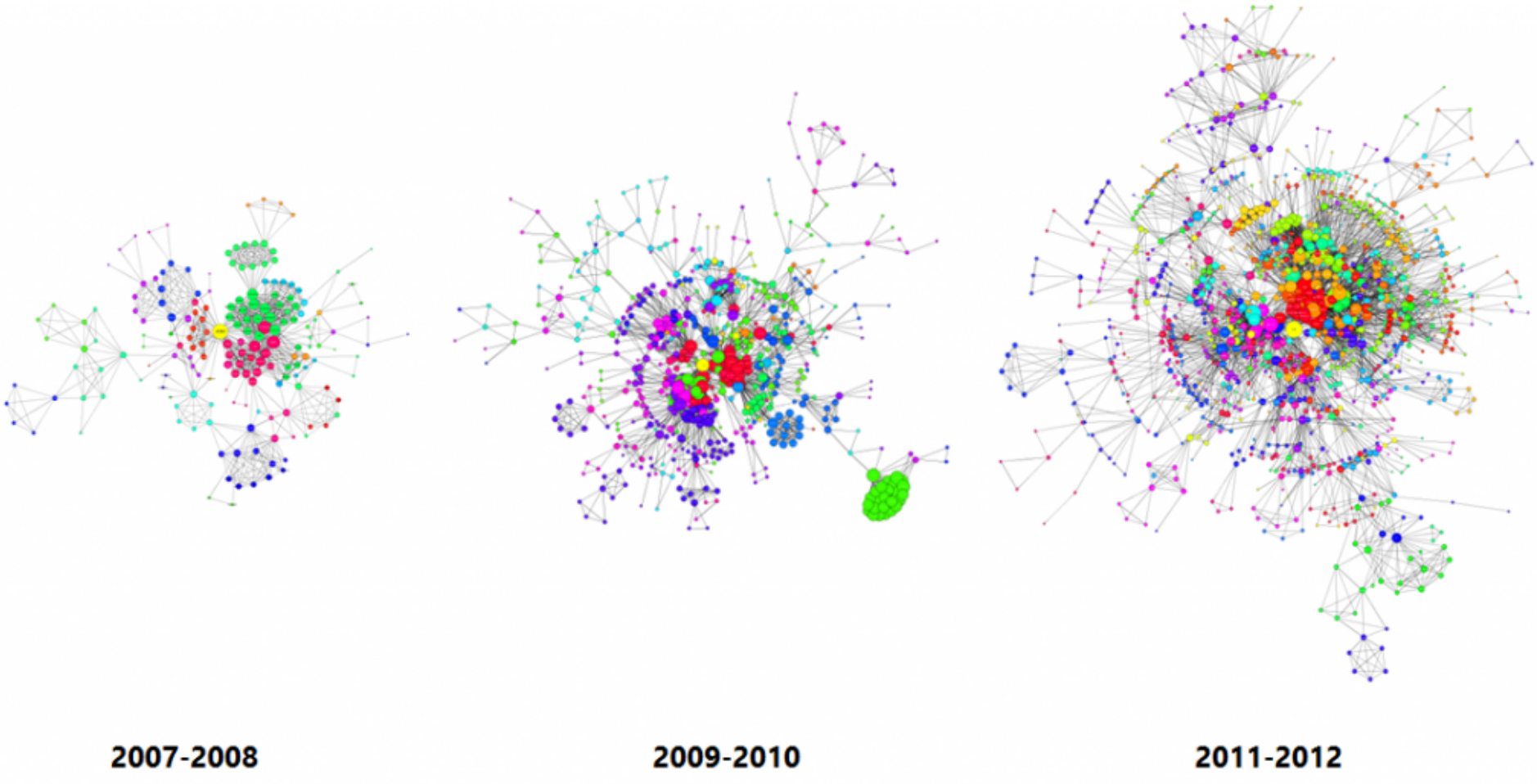


Defining a network
Example



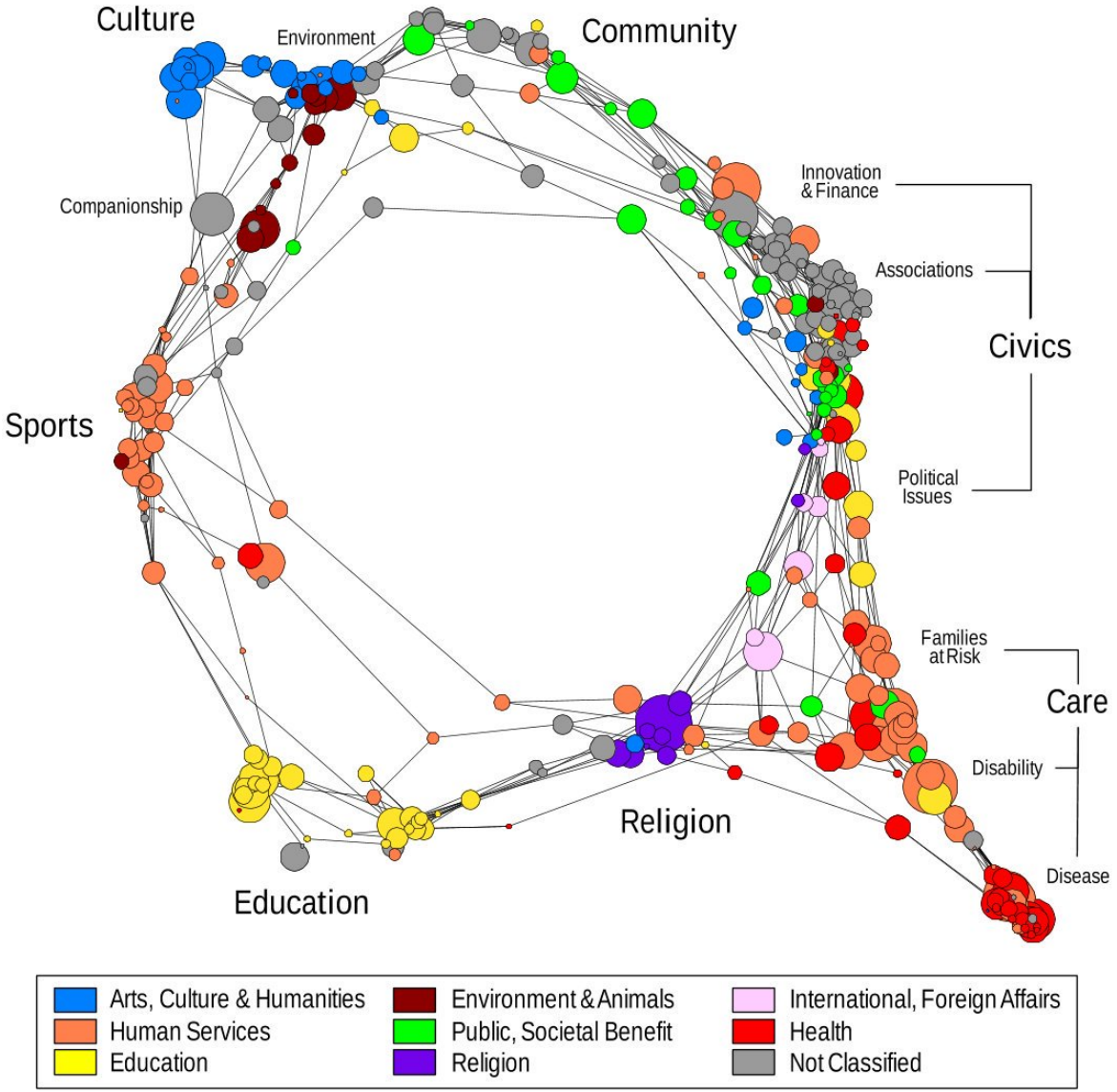
Source: Critical Path Method [www.pmi.org]

Defining a network
Example



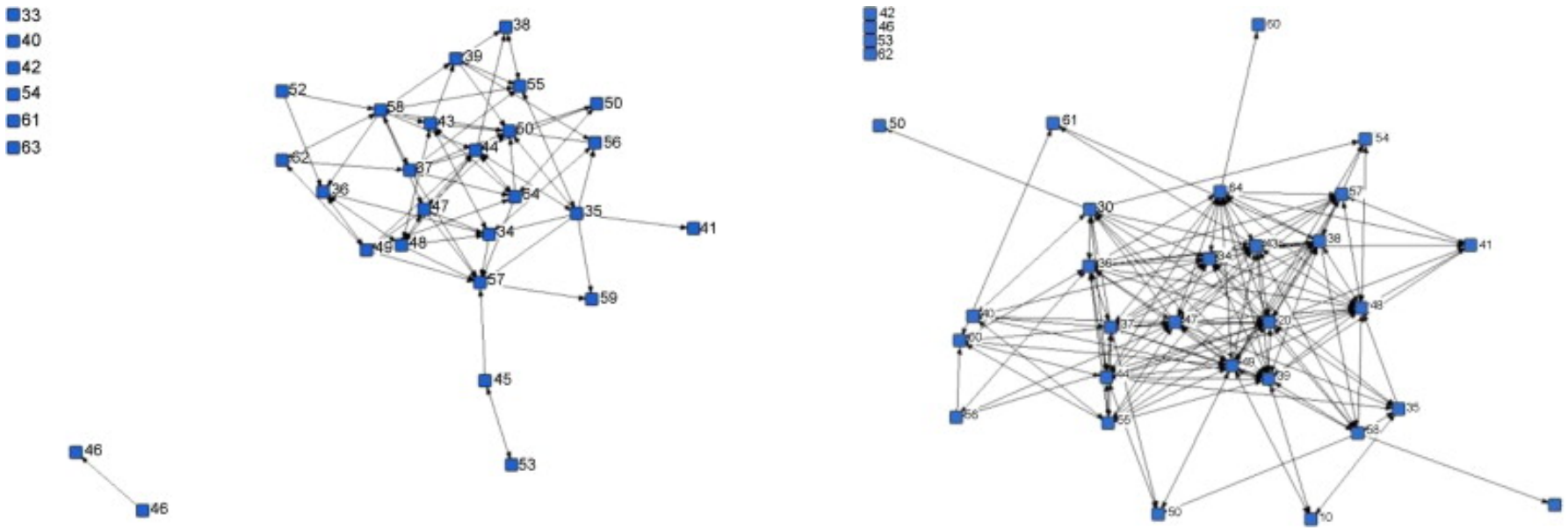
Source: Apple's inventor network [http://www.kenedict.com]

Defining a network
Example



Source: 357 topics from the websites of 125,000 non-profit organizations in the US [Klavans and Boyack, 2014]

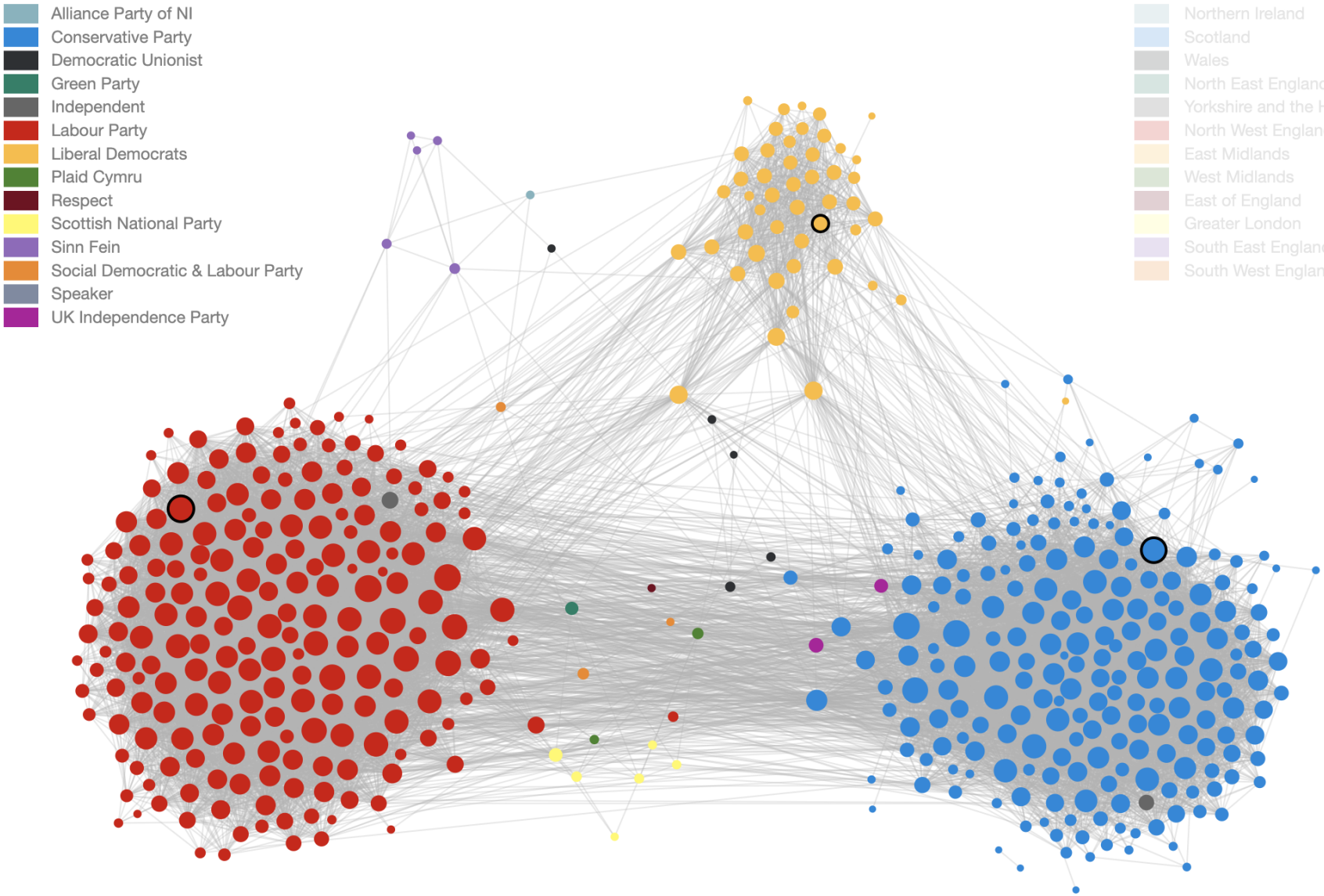
Defining a network
Example



Source: Local knowledge network in 2002 (left) 2006 (right) [Giuliani, 2013]

Defining a network

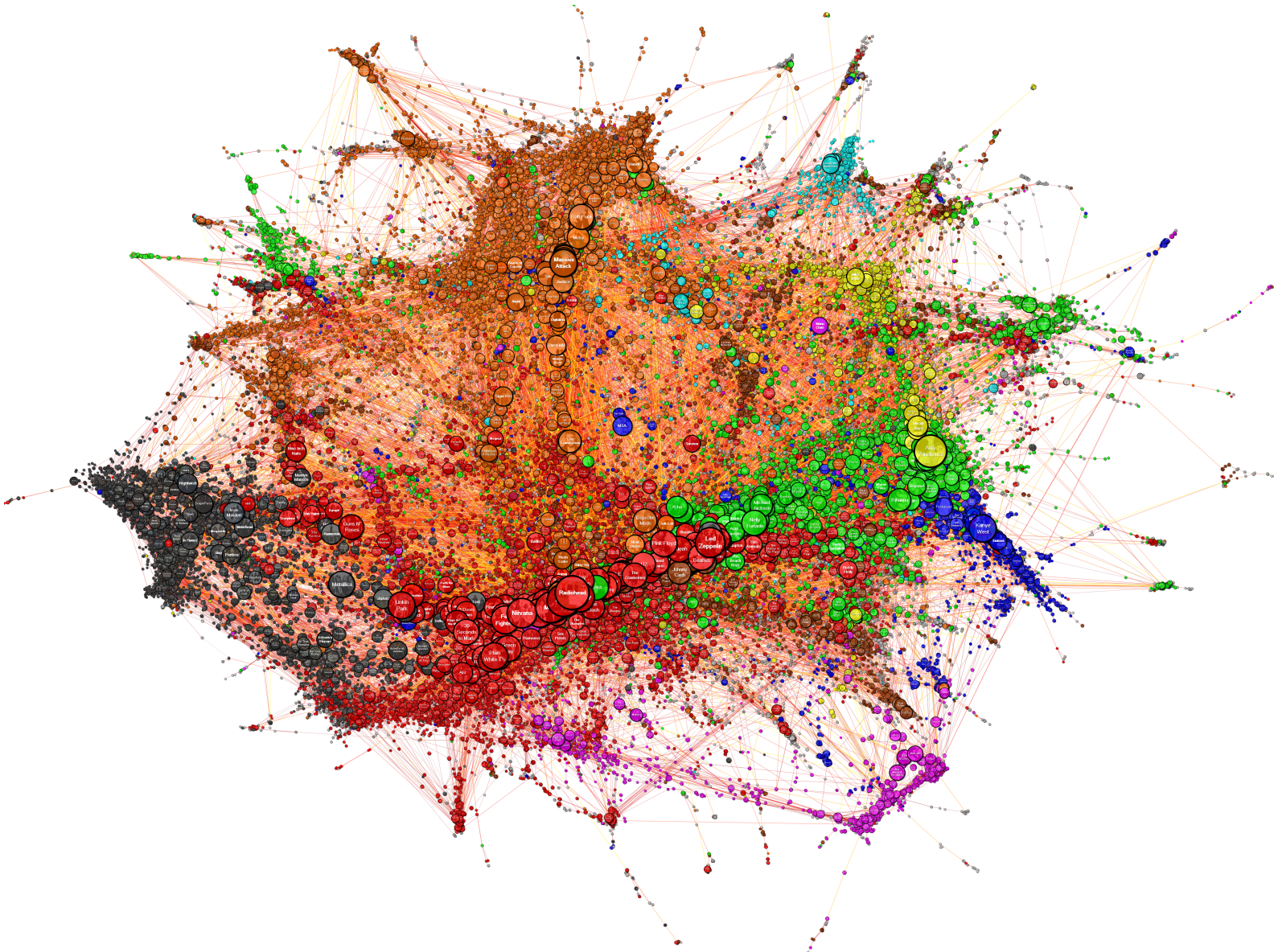
Example



Source: 2015 network of MPs based on Twitter links (`www.nesta.org.uk/blog/twitter-network-uk-mps/`)

Defining a network

Example



Source: Last.fm's artist similarity network (`http://sixdegrees.hu/last.fm/index.html`)

Defining a network
Example

Network	Nodes	Links	Directed / Undirected	N	L	‹K›
Internet	Routers	Internet connections	Undirected	192,244	609,066	6.34
WWW	Webpages	Links	Directed	325,729	1,497,134	4.60
Power Grid	Power plants, transformers	Cables	Undirected	4,941	6,594	2.67
Mobile-Phone Calls	Subscribers	Calls	Directed	36,595	91,826	2.51
Email	Email addresses	Emails	Directed	57,194	103,731	1.81
Science Collaboration	Scientists	Co-authorships	Undirected	23,133	93,437	8.08
Actor Network	Actors	Co-acting	Undirected	702,388	29,397,908	83.71
Citation Network	Papers	Citations	Directed	449,673	4,689,479	10.43
E. Coli Metabolism	Metabolites	Chemical reactions	Directed	1,039	5,802	5.58
Protein Interactions	Proteins	Binding interactions	Undirected	2,018	2,930	2.90

Source: N (nodes), L (links), ¡K¿ (average degree) [Barabási, 2016]

‘Social’ Network Analysis and its origins

‘Social’ Network Analysis and its origins

Social Network Analysis is a perspective that encompasses theories, models, and applications that are expressed in terms of relations among social units (e.g. individuals, groups, communities, organisations, etc.)
[Wasserman and Faust, 1994]

‘Social’ Network Analysis and its origins

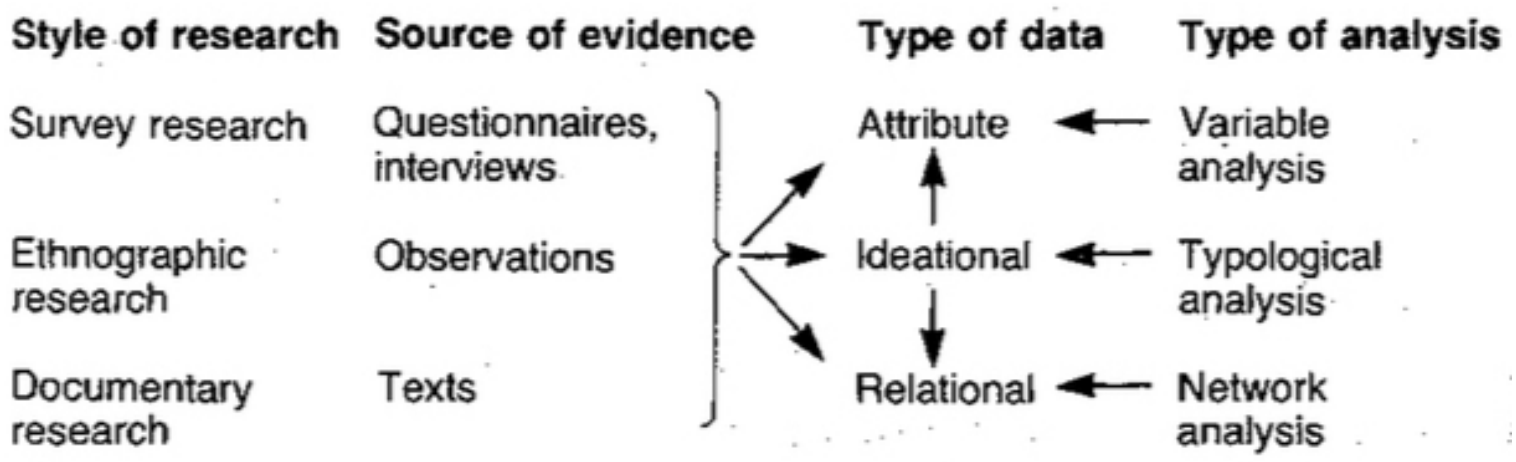


Figure 1.1 *Types of data and analysis*

Source: [Scott, 2000]

‘Social’ Network Analysis and its origins

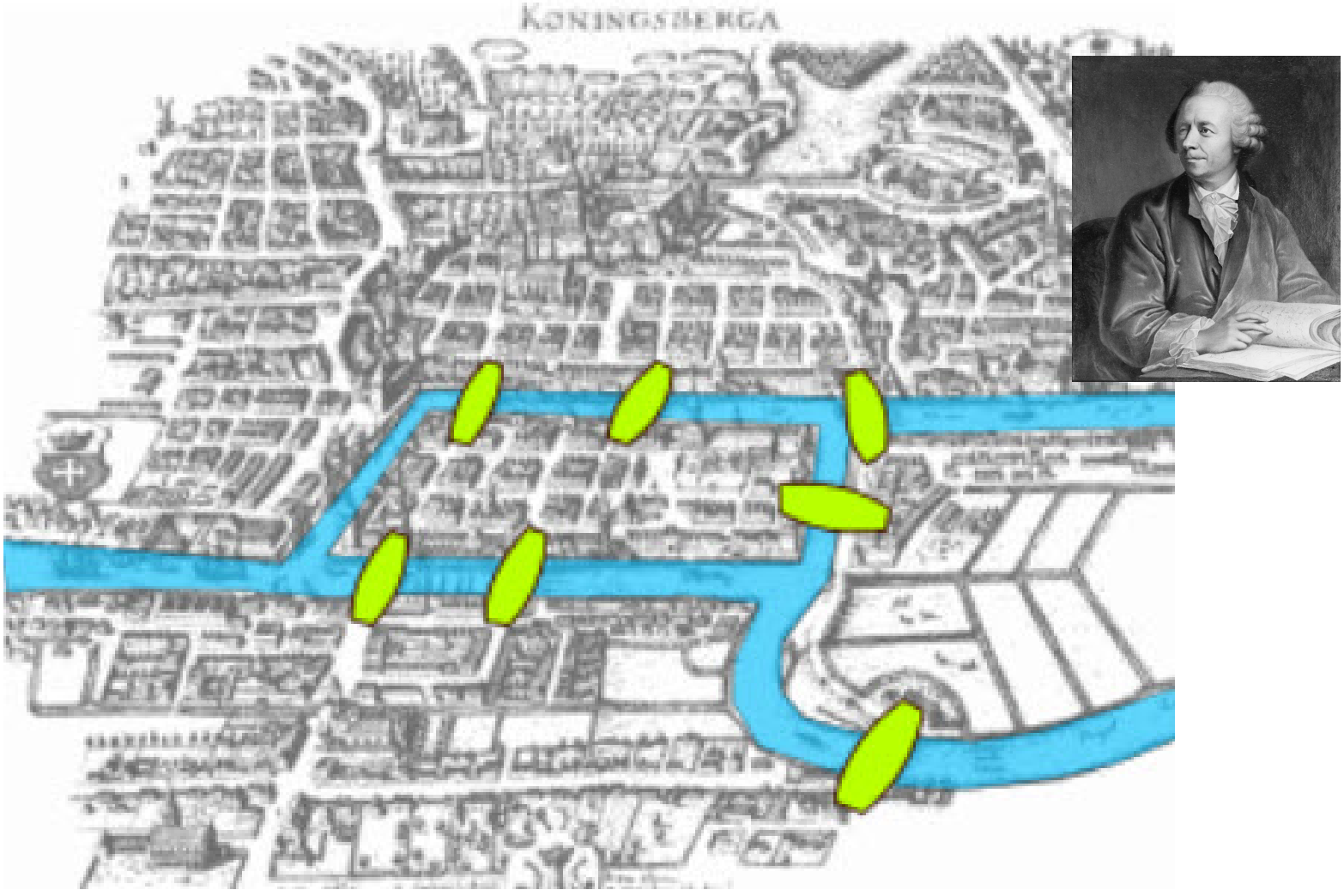
- Mathematics
- Anthropology
- Psycology

‘Social’ Network Analysis and its origins

Mathematics: development of the mathematical notation (graph theory) to describe and identify solutions of relatively complex problems

- Euler’s Konigsberg Bridge problem (1736)
- Markov-chain of probabilities

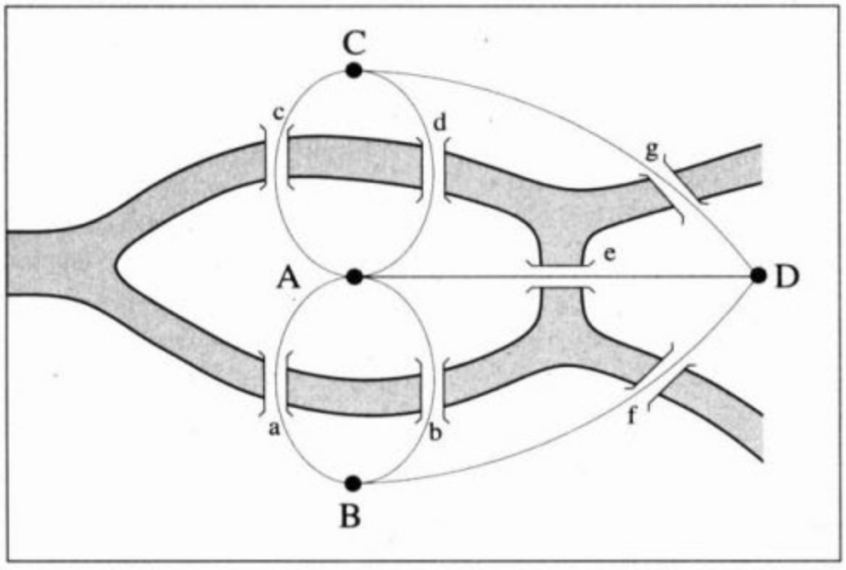
‘Social’ Network Analysis and its origins



Source: Euler and the seven bridges of Konigsberg [http://eulerarchive.maa.org]

When poll is active, respond at **PollEv.com/rotolo**

Does a path crossing all bridges once exist?



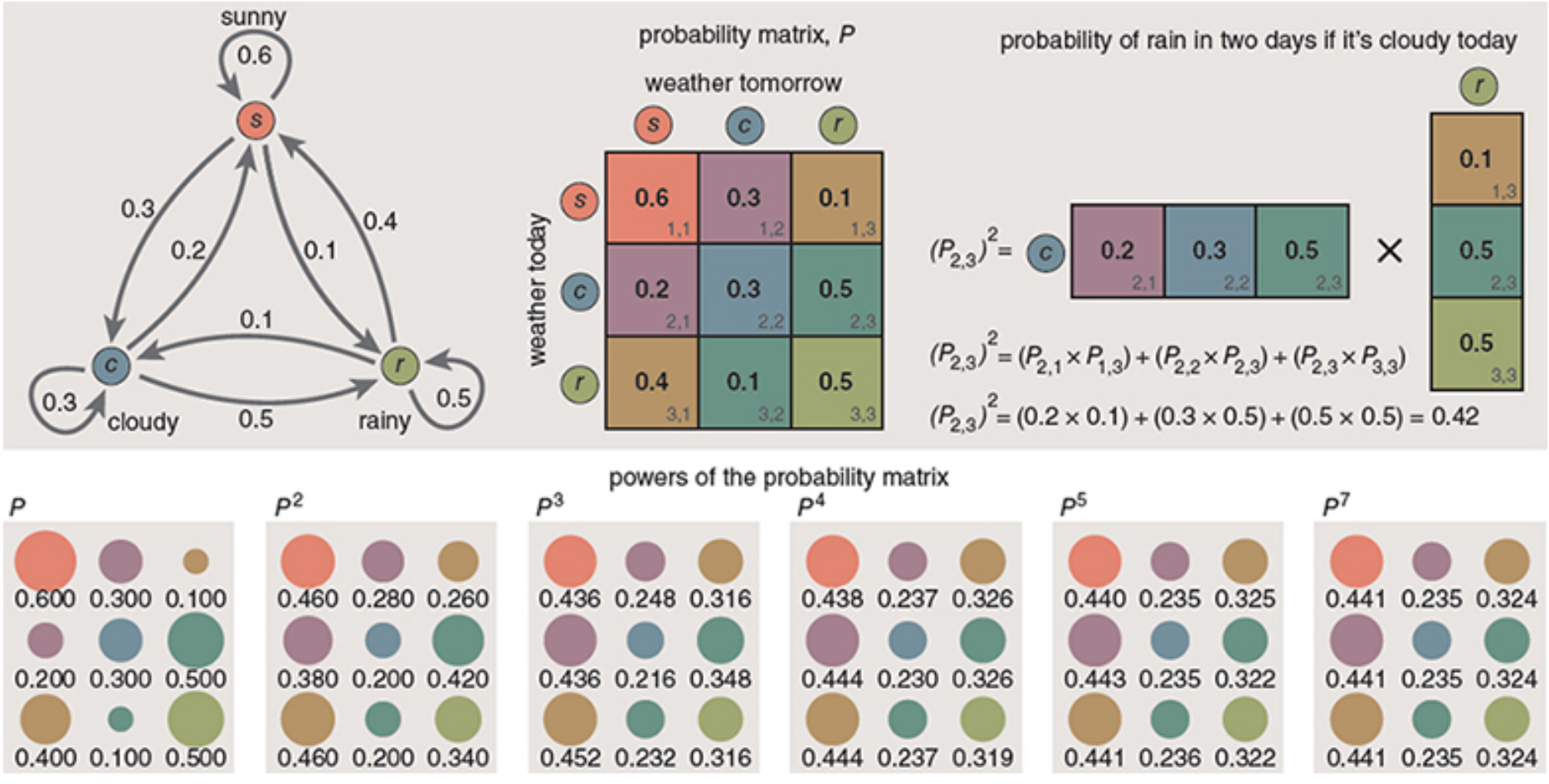
Yes

No

I do not know

Total Results: 0

‘Social’ Network Analysis and its origins

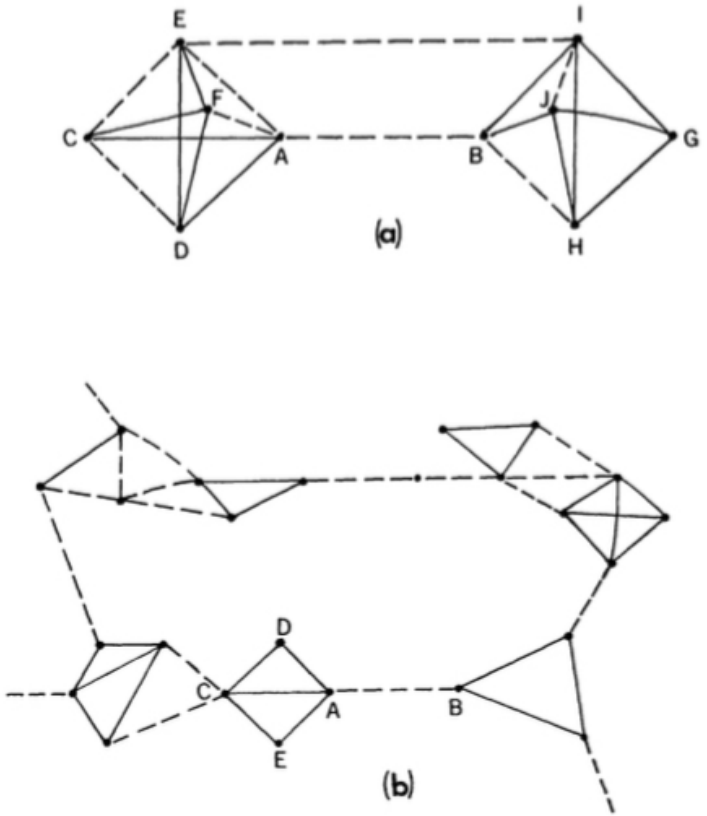


Source: Markov-chain for weather forecasting [www.americanscientist.org]

‘Social’ Network Analysis and its origins

Anthropology

- 1950s: researchers of the Dept. of Social Anthropology at Manchester University examined the structure of community relations in tribal and villages societies
- 1960s-1970s: researchers at Harvard developed the mathematical component of social network analysis (e.g. network measures)

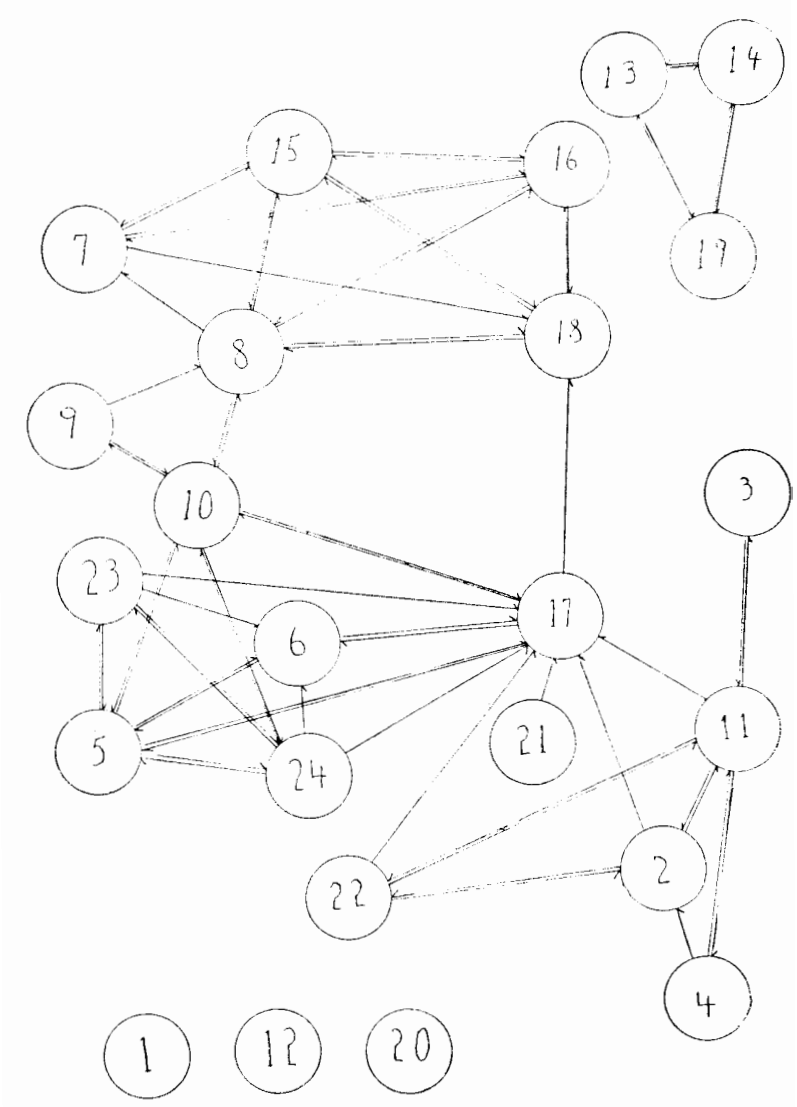


Source: The strength of weak ties [Granovetter, 1973]

‘Social’ Network Analysis and its origins

Physicology

- Gestalt studies in the 1920s on how mind works (organised patterns that structure thoughts and perceptions)
- Jacob Moreno and the development of sociometry (early 1930s)
- Kurt Lewin studied group behavior (individual-goal connections) (1930s)
- Fritz Heider’s balance theory (1940s)



Source: Sociogram of relationships between people [Moreno, 1934]

Branches of network analysis

Branches of network analysis

Descriptive network analysis

- An observed network is analysed by means of measures
- Network-level measures
- Node-level measures

Modelling and inference of networks

- Mathematical models
  Based on 'simple' probabilist rules to capture specific mechanisms (e.g. Erdós-Rényi networks, 'the rich get richer')
- Statistical models
  The observed network is considered as one of the possible realisation of a process – a model that aims to fit to the observed data is specified (e.g. explanatory power of certain variables)

Branches of network analysis

Network Analysis is not a theory *per se*, but it a methodological tool to support the development of theories [Borgatti and Halgin, 2011]

- Network theory: mechanisms and processes that interact with network structures to produce certain outcomes for individuals, groups, and organisations (e.g. firms' performance, individuals' creativity)

- Theory of networks: mechanisms and processes that explain why certain networks have certain structures (i.e. antecedents of network properties)

Branches of network analysis

	Dependent variable	
Independent variable	Nonnetwork variable as outcome	Network variable as outcome
Nonnetwork variable as antecedent	(Nonnetwork theory)	Theory of networks
Network variable as antecedent	Network theory	Network theory of networks

Source: [Borgatti and Halgin, 2011]

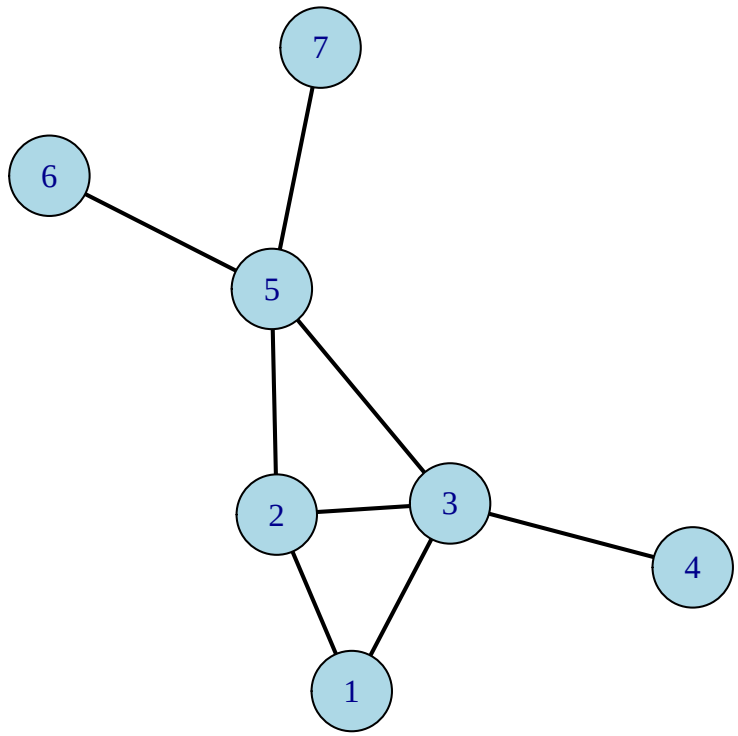
Key definitions

Key definitions

Undirected vs. directed and unweighted vs. weighted networks

- Tie directionality

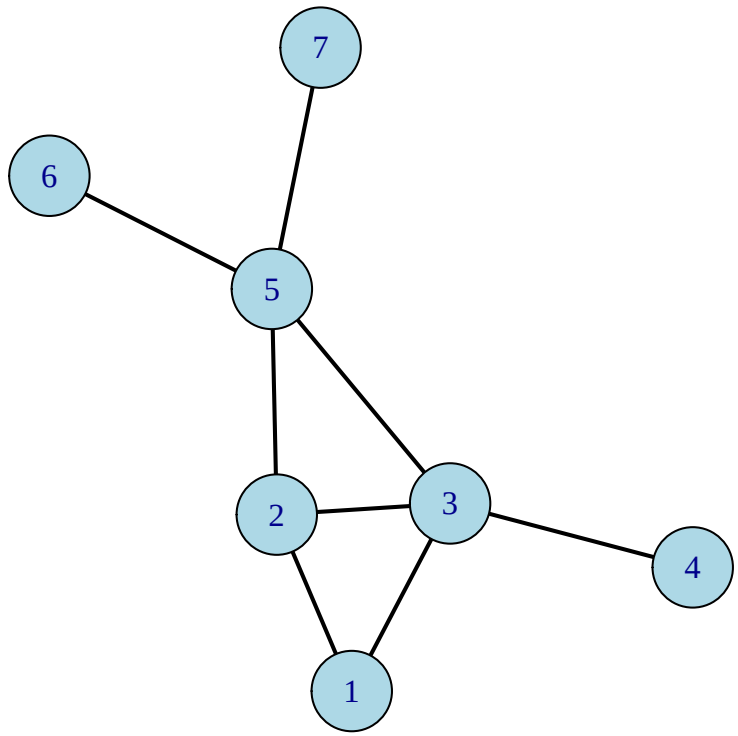
- Tie value



Key definitions

Undirected vs. directed and unweighted vs. weighted networks

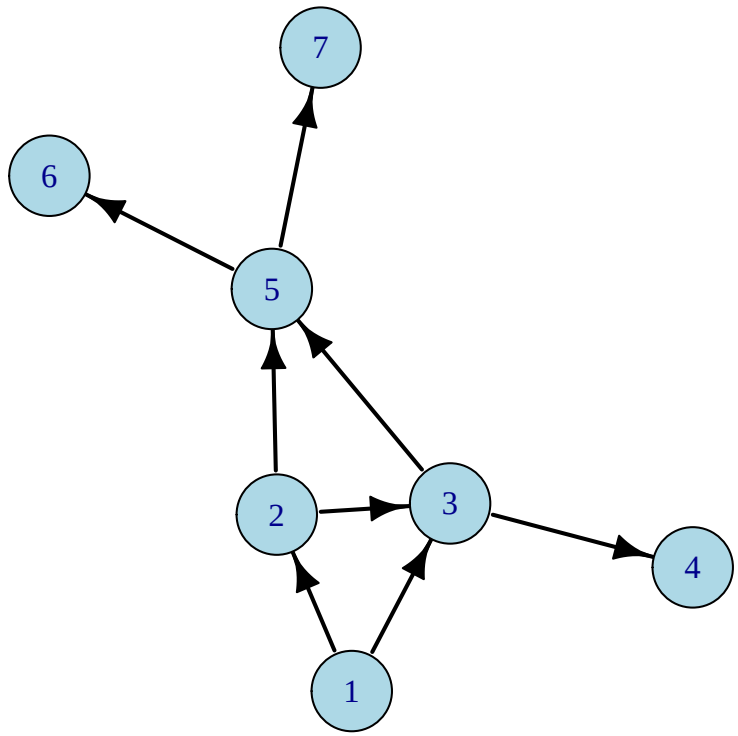
- Tie directionality
  - Undirected networks: ties with no directionality (e.g. co-occurrence, alliances)
- Tie value



Key definitions

Undirected vs. directed and unweighted vs. weighted networks

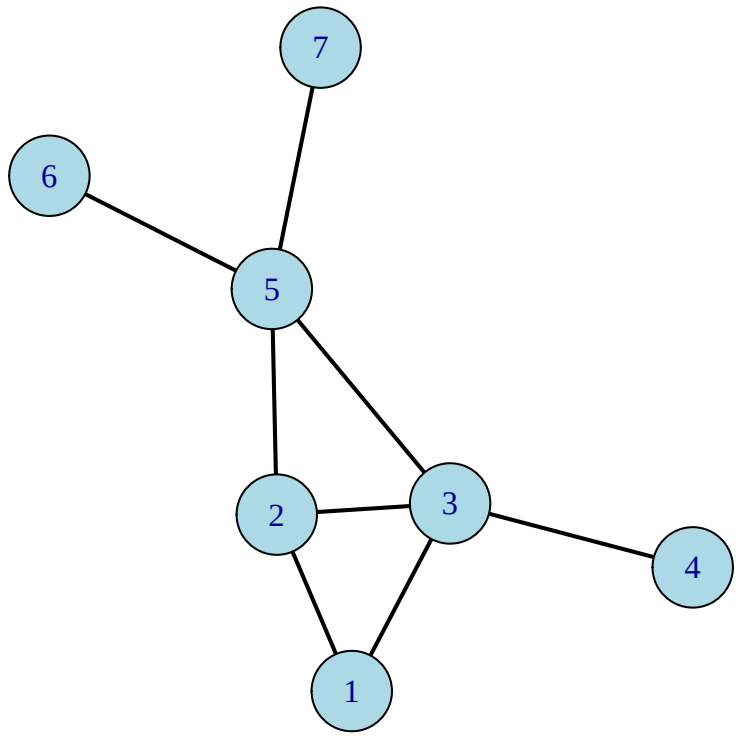
- Tie directionality
  - Undirected networks: ties with no directionality (e.g. co-occurrence, alliances)
  - Directed networks: ties with directionality (e.g. communication, transactions)
- Tie value



Key definitions

Undirected vs. directed and unweighted vs. weighted networks

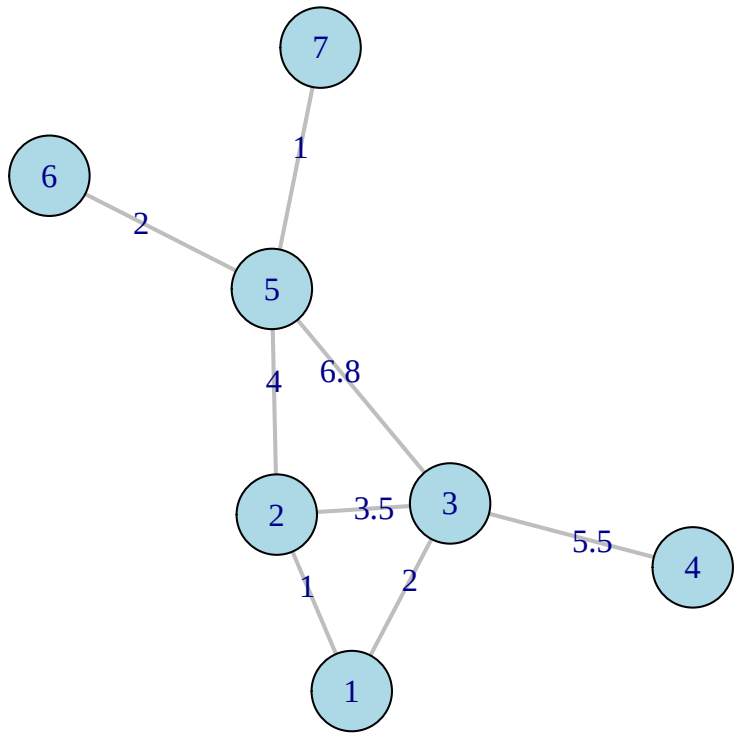
- Tie directionality
  - Undirected networks: ties with no directionality (e.g. co-occurrence, alliances)
  - Directed networks: ties with directionality (e.g. communication, transactions)
- Tie value
  - Unweighted networks: ties with no weight (e.g. presence/absence)



Key definitions

Undirected vs. directed and unweighted vs. weighted networks

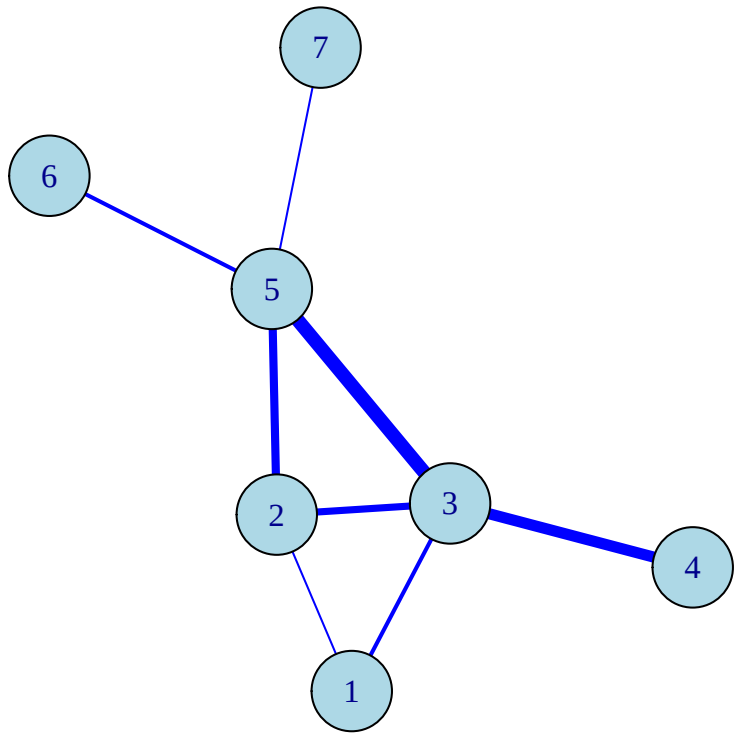
- Tie directionality
  - Undirected networks: ties with no directionality (e.g. co-occurrence, alliances)
  - Directed networks: ties with directionality (e.g. communication, transactions)
- Tie value
  - Unweighted networks: ties with no weight (e.g. presence/absence)
  - Weighted networks: ties with weights (e.g. frequency of interaction, amount of money associated with a transaction)



Key definitions

Undirected vs. directed and unweighted vs. weighted networks

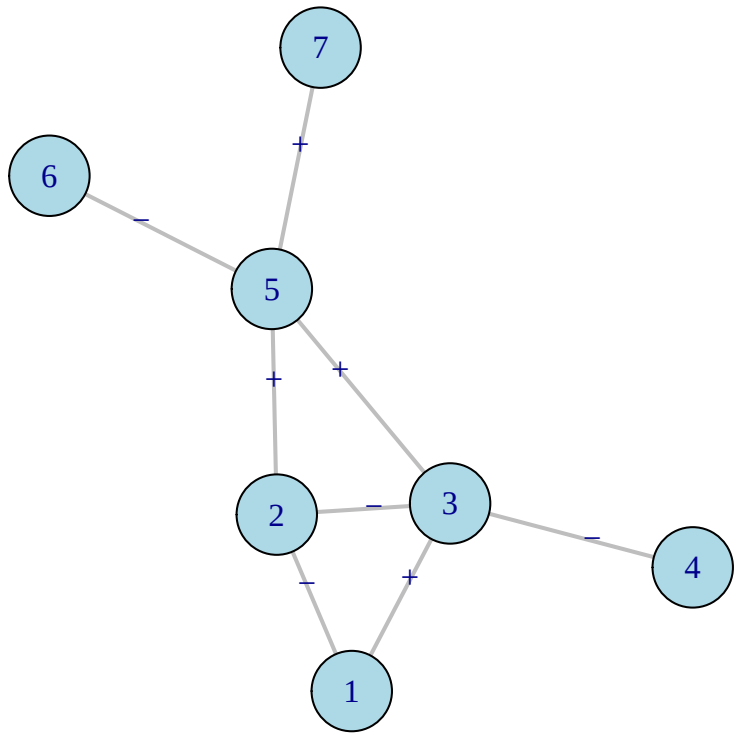
- Tie directionality
  - Undirected networks: ties with no directionality (e.g. co-occurrence, alliances)
  - Directed networks: ties with directionality (e.g. communication, transactions)
- Tie value
  - Unweighted networks: ties with no weight (e.g. presence/absence)
  - Weighted networks: ties with weights (e.g. frequency of interaction, amount of money associated with a transaction)



Key definitions

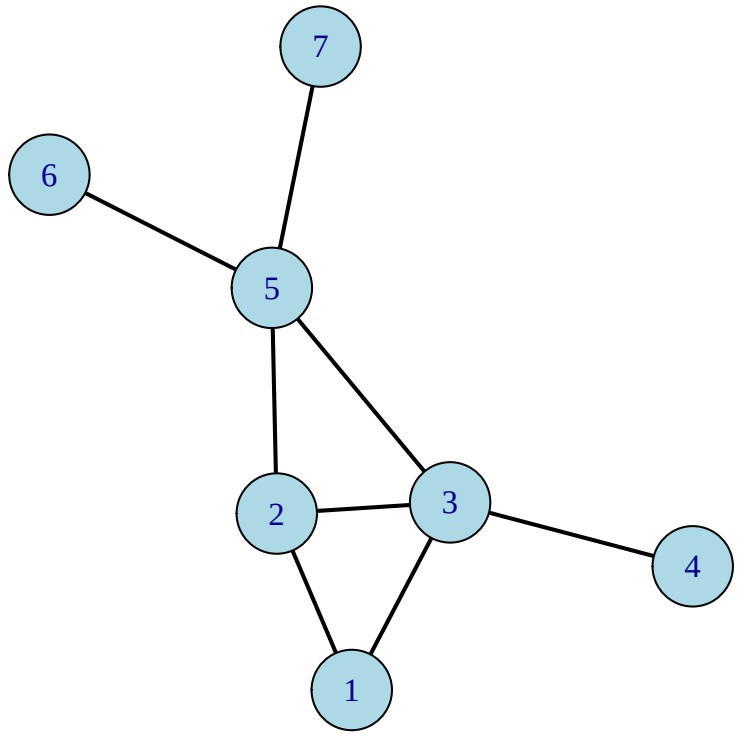
Undirected vs. directed and unweighted vs. weighted networks

- Tie directionality
  - Undirected networks: ties with no directionality (e.g. co-occurrence, alliances)
  - Directed networks: ties with directionality (e.g. communication, transactions)
- Tie value
  - Unweighted networks: ties with no weight (e.g. presence/absence)
  - Weighted networks: ties with weights (e.g. frequency of interaction, amount of money associated with a transaction)
  - Signed networks: positive and negative ties (e.g. loves vs. hates)



Key definitions
Adjacency matrix

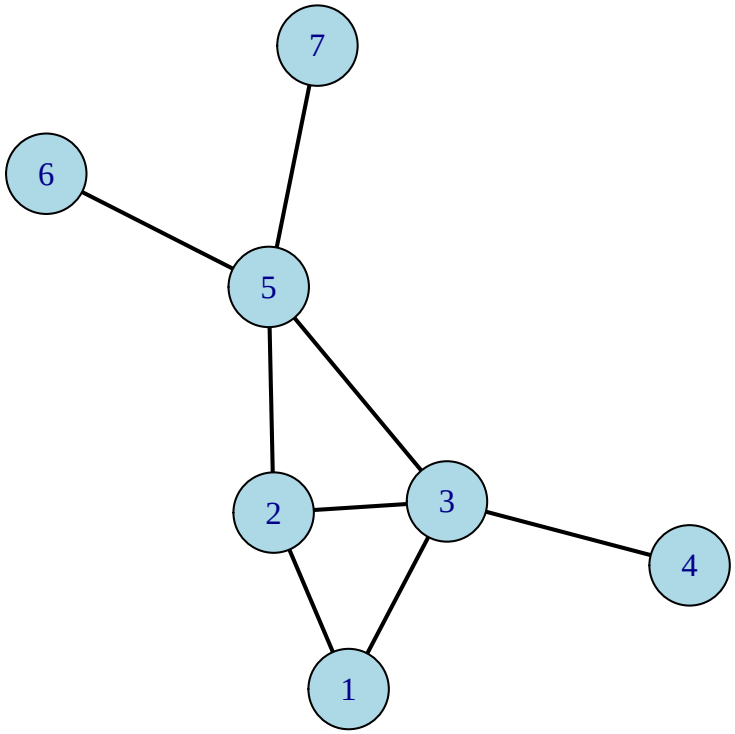
- The (symmetric or asymmetric) matrix representing the connections among nodes is called adjacency matrix
- Relational data can be complemented with attributes of nodes (e.g. gender, education)



$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & a_{ij} & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{pmatrix}$$

Key definitions

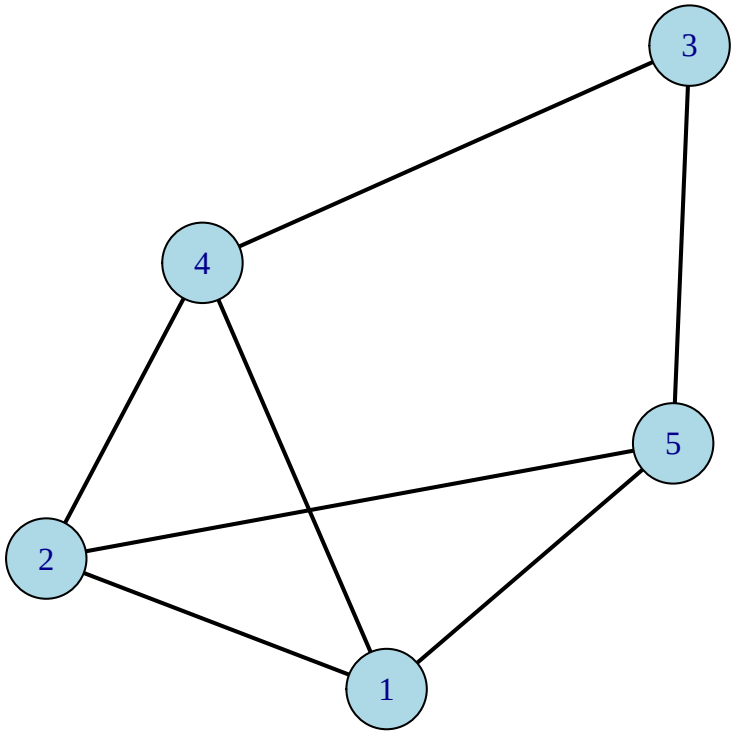
Adjacency matrix: Undirected and unweighted network



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Key definitions

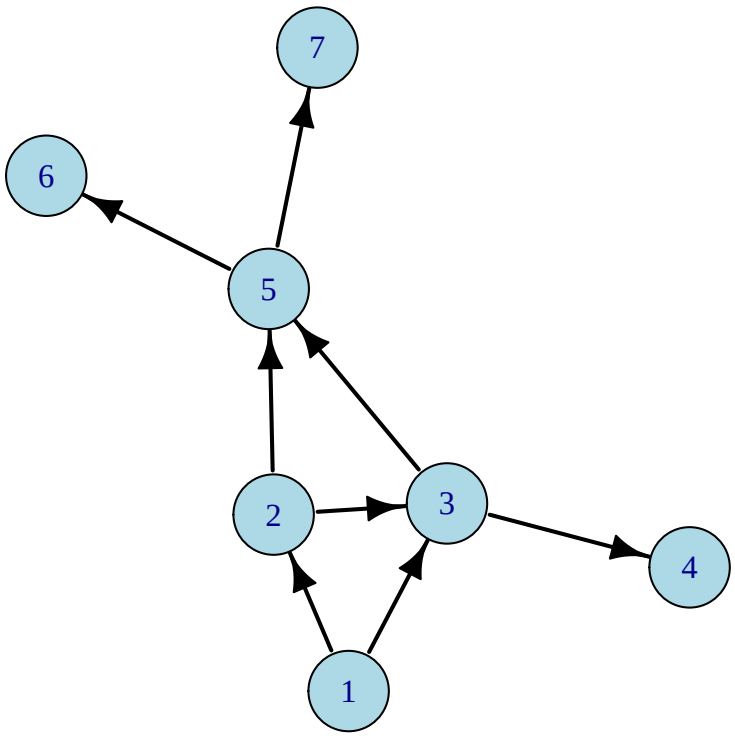
Adjacency matrix: Undirected and unweighted network (exercise)



What is the adjacency matrix of this network?

Key definitions

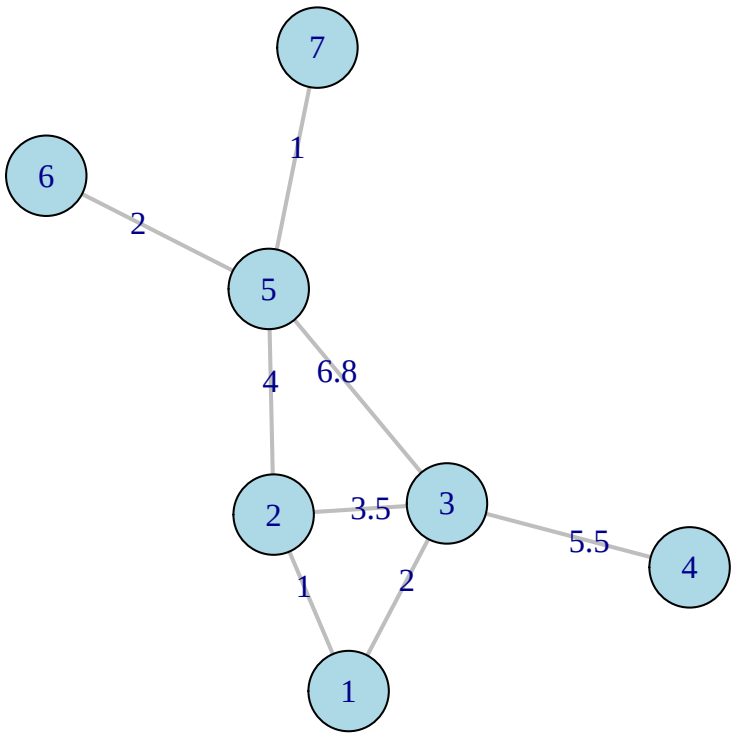
Adjacency matrix: Directed and unweighted network



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Key definitions

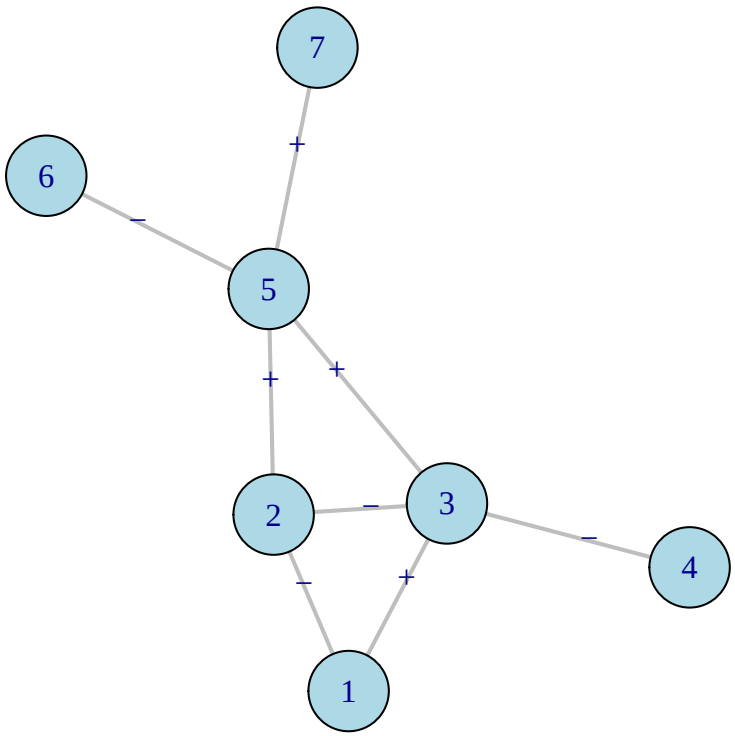
Adjacency matrix: Undirected and weighted network



$$\mathbf{A} = \begin{pmatrix} 0 & 1.0 & 2.0 & 0.0 & 0.0 & 0 & 0 \\ 1 & 0.0 & 3.5 & 0.0 & 4.0 & 0 & 0 \\ 2 & 3.5 & 0.0 & 5.5 & 6.8 & 0 & 0 \\ 0 & 0.0 & 5.5 & 0.0 & 0.0 & 0 & 0 \\ 0 & 4.0 & 6.8 & 0.0 & 0.0 & 2 & 1 \\ 0 & 0.0 & 0.0 & 0.0 & 2.0 & 0 & 0 \\ 0 & 0.0 & 0.0 & 0.0 & 1.0 & 0 & 0 \end{pmatrix}$$

Key definitions

Adjacency matrix: Undirected and signed network

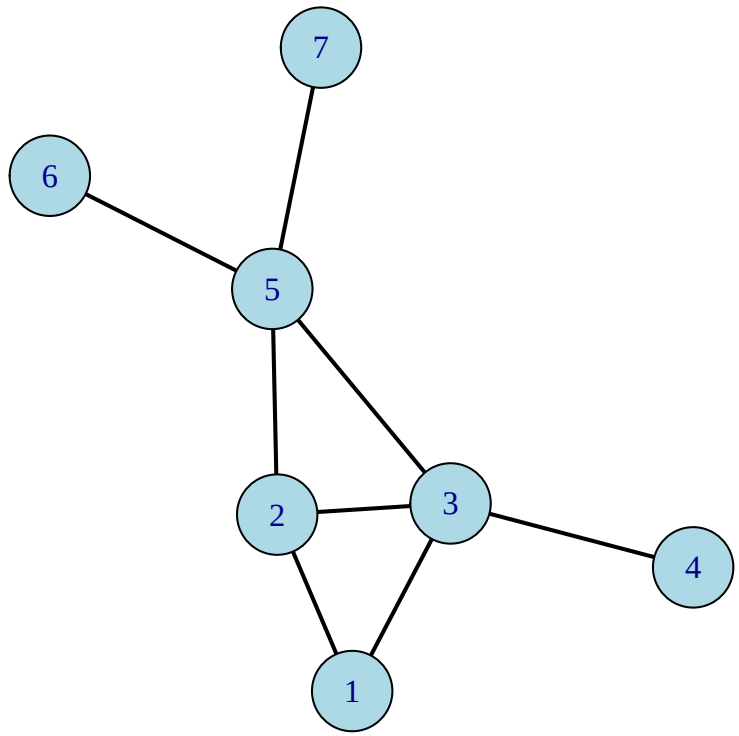


$$\mathbf{A} = \begin{pmatrix} . & - & + & . & . & . & . \\ - & . & - & . & + & . & . \\ + & - & . & - & + & . & . \\ . & . & - & . & . & . & . \\ . & + & + & . & . & - & + \\ . & . & . & . & - & . & . \\ . & . & . & . & + & . & . \end{pmatrix}$$

Key definitions

Adjacency, dyads, and triads

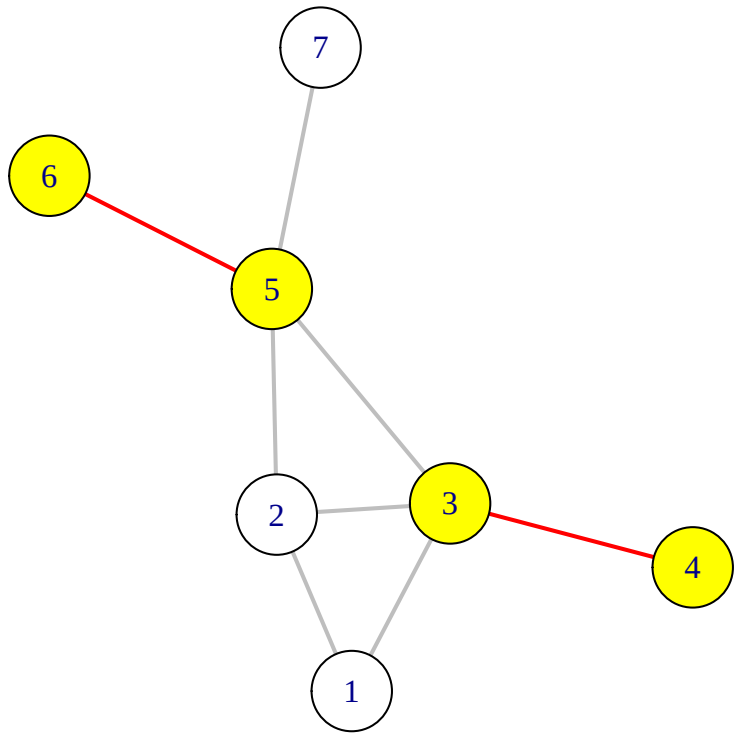
- Two nodes, n_i and n_j are called adjacent if an edge $e_k = (n_i, n_j)$ between them exists



Key definitions

Adjacency, dyads, and triads

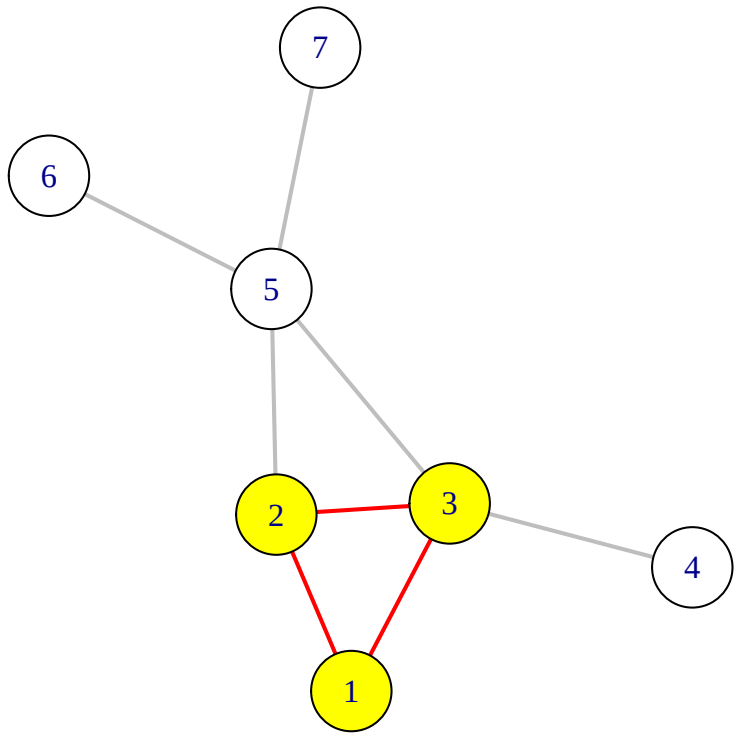
- Two nodes, n_i and n_j are called adjacent if an edge $e_k = (n_i, n_j)$ between them exists
- A dyad is a pair of nodes and the edge between them



Key definitions

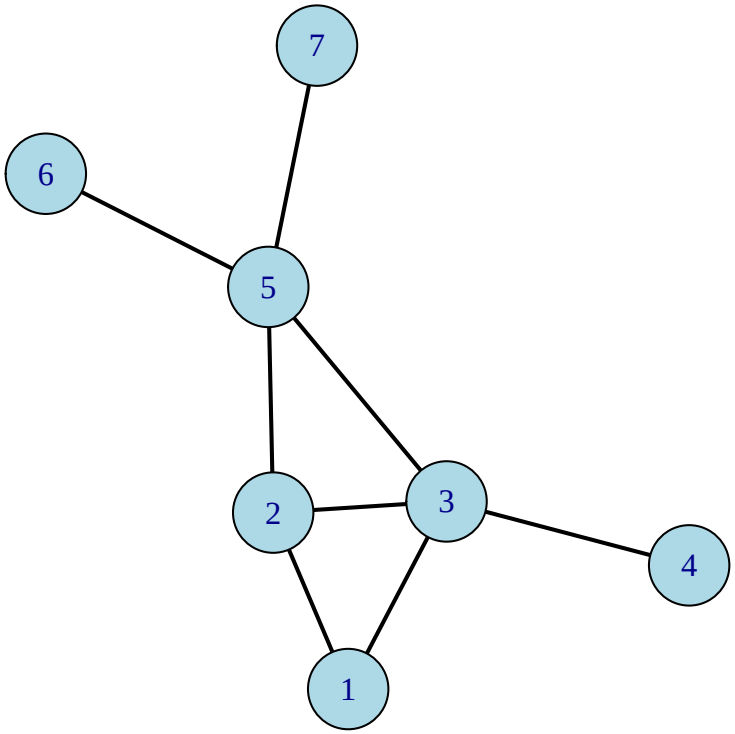
Adjacency, dyads, and triads

- Two nodes, n_i and n_j are called adjacent if an edge $e_k = (n_i, n_j)$ between them exists
- A dyad is a pair of nodes and the edge between them
- A triad is a set of three nodes and the edges between them



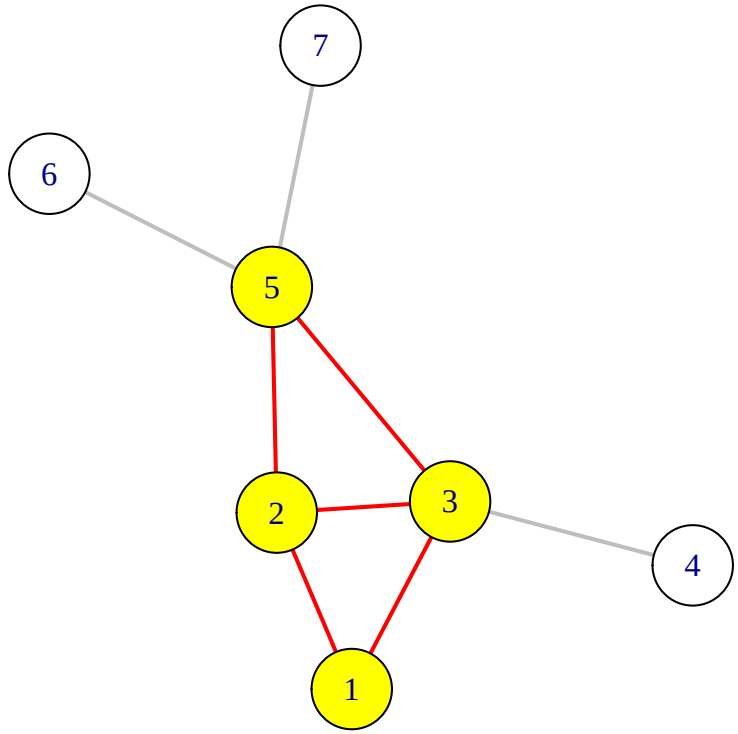
Key definitions
Subgraphs

- A subgraph of G is a graph $G_S(N_S, E_S)$ where $N_S \subseteq N$ and $E_S \subseteq E$



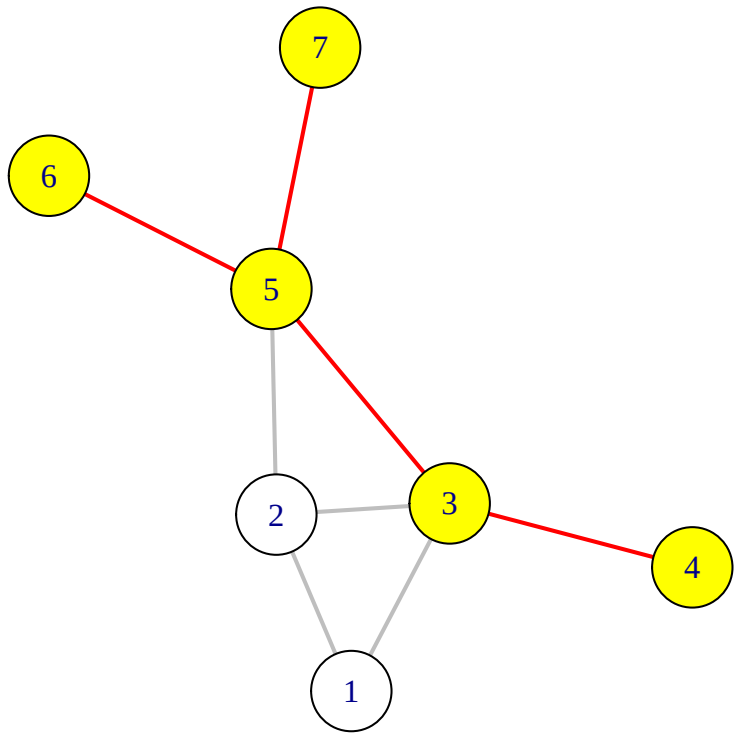
Key definitions
Subgraphs

- A subgraph of G is a graph $G_S(N_S, E_S)$ where $N_S \subseteq N$ and $E_S \subseteq E$
  - Node-generated: $N_S \subseteq N$ and E_S includes the lines linking N_S



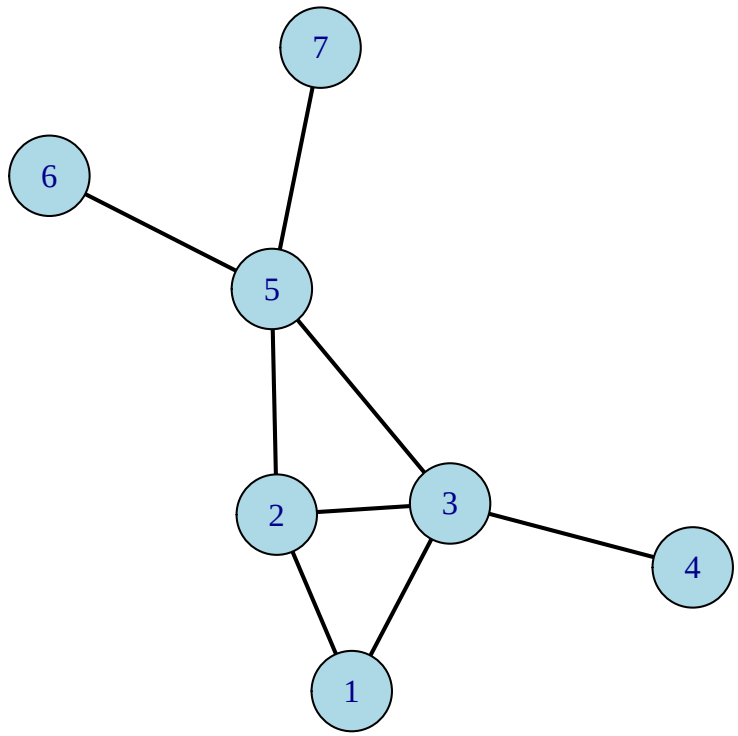
Key definitions
Subgraphs

- A subgraph of G is a graph $G_S(N_S, E_S)$ where $N_S \subseteq N$ and $E_S \subseteq E$
  - Node-generated: $N_S \subseteq N$ and E_S includes the lines linking N_S
  - Line-generated: $E_S \subseteq E$ and N_S includes all the nodes linked by E_S



Key definitions
Walks, trails, paths

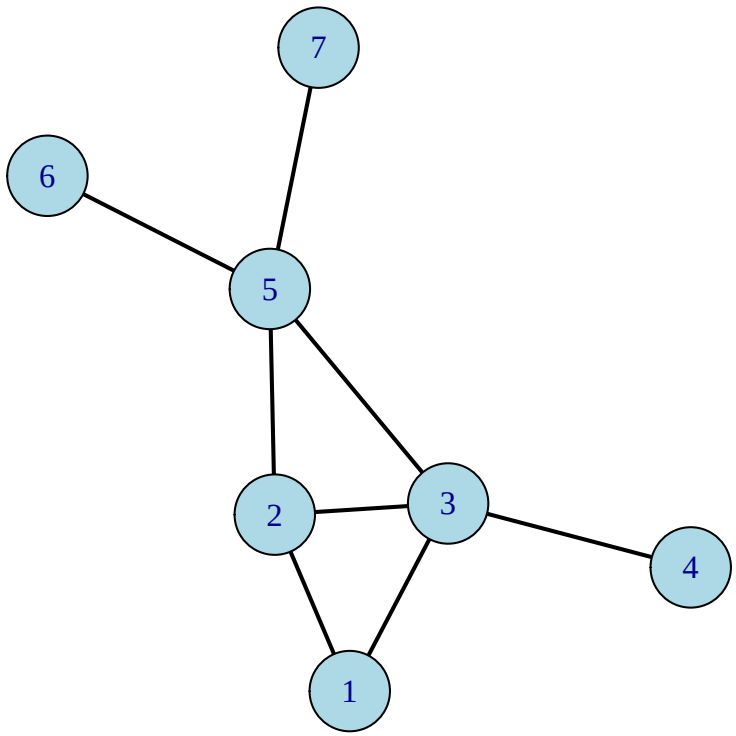
- A walk is a sequence of nodes and lines in which each node is adjacent with the nodes following and preceding it in the sequence



$Walk = (n_1, n_3, n_2, n_1, n_3, n_4)$

Key definitions
Walks, trails, paths

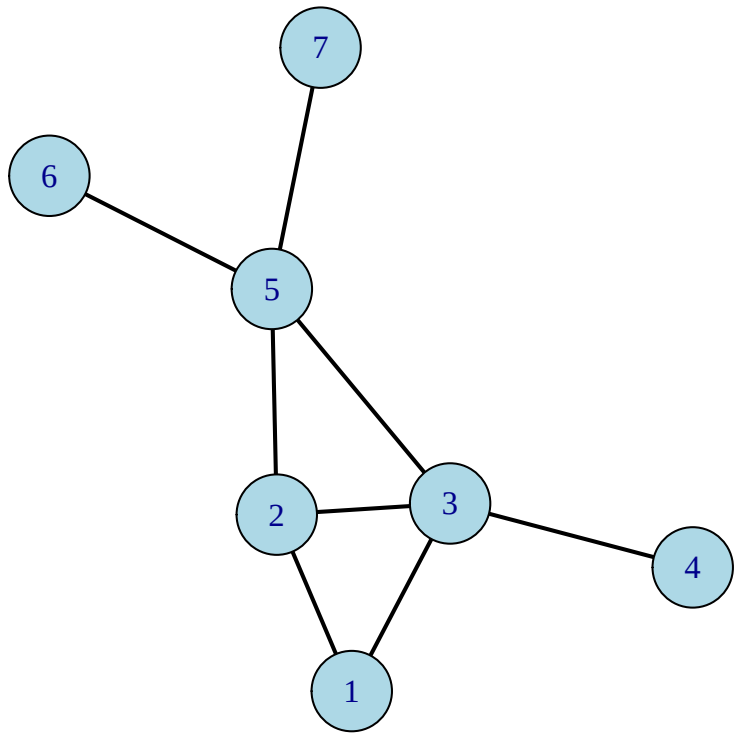
- A walk is a sequence of nodes and lines in which each node is adjacent with the nodes following and preceding it in the sequence
- A trail is a walk in which all links are distinct, but nodes can be repeated



$Trail = (n_3, n_5, n_2, n_1, n_3, n_4)$

Key definitions
Walks, trails, paths

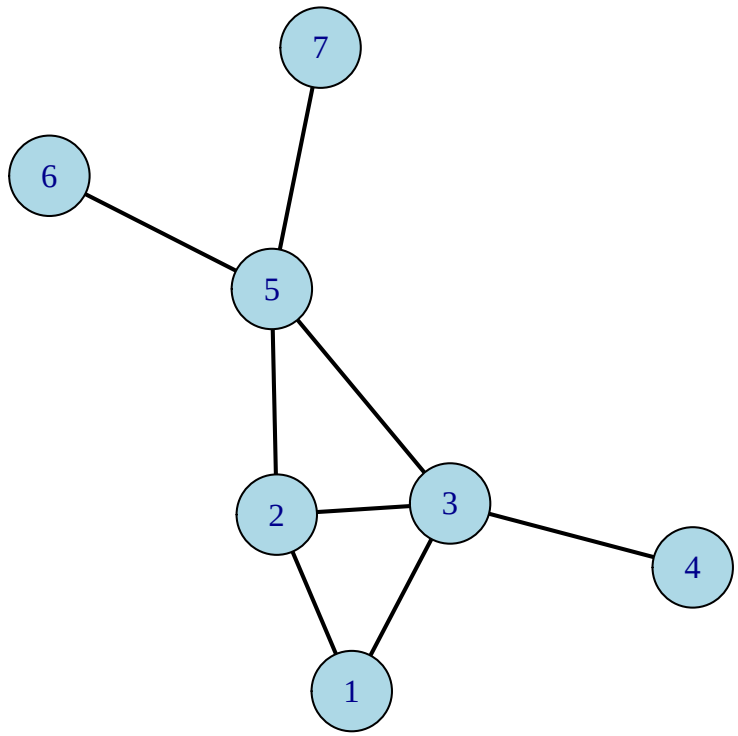
- A walk is a sequence of nodes and lines in which each node is adjacent with the nodes following and preceding it in the sequence
- A trail is a walk in which all links are distinct, but nodes can be repeated
- A path is a walk in which all nodes and links are distinct



$Path = (n_1, n_2, n_3, n_4)$

Key definitions
Walks, trails, paths

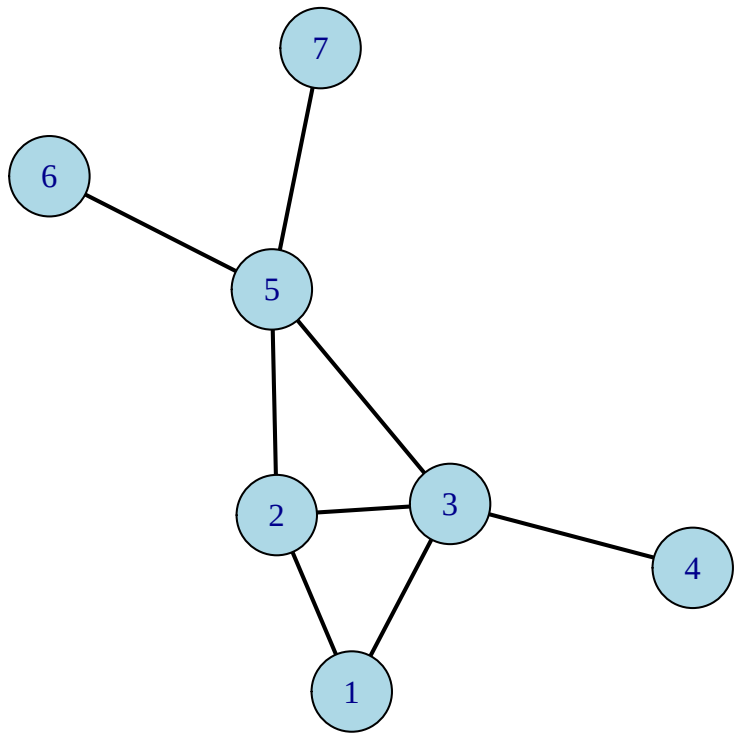
- A closed walk is a walk that begins and ends with the same node



$ClosedWalk = (n_1, n_3, n_5, n_2, n_1)$

Key definitions
Walks, trails, paths

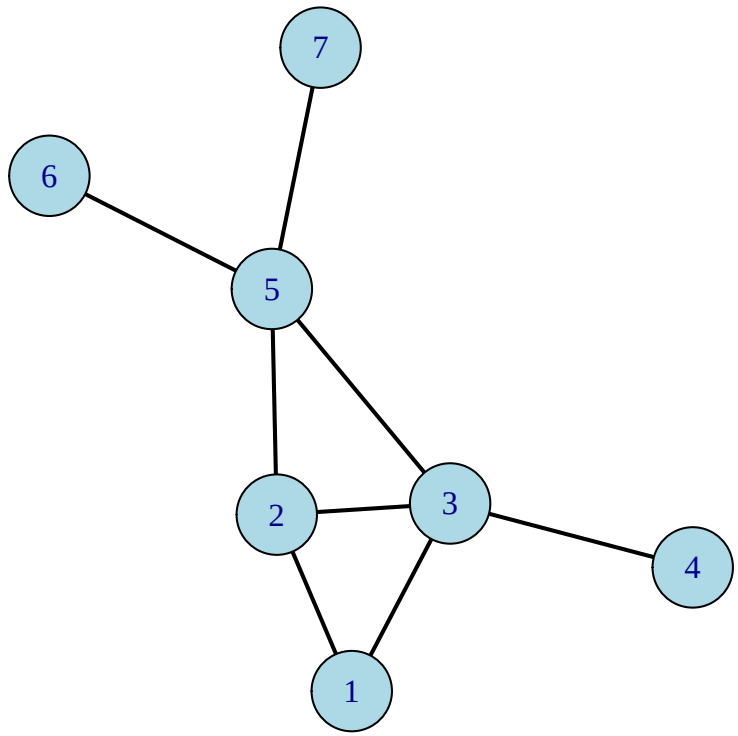
- A closed walk is a walk that begins and ends with the same node
- A tour is a walk in which each link is used at least once



$Tour = (n_1, n_2, n_5, n_7, n_5, n_6, n_5, n_3, n_4, n_3, n_3)$

Key definitions
Walks, trails, paths

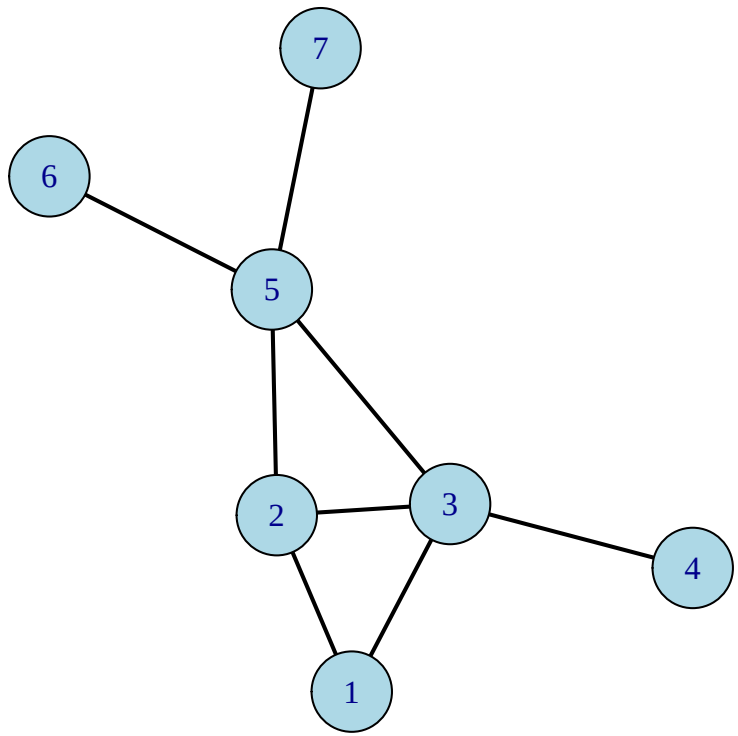
- A closed walk is a walk that begins and ends with the same node
- A tour is a walk in which each link is used at least once
- A cycle is a closed walk of at least three nodes in which links are distinct



$Cycle = (n_1, n_3, n_5, n_2, n_1)$

Key definitions
Geodesic distance

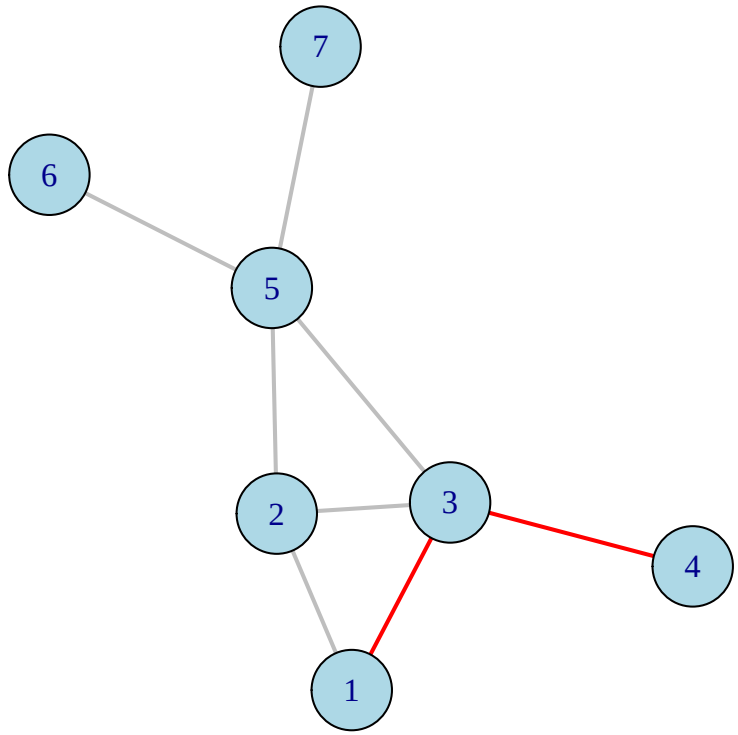
- The shortest path between two nodes in the network is called geodesic distance $d(n_i, n_j)$
- If a network is not connected, the geodesic distance is infinite in at least one pair of nodes



$d(n_i, n_j)$

Key definitions
Geodesic distance

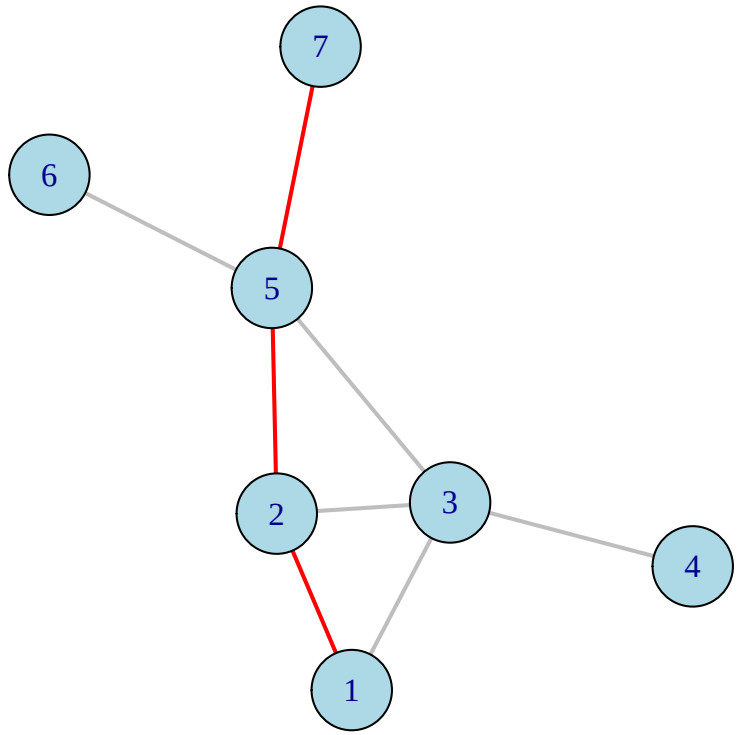
- The shortest path between two nodes in the network is called geodesic distance $d(n_i, n_j)$
- If a network is not connected, the geodesic distance is infinite in at least one pair of nodes



$d(1, 4) = 2$

Key definitions
Geodesic distance

- The shortest path between two nodes in the network is called geodesic distance $d(n_i, n_j)$
- If a network is not connected, the geodesic distance is infinite in at least one pair of nodes

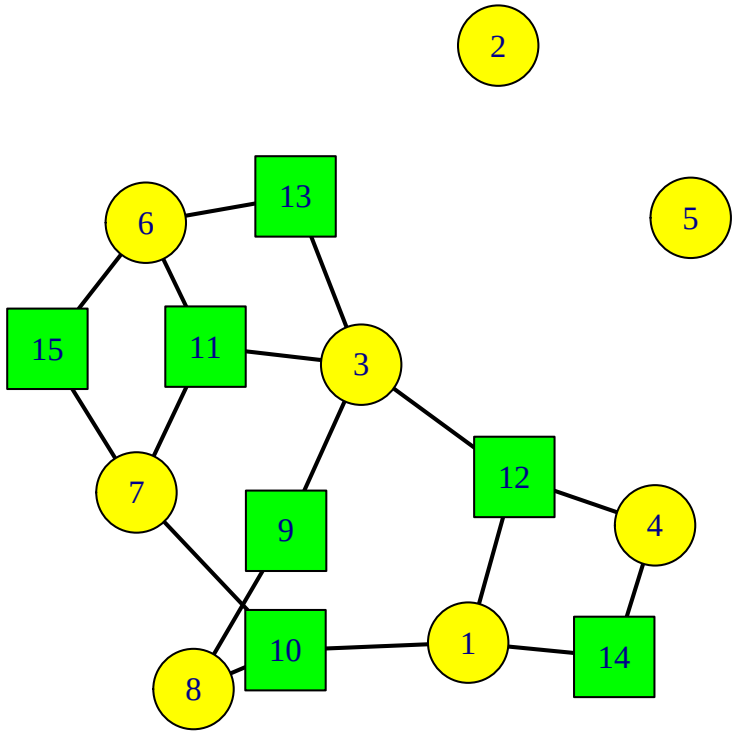


$d(1, 7) = 3$

Key definitions

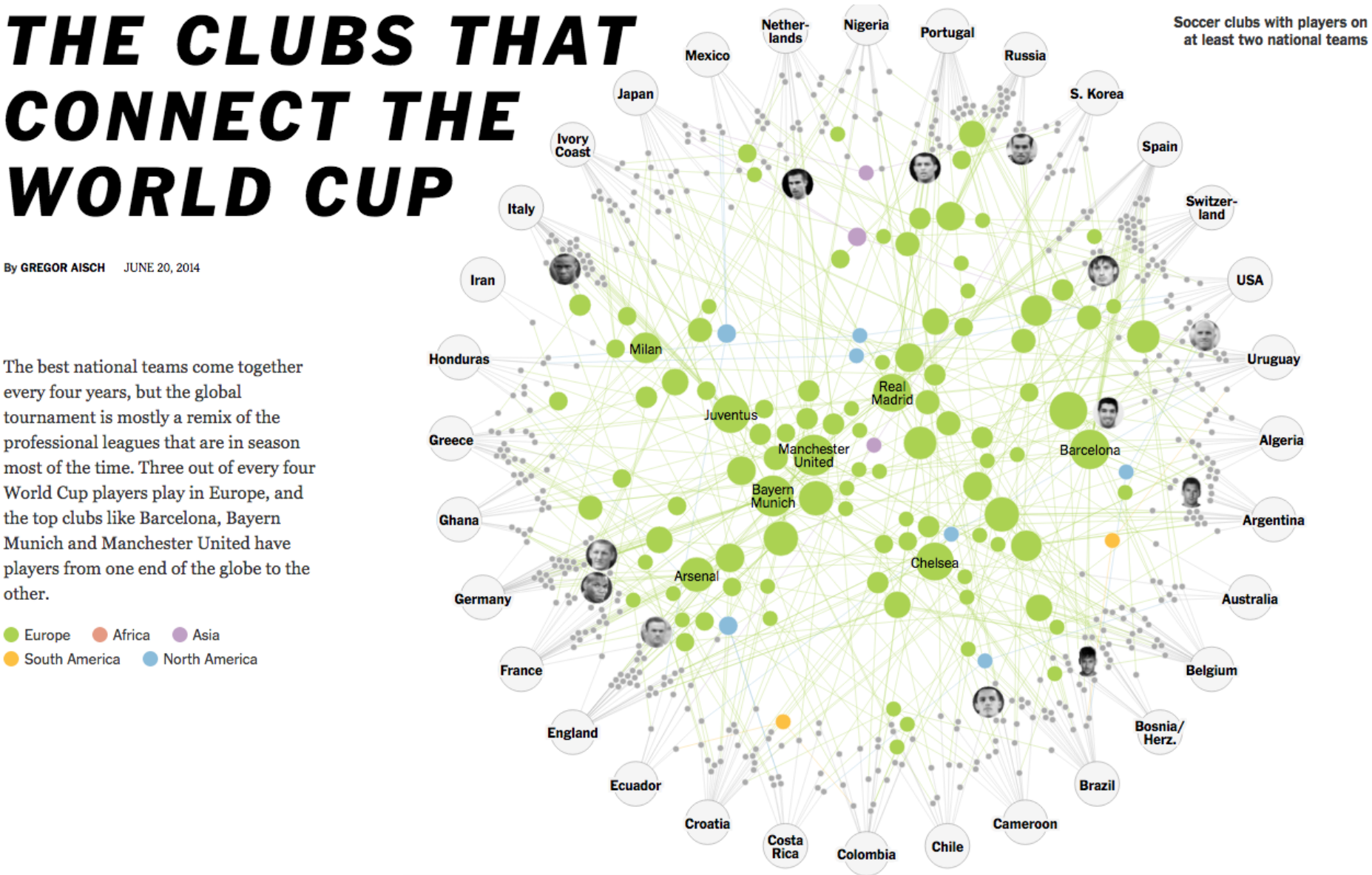
Bipartite graphs/two-mode networks

- A bipartite graph or two-mode network is a network where two subsets of nodes (N_1, N_2) exist and the links are only between nodes of distinct subsets
- Examples: author-affiliation, actor-movie, firm-patent, R&D projects and partners, etc.
- More than two partitions of nodes can exist: s-partite graph or s-mode network



Key definitions

Multigraphs/multiplex networks

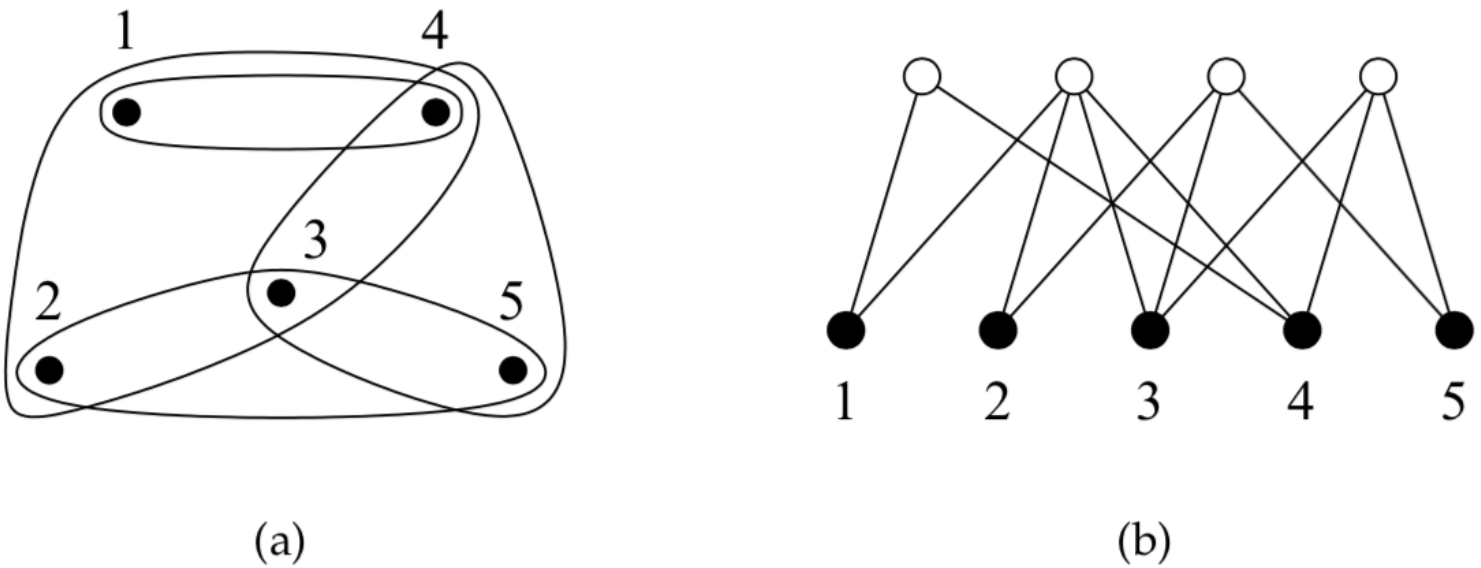


Source: `www.nytimes.com/interactive/2014/06/20/sports/worldcup/how-world-cup-players-are-connected.html`

Key definitions

Bipartite graphs vs. Hypergraphs

- In some networks, links can join more than two vertices (e.g. more than two authors on a publication)
- We can represent these with hypergraphs, where circles around groups of vertices are called hyperedges

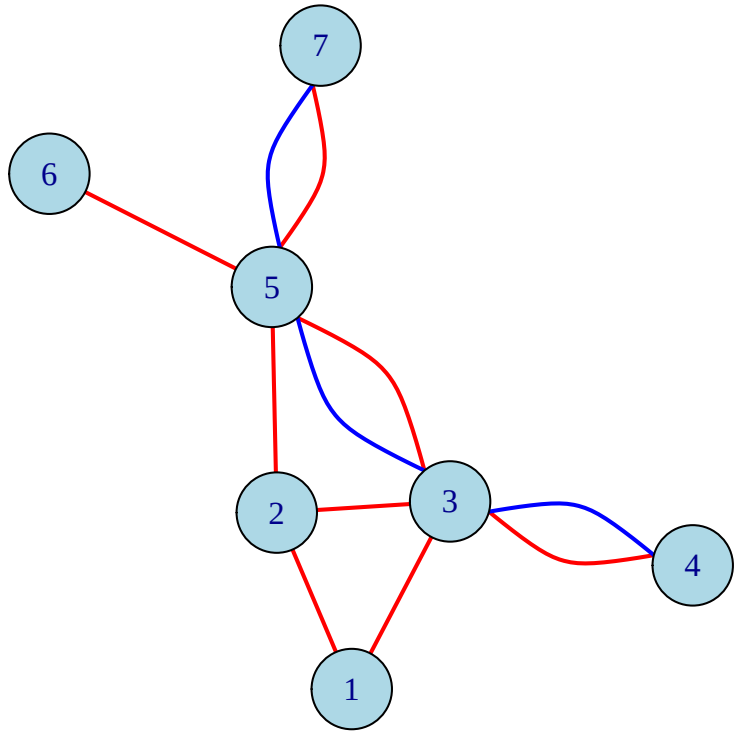


Source: [Newman, 2010]

Key definitions

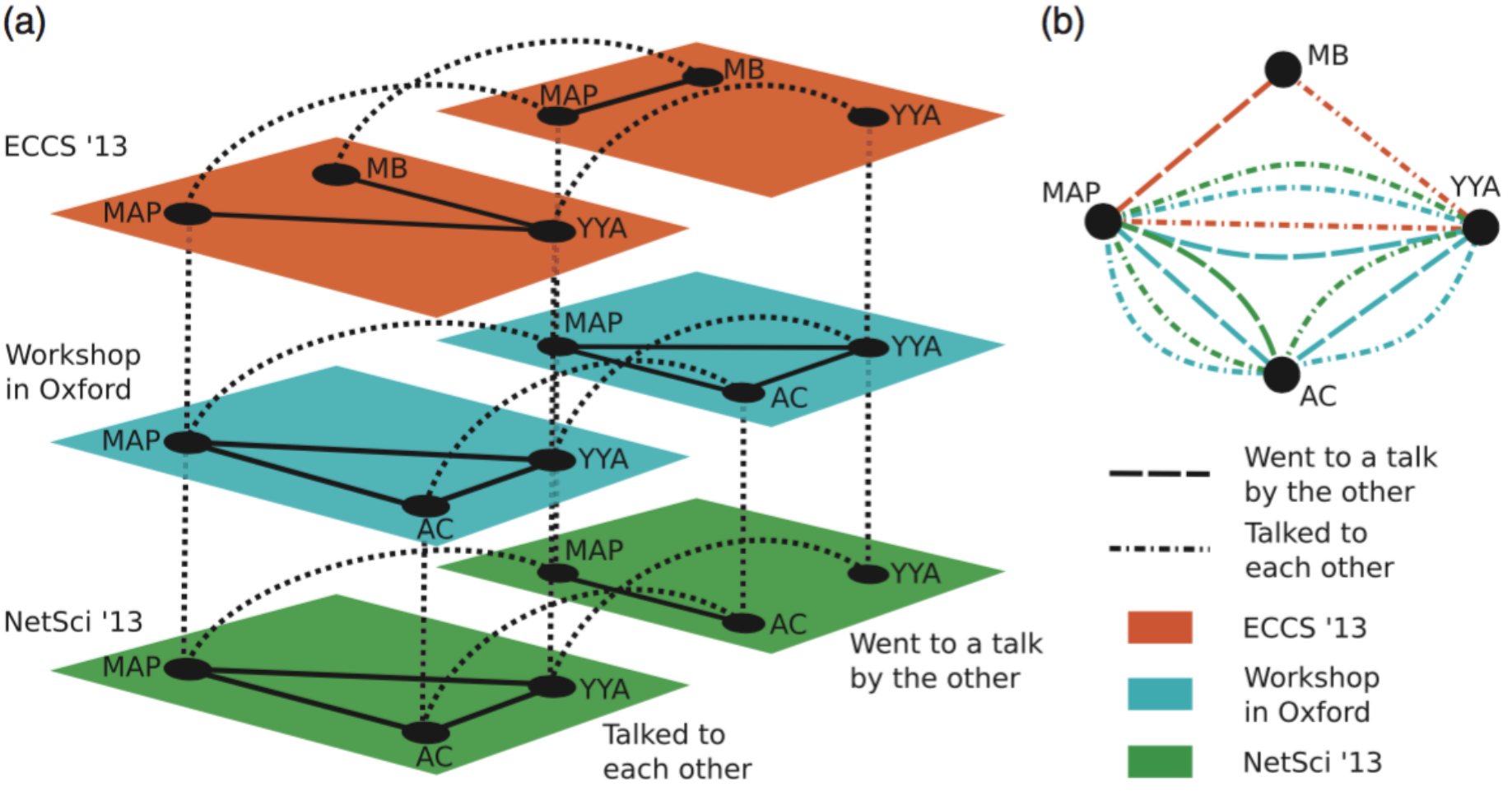
Multigraphs/multiplex networks

- A multigraph or multiplex network is a network where nodes are linked by different sets of edges $(E_1, E_2, ...E_R)$
- Examples: "who do you seek advice on job" vs. "who is your friend"



Key definitions

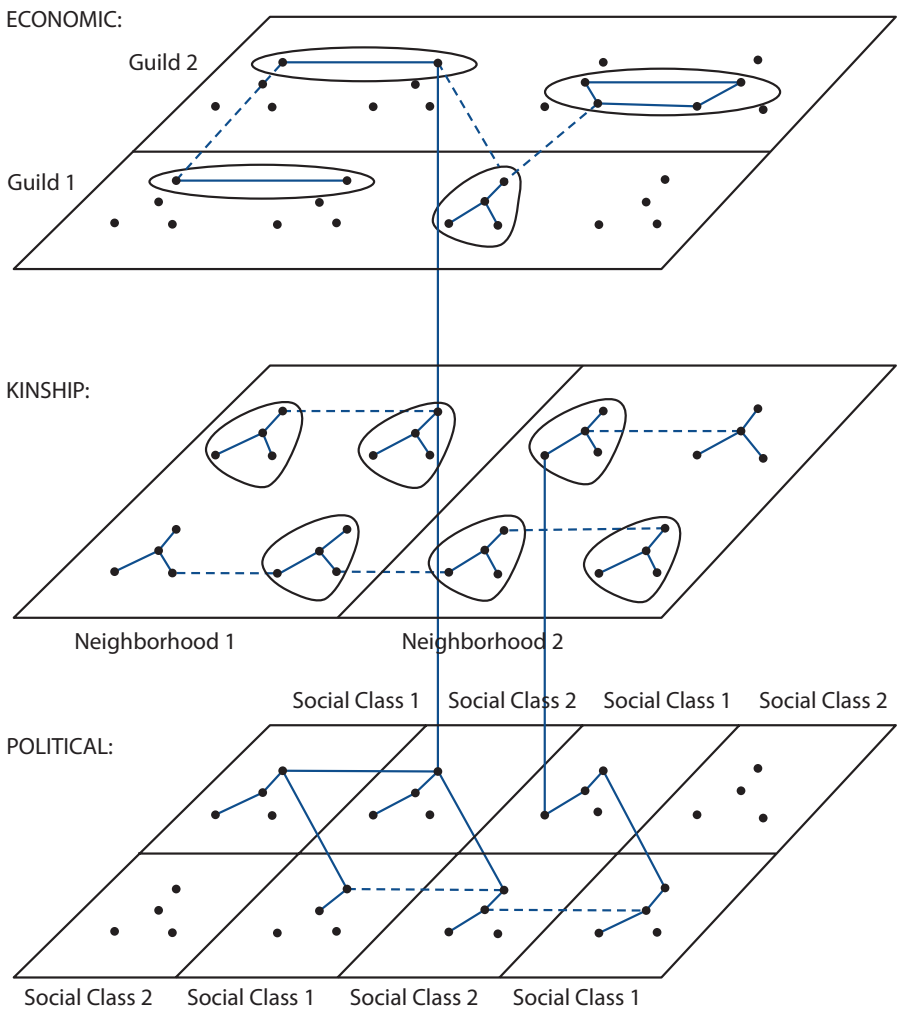
Multigraphs/multiplex networks



Source: [Kivelä et al., 2014]

Key definitions

Multigraphs/multiplex networks

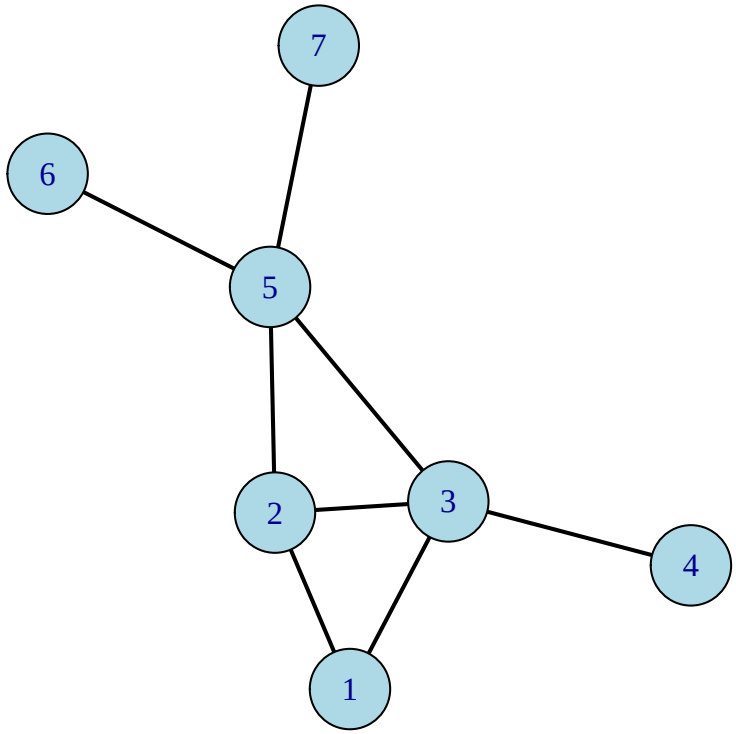


Source: Renaissance Florence [Powell and Sandholtz, 2012]

Key definitions

Ego-network

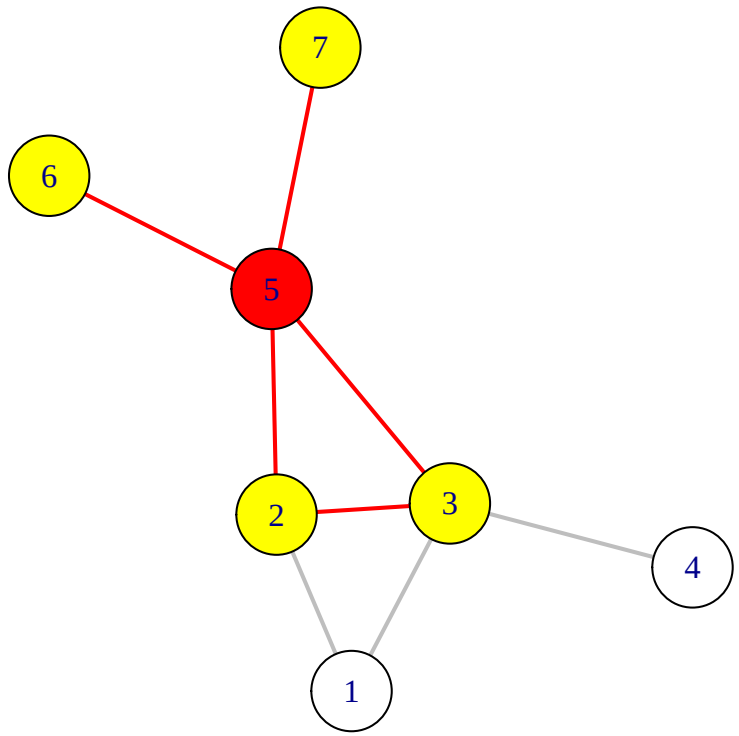
- The ego-network of a node i is the set of nodes that are directly ($d(i,j) = 1$) linked to i (ego) and all the lines among these nodes



Key definitions

Ego-network

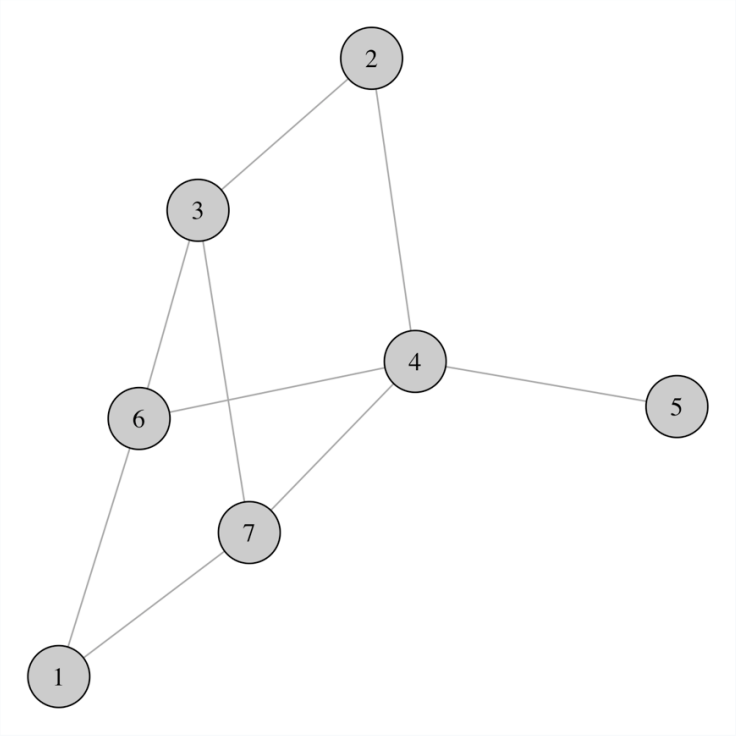
- The ego-network of a node i is the set of nodes that are directly ($d(i,j) = 1$) linked to i (ego) and all the lines among these nodes



Key definitions

When poll is active, respond at **PollEv.com/rotolo**

2. Is the sequence of nodes "1-7-4-5" a shortest path



Yes

No

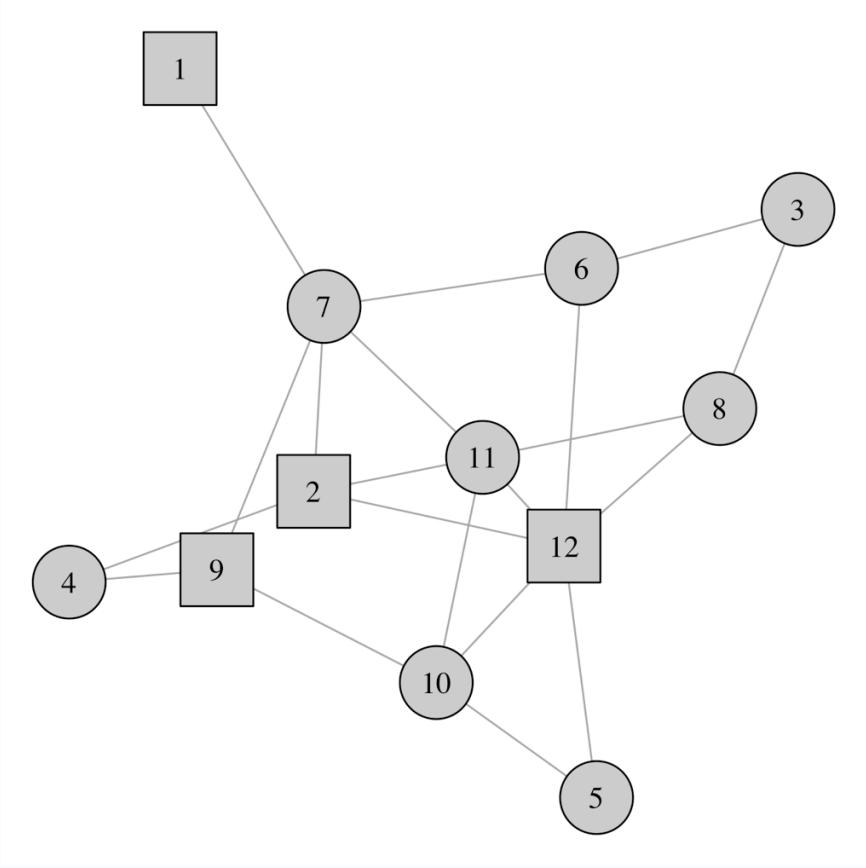
I do not know

Total Results: 0

Key definitions

When poll is active, respond at **PollEv.com/rotolo**

3. Is this network bipartite?



Yes

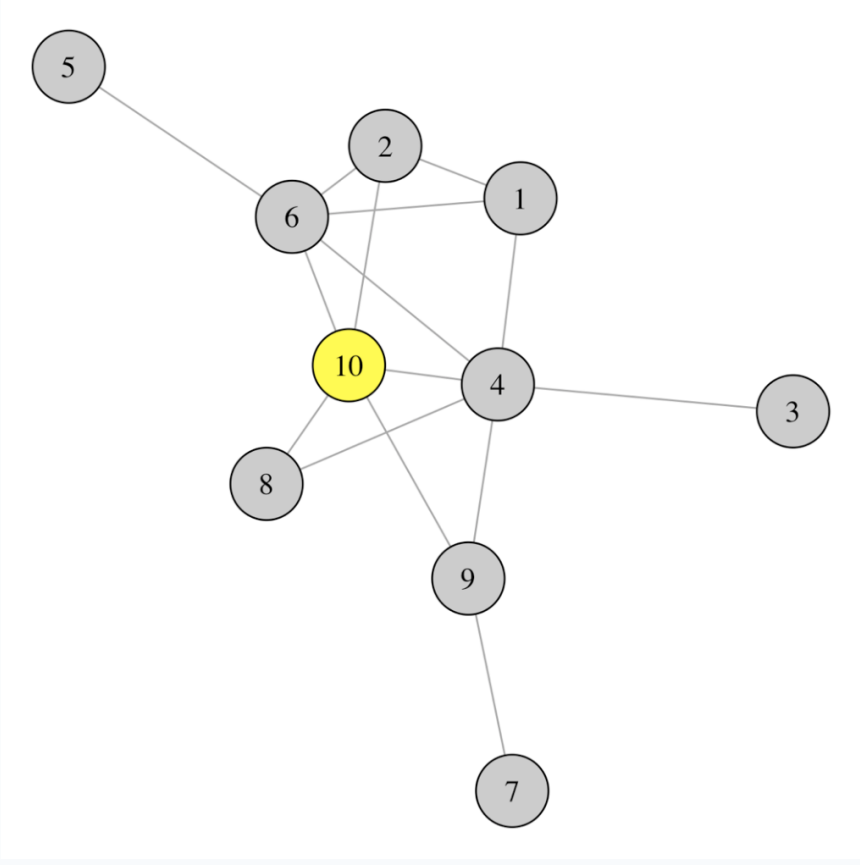
No

I do not know

Total Results: 0

When poll is active, respond at **PollEv.com/rotolo**

4. What is the ego-network of the node 10?



- The node 10 and the nodes [2, 4, 6, 9, 8]
- The nodes [2, 4, 6, 9, 8] and all the links between these nodes
- The node 10, the nodes [2, 4, 6, 9, 8], and all the links between these nodes
- It cannot be defined

Total Results: 0

Next time ...

Next time ...

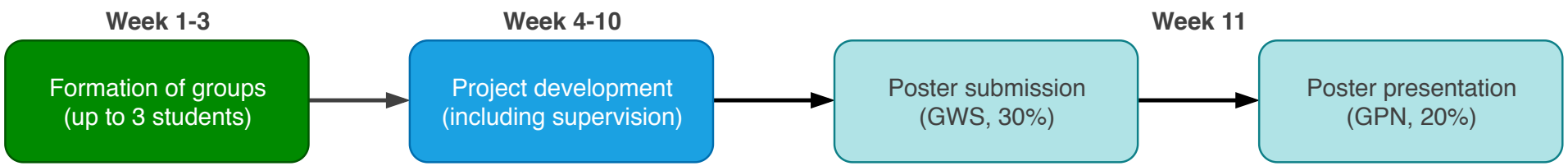
- **Seminar: Network definition**
  - Short intro to R
  - Basic commands of the igraph package

- **Lecture: Network data collection**
  - Main approaches to collect and sample network data
  - Network boundary specification problem

Assessment modes

Coursework

- Groups up to 3 students
- Report your group by Week 3: https://bit.ly/2KLKCJ6
- Small-scale network analysis project using novel or existing network data
- Any topic!
- Assessment modes:
  - **Group Written Submission (GWS)** [30% weighting]
    Infographic poster (A1-size, PDF)
  - **Group Presentation (GPN)** [20% weighting]
    10-minute video recording presenting the infographic poster (GWS)
- Submission in Week 11 (see Canvas)
- Marking criteria (see Canvas)



Questions

References I

Barabási, A.-L. (2016).

Network science.
Cambridge university press, Cambridge, MA.

Bondy, J. A. and Murty, U. S. R. (1976).

Graph theory with applications, volume 290.
North-Holland, New York, NY.

Borgatti, S. P. and Halgin, D. S. (2011).

On Network Theory.
Organization Science, 2(1):71–87.

de Nooy, W., Mrvar, A., and Batagelj, V. (2005).

Exploratory Social Network Analysis with Pajek, volume 53.
Cambridge University Press, Cambridge, UK.

Giuliani, E. (2013).

Network dynamics in regional clusters: Evidence from Chile.
Research Policy, 42(8):1406–1419.

Granovetter, M. S. (1973).

The strength of weak ties.
American Journal of Sociology, 78(6):1360–1380.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014).

Multilayer networks.
Journal of Complex Networks, 2(3):203–271.

Klavans, R. and Boyack, K. W. (2014).

Mapping altruism.
Journal of Informetrics, 8(2):431–447.

References II

Moreno, J. L. (1934).

Who Shall Survive? A New Approach to the Problem of Human Interrelations, volume 58.
Nervous and Mental Disease Publishing, Washington, DC, US.

Newman, M. (2010).

Networks. An introduction.
Oxfor University Press, Oxford, UK.

Powell, W. W. and Sandholtz, K. W. (2012).

Amphibious entrepreneurs and the emergence of organizational forms.
Strategic Entrepreneurship Journal, 6(2):94–115.

Scott, J. (2000).

Social network analysis: A handbook.
SAGE Publications, London.

Wasserman, S. and Faust, K. (1994).

Social Network Analysis: Methods and Applications., volume 24.
Cambridge University Press, New York, NY.