



# Social Networks

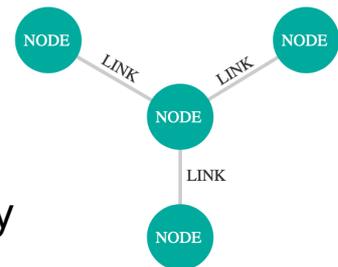
Algorithms, Data and Security  
A.Y. 2023/24

**Valeria Cardellini**

Global Governance, 3rd year  
Science and Technology Major

## Networks

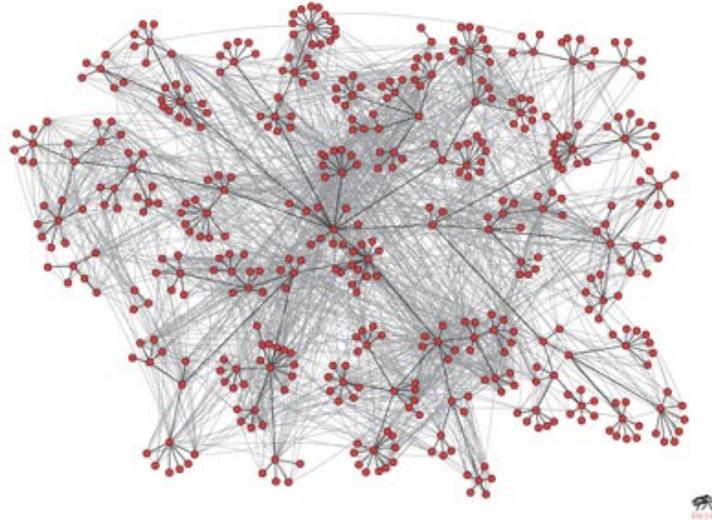
- Networks are a representation of **interaction** structure among units
  - These units are **nodes**
  - Some pairs of nodes are connected by **links**
  - In case of social and economic networks, nodes are individuals or organizations
- The study of networks can encompass the study of all kinds of interactions, e.g.,
  - Transportation
  - Communication
  - **Social**
  - Spread of epidemics



## Examples of social networks

---

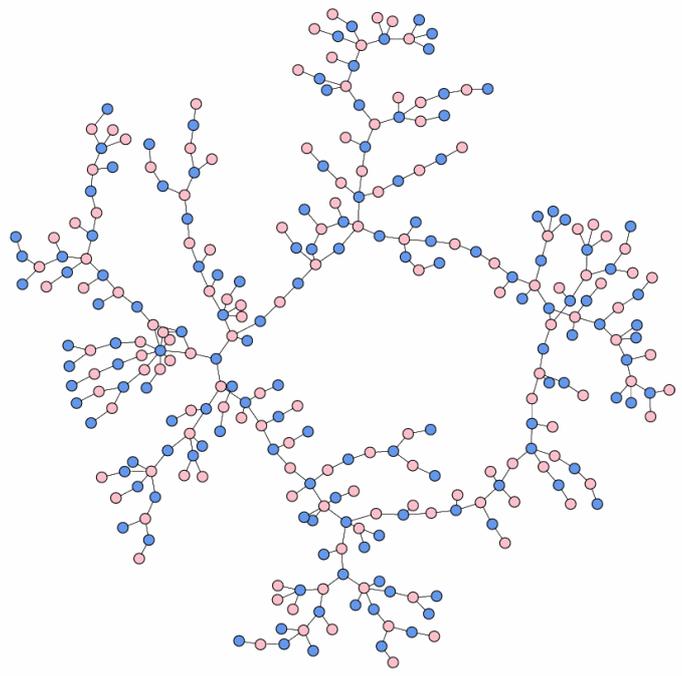
- Corporate e-mail communication
  - Links indicate people connected by e-mail exchange



## Examples of social networks

---

- High school dating



## Examples of social networks

---

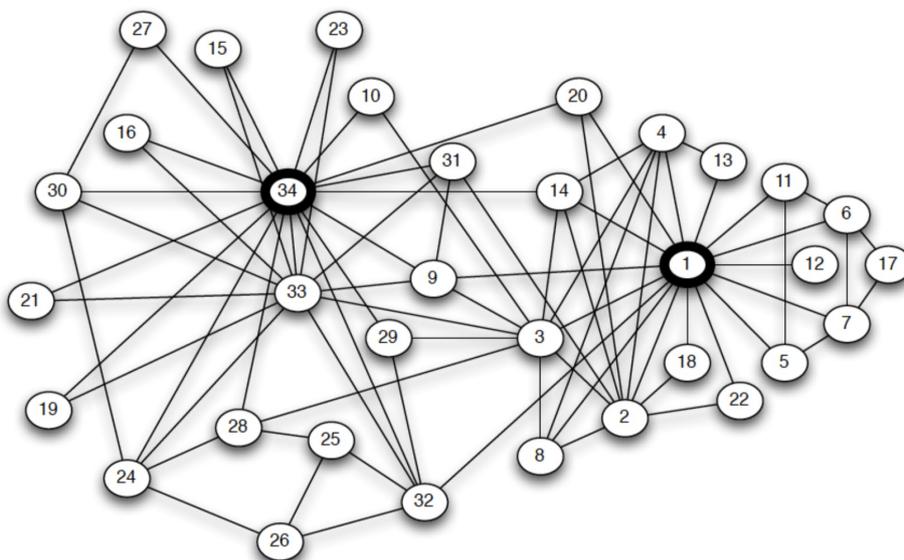
- Trails of Flickr users in Manhattan



## Examples of social networks

---

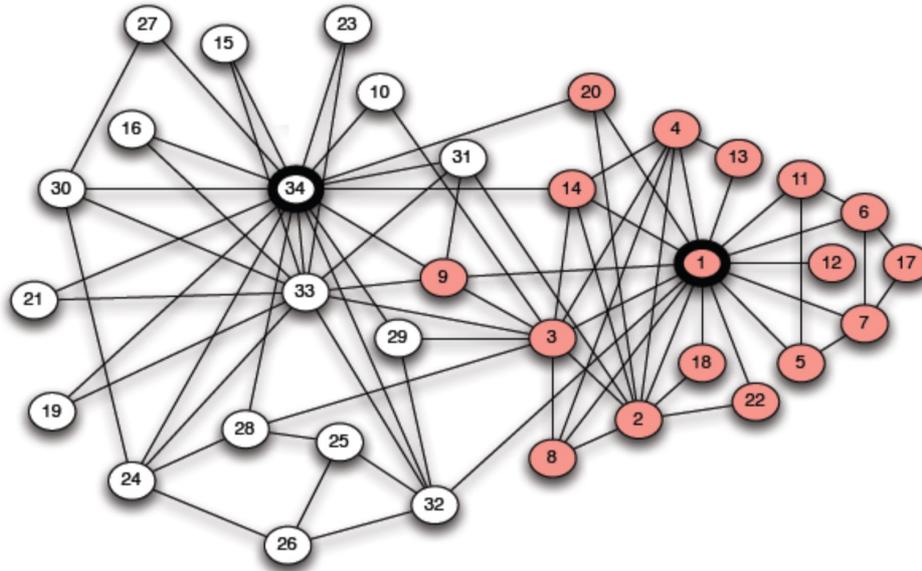
- The social network of friendships within a 34-person karate club



## Examples of social networks

---

- The social network of friendships within a 34-person karate club provides clues to the fault lines that eventually split the club apart



Valeria Cardellini - ADS 2023/24

6

## Networks: behavior and dynamics

---

- Networks are complex and highly connected systems
- Connectedness at the level of **structure**: who is linked to whom (**graph theory**)
  - Structural features of networks are important first step to understand them, but not the only one
- Connectedness at the level of **behavior**: each individual's actions have implicit consequences for the outcomes of everyone in the system (**game theory**)

Valeria Cardellini - ADS 2023/24

7

## Networks: behavior and dynamics

---

- Models of networked behavior must take **strategic** behavior and reasoning into account
  - Evaluate your actions not in isolation, but with the expectation that the world will react to what you do
- Investigate **dynamics** of aggregate behavior

## Network dynamics: population effects

---

- Observing a large population over time, we see a recurring pattern by which new ideas, beliefs, opinions, innovations, technologies, products, and social conventions emerge and evolve
- Collectively, we can refer to these as **social practices** (holding opinions, adopting products, behaving according to certain principles) that people can choose to adopt or not
- The way in which new practices spread through a population depends in large part on the fact that people **influence** each other's behavior

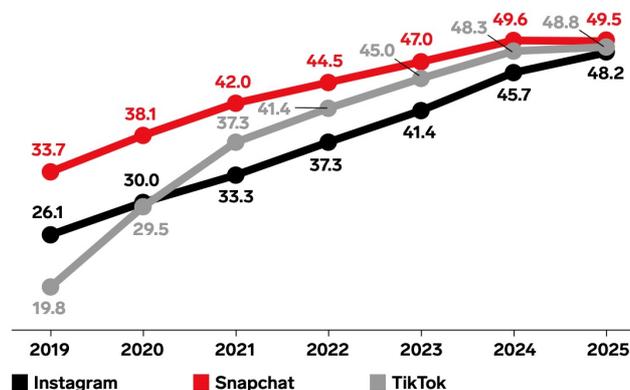
## Network dynamics: population effects

- At a surface level, we could hypothesize that people imitate the decisions of others because of the human tendency to **conform**
- But there are multiple reasons why even purely rational agents - individuals with no a priori desire to conform to what others are doing – will copy the behavior of others
- One class of reasons is based on the fact that the behavior of others conveys information: **information cascades**

## Network dynamics: population effects

- Example: adoption of particular practices over time

**US Gen Z Instagram, Snapchat, and TikTok Users, 2019-2025**  
millions



Note: individuals born between 1997-2012 who access their account via any device at least once per month  
Source: eMarketer, May 2021

## Network dynamics: population effects

---

- **Cascading adoption** of a new technology or service can be the result of individual incentives to use the most widespread technology
  - Either based on the informational effects of seeing many other people adopt the technology
  - Or the direct benefits of adopting what many others are already using

## Network dynamics: structural effects

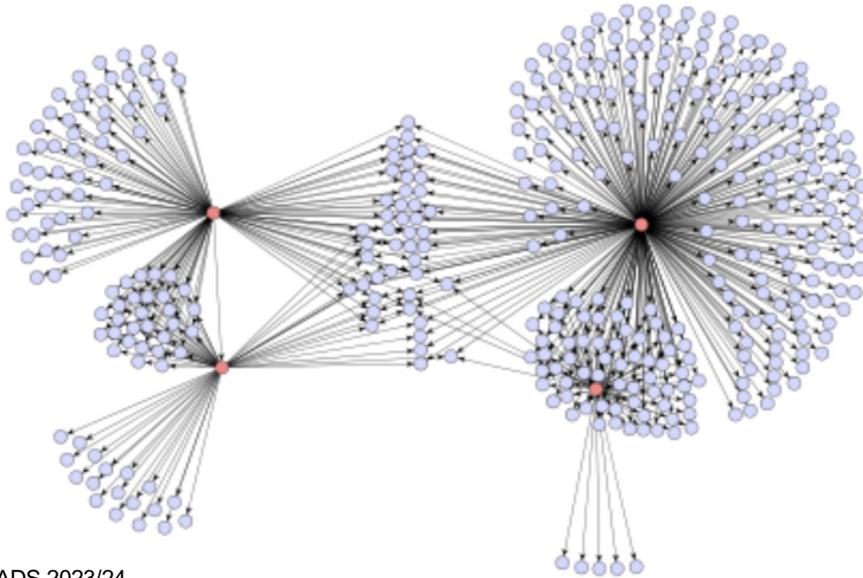
---

- Taking network structure into account provides important further insights into how such kinds of influence take place
- The underlying mechanisms - based on information and direct benefits - are present both at the global level of **whole population**, and also at a **local level** in the network
  - Between an individual and their set of friends or colleagues
- When people are influenced by the behaviors of their **neighbors** in the network, the adoption of a new product or innovation can **cascade** through the network structure

## Network dynamics: structural effects

---

- Example: e-mail recommendations for a Japanese graphic novel spread in a kind of informational or social contagion



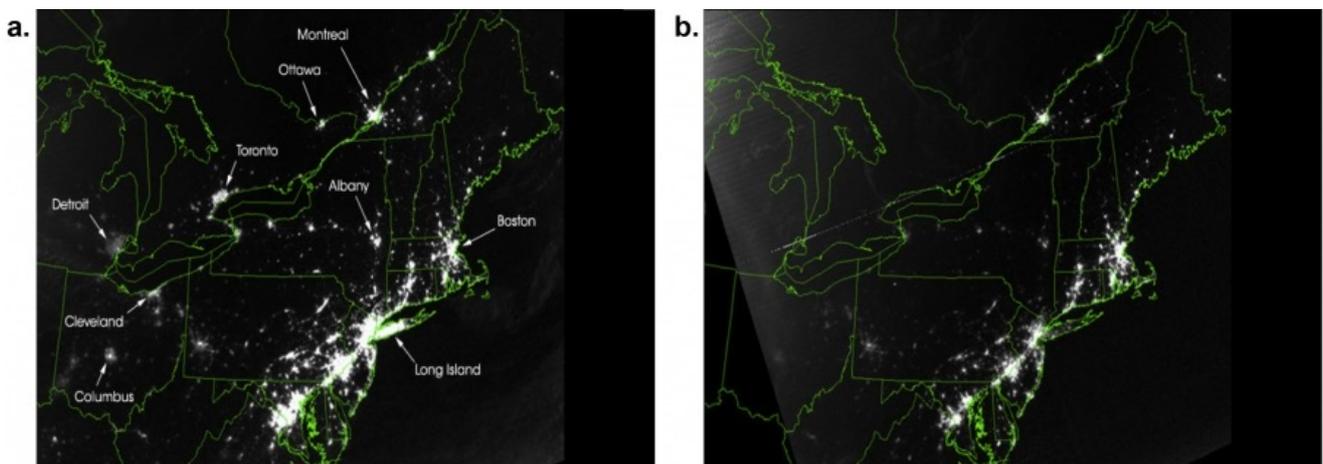
Valeria Cardellini - ADS 2023/24

14

## Network dynamics: structural effects

---

- Example: 2003 North American blackout



a) Satellite image on Northeast United States on August 13th, 2003, at 9:29pm (EDT), 20 hours before the 2003 blackout

b) The same as above, but 5 hours after the blackout

Valeria Cardellini - ADS 2023/24

15

## Vulnerability due to interconnectivity

---

- 2003 blackout: typical example of **cascading failure**, where a local failure may not stay local any longer
  - When a network acts as a transportation system, a local failure shifts loads to other nodes
  - If the extra load is negligible, the system can seamlessly absorb it and the failure goes unnoticed
  - If the extra load is too much for the neighboring nodes, they will redistribute the load to their neighbors
  - In no time, we are faced with a cascading event, whose magnitude depends on the position and the capacity of the nodes that failed initially

## Networks: interdisciplinarity

---

- Theory of network structure and behavior addresses simultaneous challenges deriving from
  - **Economics**: theories for strategic interaction among small numbers of parties, as well as for cumulative behavior of large, homogeneous populations
  - **Sociology**: some of fundamental insights into structure of social networks, but network methodology refined only in domains and scales where data-collection has traditionally been possible (well-defined groups with tens to hundreds of people)
  - **Computer Science**: with rise of Web and social media, we deal with design constraints on large computing systems which are not only technological but also human (complex feedback that human audiences create when humans collectively use Web for communication, self-expression, and creation of knowledge)

## Network science

---

- Academic field which studies **complex networks**
  - Including telecommunication, computer, biological, cognitive and semantic, social networks
- Considers distinct elements or actors represented by **nodes** (or **vertices**) and the connections between elements or actors as **links** (or **edges**)
- Theories and methods from multiple disciplines
  - Graph theory from mathematics, statistical mechanics from physics, data mining and information visualization from computer science, inferential modeling from statistics, and social structure from sociology

## Elements of any social network

---

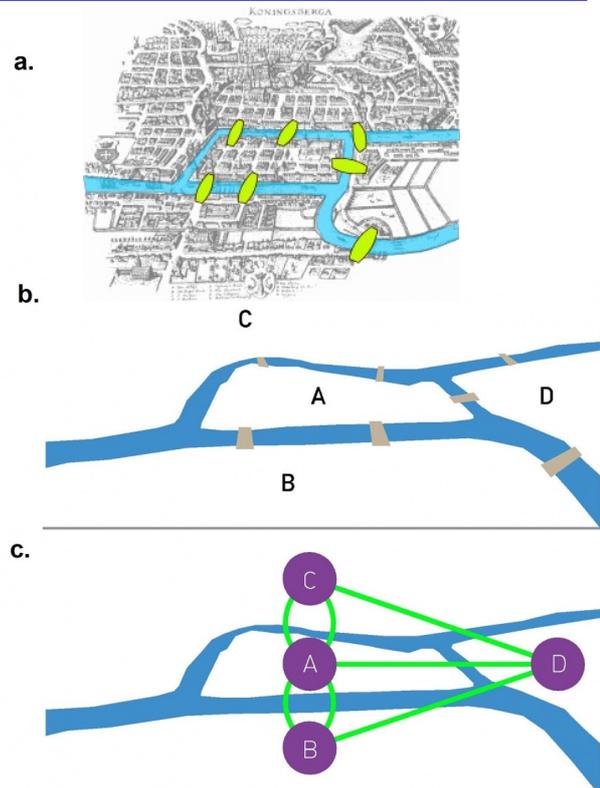
- We focus on social networks
- Components of networks consist of actors and relations
- Actors can be individuals (e.g., students, employees) or organizations
  - Those actors are modeled as **nodes** in a graph
- The connections between actors make up relations (modeled as **links** or **edges**)
  - Those relations can represent friendship, advice, hindrance or communication

# A brief history of networks (graphs)

- The earliest known paper in this field is the famous Seven Bridges of Königsberg written by Leonhard Euler in 1735
- Laid the foundations of graph theory
- City of Königsberg with 7 bridges across the river Pregel that surrounded the town
- Can one walk across all 7 bridges and never cross the same one twice?

## Seven bridges of Königsberg

- Can one walk across all 7 bridges and never cross the same one twice?
- Euler constructed a **graph** that has 4 **nodes** (A, B, C, D), each corresponding to a patch of land, and 7 **links**, each corresponding to a bridge
- He then showed that there is **no continuous path** that would cross the 7 bridges while never crossing the same bridge twice

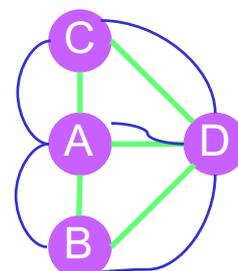
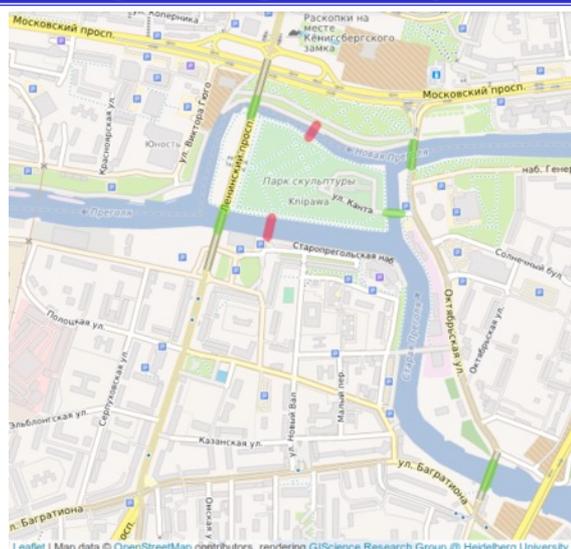


# Seven bridges of Königsberg

- Euler showed that the possibility of a walk through a graph, traversing each link exactly once, depends on the degrees of the nodes
  - **Node degree**: number of links touching it
- Necessary condition: the graph is connected and has exactly 0 or 2 nodes of odd degree
  - Not satisfied by Königsberg graph: 4 nodes with odd number of links (A, B, C, and D)
- Theorem (Euler): A **graph** is **Eulerian** if and only if:
  - All nodes have even degree
  - Or only 2 nodes have odd degree

## Seven Bridges of Königsberg: today

- 2 bridges were destroyed in World War II, while 2 others were later demolished and replaced, the 3 other bridges remain
- Now 2 nodes (B and C) have degree 2, and the other 2 have degree 3
- An Eulerian path is now possible, but it must begin on one island (A or D) and end on the other (D or A)
  - E.g., A -> B -> D -> C -> A -> D



# Networks or graphs?

---

- In scientific literature the terms network and graph are used interchangeably:

## Network Science

Network

Node

Link

## Graph Theory

Graph

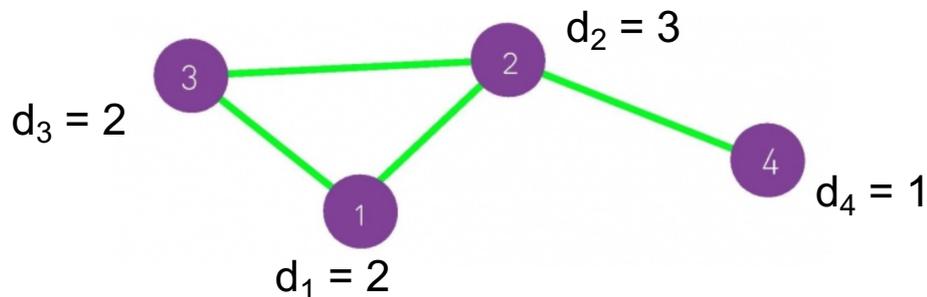
Vertex

Edge

# Graph theory: degree

---

- **Node degree**: number of links the node has to other nodes

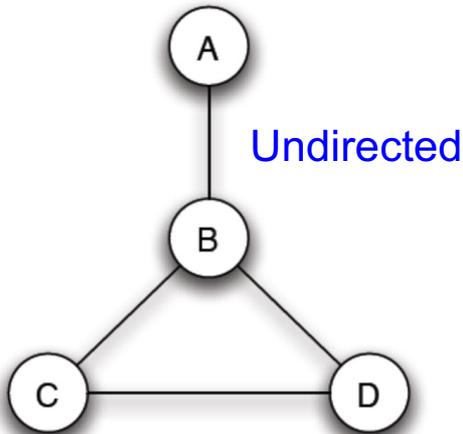


- **Average degree**: average node degree of all the nodes in the graph
  - In the example:  $(2+3+2+1)/4 = 2$

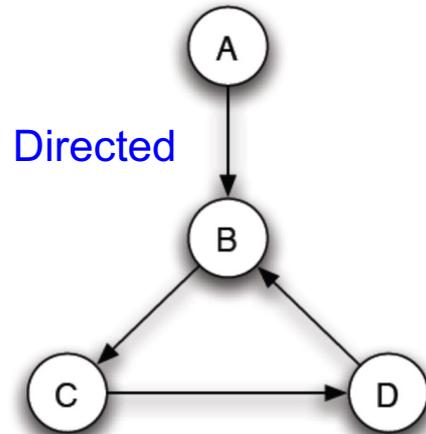
## Graph theory: undirected or directed

---

- **Undirected** or **directed** graph
  - Links can be directed or undirected
  - A graph is called directed if all of its links are directed; it is called undirected if all of its links are undirected



Valeria Cardellini - ADS 2023/24

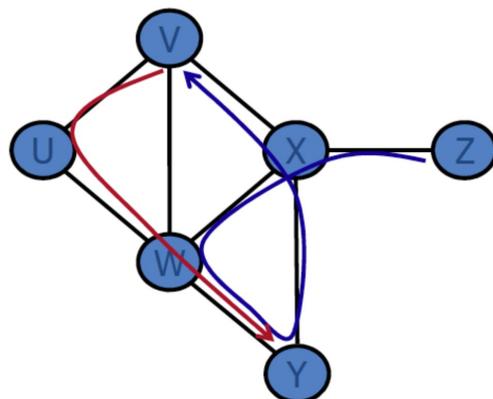


26

## Graph theory: path

---

- **Path**: a sequence of nodes with each consecutive pair in the sequence connected by a link
  - Red path:  $V \rightarrow U \rightarrow W \rightarrow Y$
- **Length** of the path: number of links the path contains
  - Red path has length 3
  - Blue path has length 5



Valeria Cardellini - ADS 2023/24

27

## Graph theory: shortest path

- Several algorithms exist to find shortest paths in a graph (i.e., paths with the fewest number of links)
  - Among which Dijkstra algorithm
- Used by Google Maps to find the shortest route
  - Shortest duration routes depending upon current traffic conditions



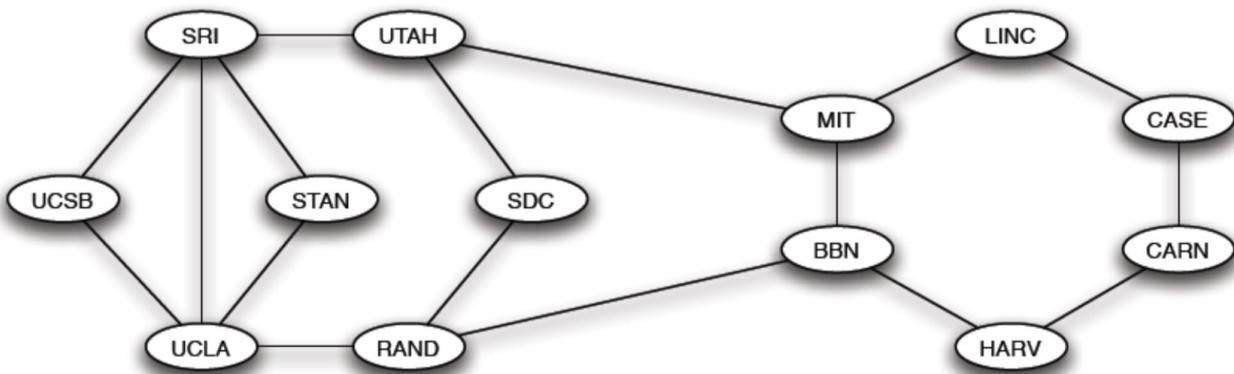
Valeria Cardellini - ADS 2023/24



28

## Graph theory: cycle

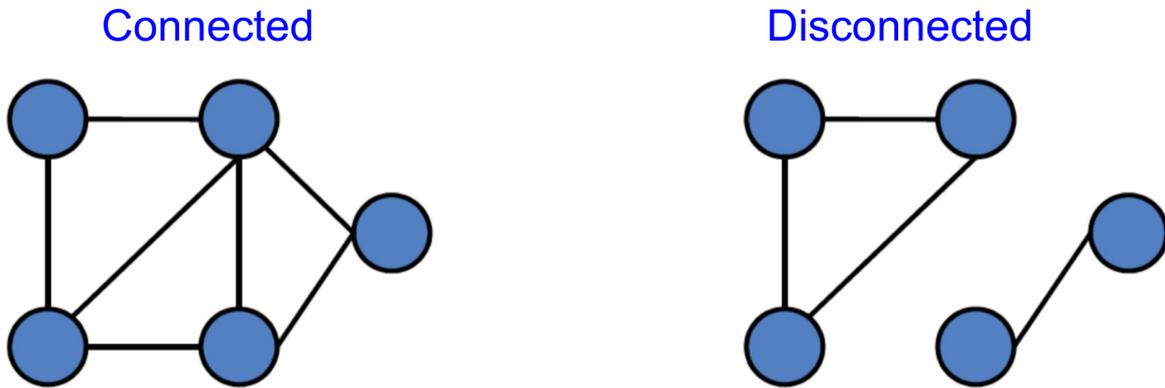
- **Cycle or loop:** a path in which the first and last nodes are the same, but otherwise all nodes are distinct



## Graph theory: connectivity

---

- An **undirected** graph is **connected** if, for every pair of nodes, there is a path between them

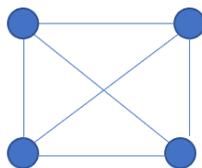


## Graph theory: connectivity

---

- In a **fully connected undirected** graph with  $N$  nodes, the number of links is equal to

$$\frac{N \times (N - 1)}{2}$$



Example:  $N=4$

Number of links is  $\frac{4 \times 3}{2} = 6$

# Social network analysis

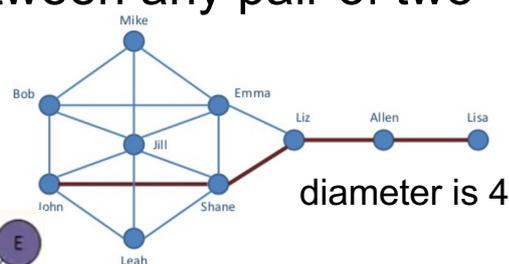
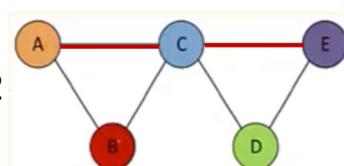
- **Social network analysis (SNA)** represents how individuals connect to each other and shows the flow of information in a network
- Can help identify influencers, key players, disconnected individuals and communities
- Consists in models that define relationships and interactions between individuals, as well as patterns that emerge from those relationships and interactions
- Importance of SNA in different fields
  - Sociology, business, engineering, economics, politics, computer science, mathematics, physics, biology, security

## SNA: network measures

- Let's analyze some measures of the network that allow us to characterize it
- Let's first consider **network size**
  - Diameter
  - Average shortest path
- **Diameter**: length of the longest shortest path through the network between any pair of two nodes

- Diameter marked in red

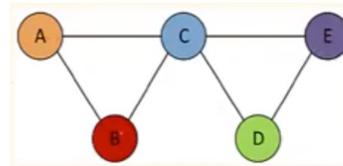
diameter is 2



diameter is 4

## SNA: network measures

- **Average shortest path length:** shortest path length (SPL) between two nodes averaged over all pairs of nodes
  - Not as extreme as diameter
  - Example:



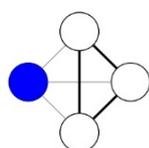
- SPL from A to B: 1
- SPL from A to C: 1
- SPL from A to D: 2
- SPL from A to E: 2
- SPL from B to C: 1
- SPL from B to D: 2
- SPL from B to E: 2
- SPL from C to D: 1
- SPL from C to E: 1
- SPL from D to E: 1

Average SPL is:

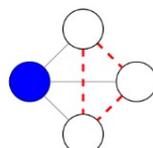
$$\frac{1 + 1 + 2 + 2 + 1 + 2 + 2 + 1 + 1 + 1}{10} = 1.4$$

## SNA: network measures

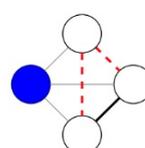
- **Clustering coefficient:** measure of the degree to which nodes in a network tend to cluster together; can be local or global
  - **Clustering coefficient of a node:** number of links between neighbors of that node divided by the maximum number of possible links between them (if node  $v$  has  $m_v$  neighbors,  $\max = m_v * (m_v - 1) / 2$ )
  - **Average clustering coefficient:** averaged over all nodes of the network
  - E.g., clustering coefficient of blue node



$$c = 1$$



$$c = 0$$



$$c = 1/3$$

3 neighbors

max number of links:  
 $3 * 2 / 2 = 3$

# How to build a social network?

---

- Network science aims to build models that reproduce properties of real networks
- Goal: how to build a **model for network formation**
- We consider 3 models
  - **Erdos-Renyi** model
  - **Small-world** model (or Watts-Strogatz)
  - **Scale-free** model (or Barabási-Albert)
- Properties of networks we consider
  - *Average clustering coefficient*
  - *Average shortest path length*

## Network models: Erdos-Renyi

---

- **Erdos-Renyi** is the classical **random** network: nodes are randomly connected to each other (i.e., choices independent of current network structure)
  - $N$ : number of nodes in the network
  - $p$ : probability that two nodes are connected by a link
  - Node degree follows **binomial distribution**

$P(\text{deg}(v) = k)$  : probability that node  $v$  has degree  $k$

$$P(\text{deg}(v) = k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

- Average clustering coefficient:  $p$
- Average shortest path length:  $\frac{\log(N)}{\log((N-1)p)}$

## Network models: Erdos-Renyi

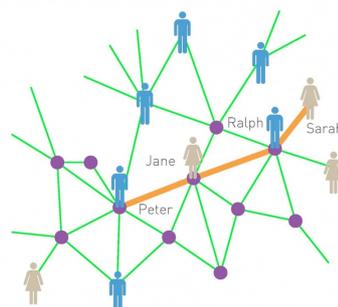
---

- Erdos-Renyi properties
  - Small average shortest path length and low clustering coefficient
  - Homogeneous network, i.e., all nodes are similar
  - But low clustering and homogeneity are not observed in many real-world networks!
- Properties observed in **real-world networks**:
  - **High clustering coefficient**
  - **Presence of hubs** (high-degree nodes), i.e., non-homogeneous network
  - **Small-world property**: average distance between two nodes depends logarithmically on  $N$

## Network models: Watts-Strogatz

---

- **Watts-Strogatz**: *small-world* random network
- *Small-world phenomenon*, also known as *six degrees of separation*, states that if you choose any two individuals anywhere on Earth, you will find a path of at most six acquaintances between them
  - The distance between any two nodes in a network is unexpectedly small



# Small world: experimental confirmation

- The first small-world study comprised several experiments conducted in 1967 by Stanley Milgram to measure the **average path length for social networks of people** in the US
  - Letters sent from randomly selected residents of Wichita (Kansas) and Omaha (Nebraska) to two target persons in Boston
  - Senders were asked to forward the letter to a friend, relative or acquaintance who is most likely to know the target person
- The research was groundbreaking in that it suggested that human society is a **small-world network characterized by short path lengths**
- The experiments are often associated with the phrase “six degrees of separation”, although Milgram did not use this term himself

# Small world: experimental confirmation



One possible path of a message in the “Small World” experiment by Stanley Milgram

[https://en.wikipedia.org/wiki/Small-world\\_experiment](https://en.wikipedia.org/wiki/Small-world_experiment)



## Small world: experiment repeated

---

- Global, Internet-based social search experiment
  - 18 target persons in 13 different countries
  - 60,000+ participants: by email pass the message to a social acquaintance whom they considered “closer” than themselves to the target
  - 24,163 message chains
  - 384 reached their targets
  - average path length 4.0
  - Actual success depends sensitively on individual incentives

Dodds, Muhamad, Watts, [An Experimental Study of Search in Global Social Networks](#), *Science*, 2003

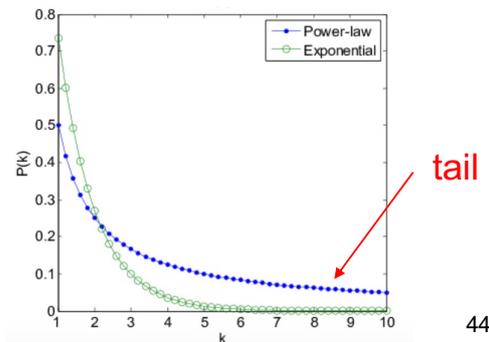
## Network models: Watts-Strogatz

---

- **Watts-Strogatz:** *small-world* random network
- Watts-Strogatz properties:
  - Small average shortest path length
  - High clustering coefficient, independent of the number of vertices
- Most nodes are not neighbors of one another, but most nodes can be reached from every other by a small number of hops
  - Example: social networks
- Limitation: does not account for the formation of hubs

# Network models: Barabási-Albert

- **Barabási-Albert**: a *scale-free* network
  - Node degree follows a **power law** distribution
$$P(\text{deg}(v) = k) \sim k^{-\alpha} \quad \text{where } 2 < \alpha < 3$$
- **Power law**: event frequency varies as a power of some attribute of that event
  - Examples: distribution of wealth among individuals, number of cities with a certain population size
  - A property of power law is **scale invariance**
  - Power law is also **heavy-tailed**
    - Tail of distribution contains a great deal of probability (i.e., heavier tail than that of exponential distribution)

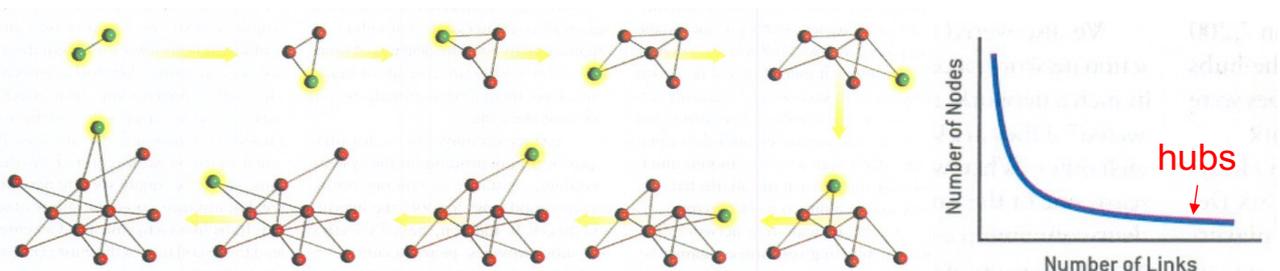


Valeria Cardellini - ADS 2023/24

44

# Network models: Barabási-Albert

- How do new nodes connect to existing nodes?  
On the basis of **preferential attachment**
  - The more connected a node is, the more likely it is to receive new links (i.e., rich-gets-richer process): thanks about connections in a social network!
  - As a result, there are a few **hubs** (nodes with a very high number of links), while most nodes have just a few links (power law behavior)



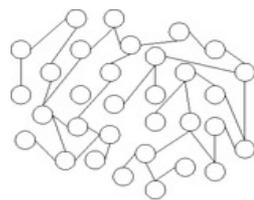
Barabasi and Bonabeau, [Scale Free-Networks](#), JSTOR 2003.

Valeria Cardellini - ADS 2023/24

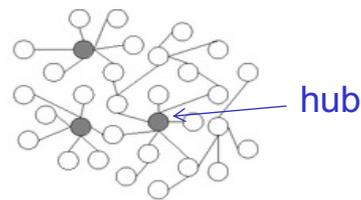
45

# Network models: Barabási-Albert

- **Scale-free property**: network diameter  $\sim \ln(\ln N)$
- Scale-free networks are **robust** against accidental node failures, but **vulnerable** to coordinated attacks against their hubs
- Many networks are conjectured to be scale-free: Internet, Web, social networks, power grids



(a) Random network



(b) Scale-free network

# Network models: comparison

	Random	Small-World	Scale-Free
Visualisation of the Topology			
Degree Distribution			
Robustness Characteristics	Responds similarly to both random and targeted attacks.		Resilient against random failures but very sensitive to targeted attacks.

Source: [appliednetsci.springeropen.com/articles/10.1007/s41109-017-0053-0](https://appliednetsci.springeropen.com/articles/10.1007/s41109-017-0053-0)

## SNA measures: centrality

---

- Let's analyze **centrality** measures

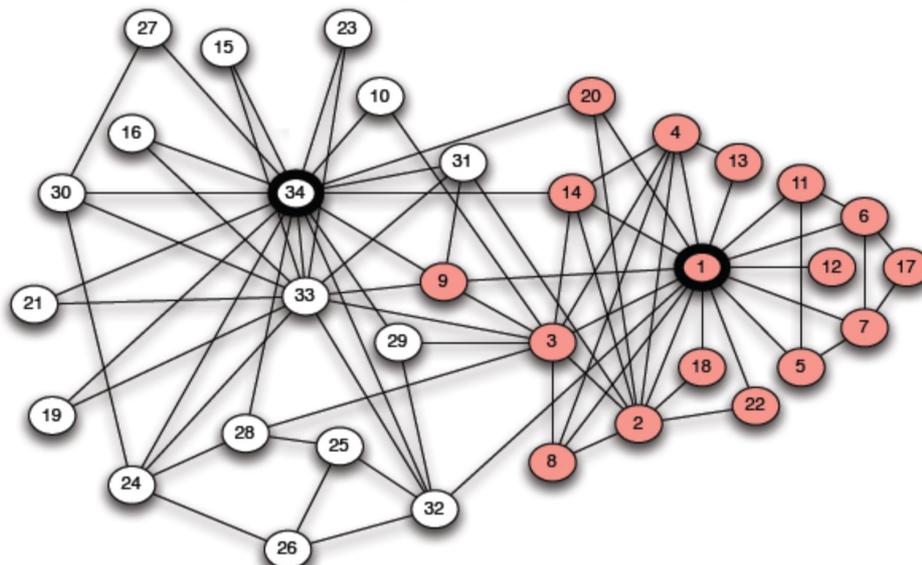
Which are the most important nodes within the network?

- To identify, e.g.,
  - Most influential persons in a social network
  - Key infrastructure nodes in Internet or urban networks
  - Super-spreaders of disease

## SNA measures: centrality

---

- The social network of friendships within a 34-person karate club provides clues to the fault lines that eventually split the club apart



## SNA measures: centrality

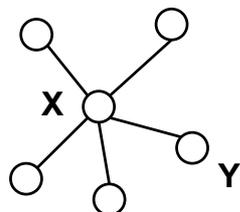
---

- Which nodes are most “central”?
- Definition of central varies by context/purpose
- Local measure:
  - Degree centrality
- Relative to rest of network:
  - Betweenness centrality
  - Closeness centrality
  - PageRank centrality
- For each measure: the higher the score, the more central the node

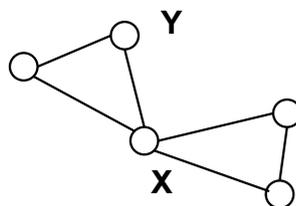
## Centrality: who’s important based on its network position

---

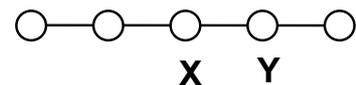
- In each of the following networks, X has higher centrality than Y according to a particular measure



degree



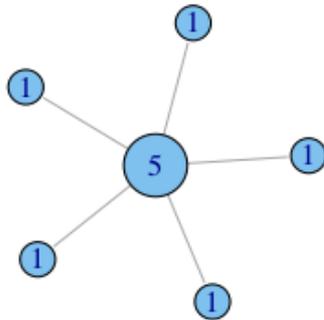
betweenness



closeness

# Degree centrality

- Idea: a node is important if it has many neighbors
- Count **how many neighbors a node has**

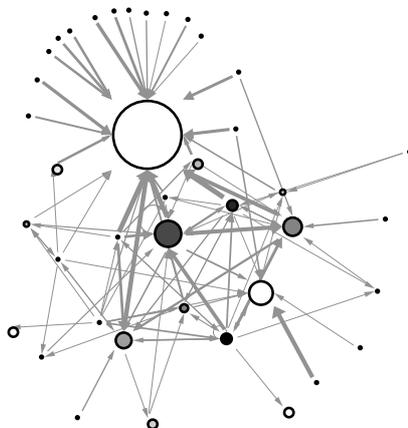


Degree centrality of a node: the **degree** of that node

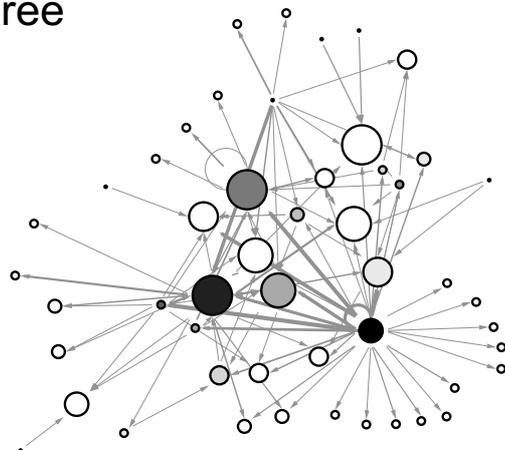
- Degree centrality on the node level can be extended to the whole graph:  $C_D$

## Degree centrality: examples

- Examples of financial trading networks
  - Nodes are sized by degree



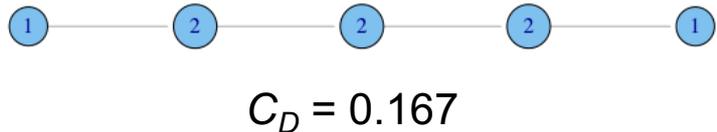
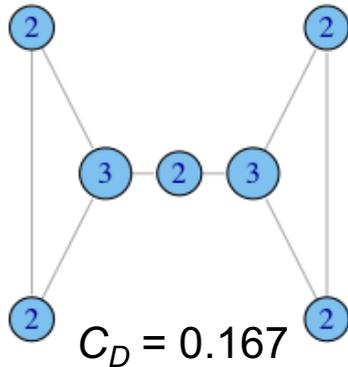
**High centralization:**  
one node trading with many others



**Low centralization:**  
trades are more evenly distributed

## When degree is not everything

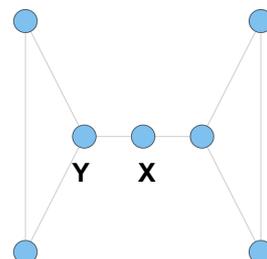
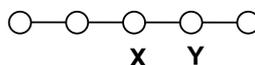
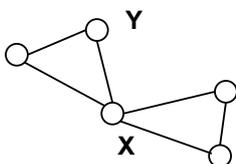
- In what ways does degree fail to capture centrality in the networks below?



- Ability to broker between groups
- Likelihood that information originating anywhere in the network reaches you

## Betweenness: another centrality measure

- Intuition: how many pairs of individuals would have to go through you in order to reach one another in the minimum number of hops?
- Betweenness centrality quantifies the **number of times a node acts as a bridge along the shortest path between each pair of nodes**
- Who has higher betweenness, X or Y?



## Betweenness: definition

---

- $C_B(i)$ , betweenness centrality of node  $i$ :
  - $V$ : set of nodes
  - $\sigma_{jk}(i)$ : number of shortest paths between nodes  $j$  and  $k$  that pass through node  $i$
  - $\sigma_{jk}$ : number of shortest paths between nodes  $j$  and  $k$

$$C_B(i) = \sum_{i \neq j \neq k \in V} \frac{\sigma_{jk}(i)}{\sigma_{jk}}$$

- Usually normalized by:

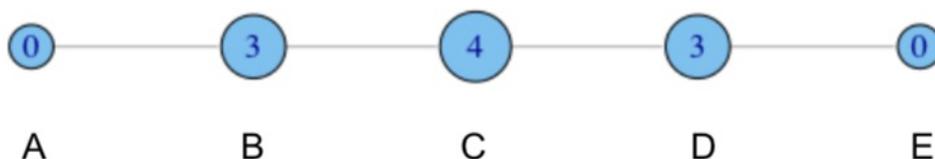
$$C'_B(i) = \frac{C_B(i)}{(N-1)(N-2)/2}$$

## Betweenness: definition and example

---

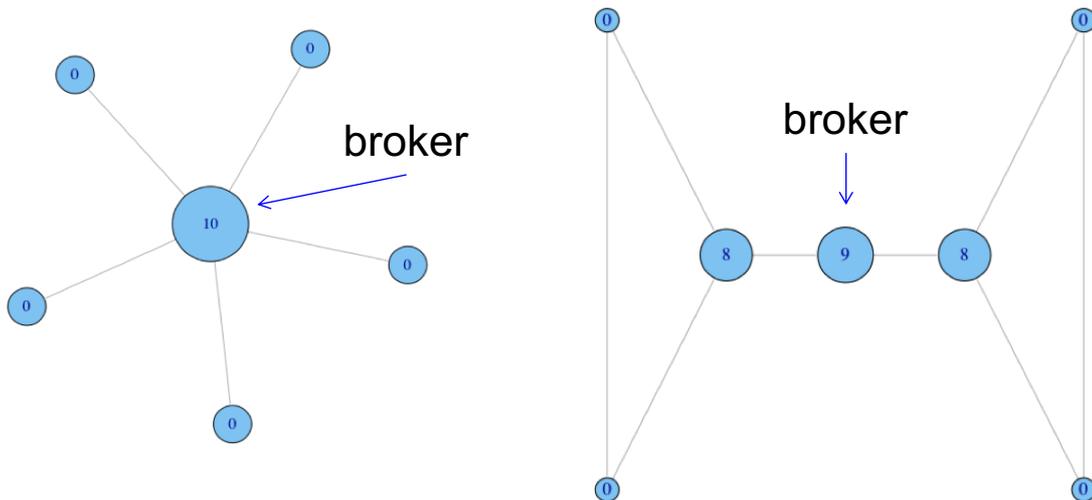
- Betweenness centrality for node  $i$  (do not consider denominator)

$$C_B(i) = \sum_{i \neq j \neq k \in V} \frac{\sigma_{jk}(i)}{\sigma_{jk}}$$



- Node B acts as a bridge 3 times along the shortest paths between A and C, A and D, A and E
- Node C acts as a bridge 4 times along the shortest paths between A and D, A and E, B and D, B and E

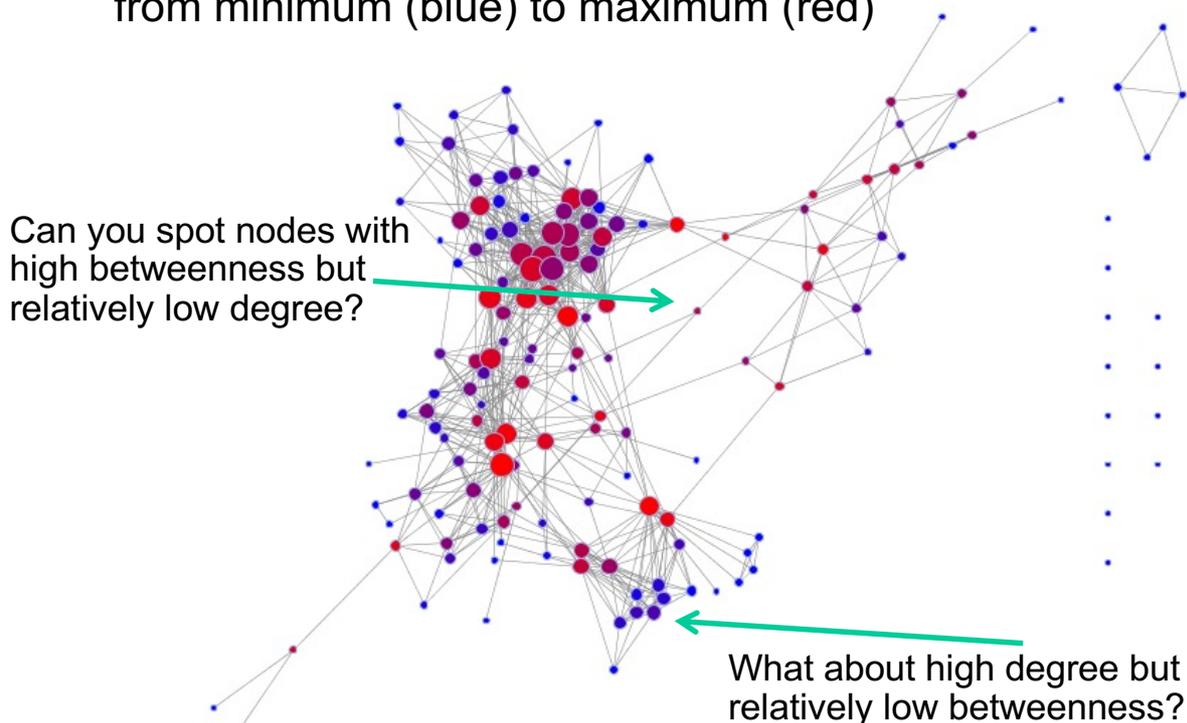
# Betweenness: example



- If the broker node is removed, the entire connection would be completely collapsed with the rest of the community, and so you will notice separated subgroups

# Betweenness: example

- Nodes sized by degree, and colored by betweenness from minimum (blue) to maximum (red)



## Closeness: another centrality measure

---

- What if it is not so important to have many direct friends (i.e., degree)?
- Or be “between” others (i.e., betweenness)?
- But one still wants to be in the “middle” of things, not too far from the center

## Closeness centrality: definition

---

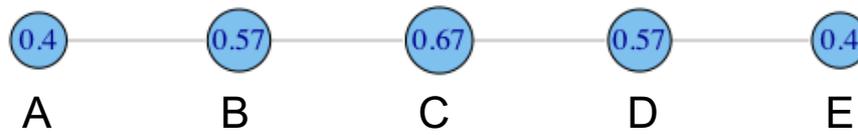
- Closeness of a node is the **average length of the shortest paths between the node and all other nodes in the network**
  - The more central a node is, the closer it is to all other nodes
- $C_C(i)$ , closeness centrality for node  $i$ :
  - $d(i,j)$ : length of the shortest path between nodes  $i$  and  $j$

$$C_C(i) = \left( \sum_{j=1, j \neq i}^N d(i, j) \right)^{-1}$$

- Normalized closeness centrality for node  $i$

$$C'_C(i) = (N - 1)C_C(i) = \frac{N - 1}{\sum_{j=1, j \neq i}^N d(i, j)}$$

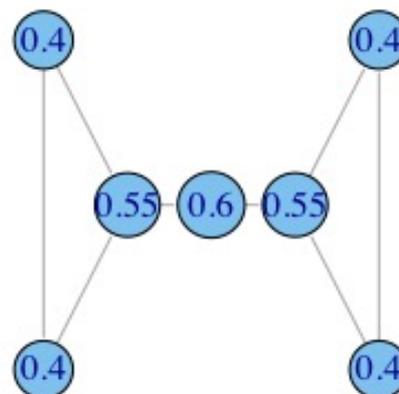
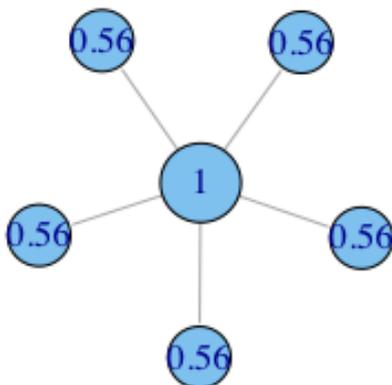
## Closeness centrality: example



$$C'_C(A) = \frac{N - 1}{\sum_{j=1, j \neq A}^N d(A, j)} = \frac{4}{1 + 2 + 3 + 4} = \frac{4}{10} = 0.4$$

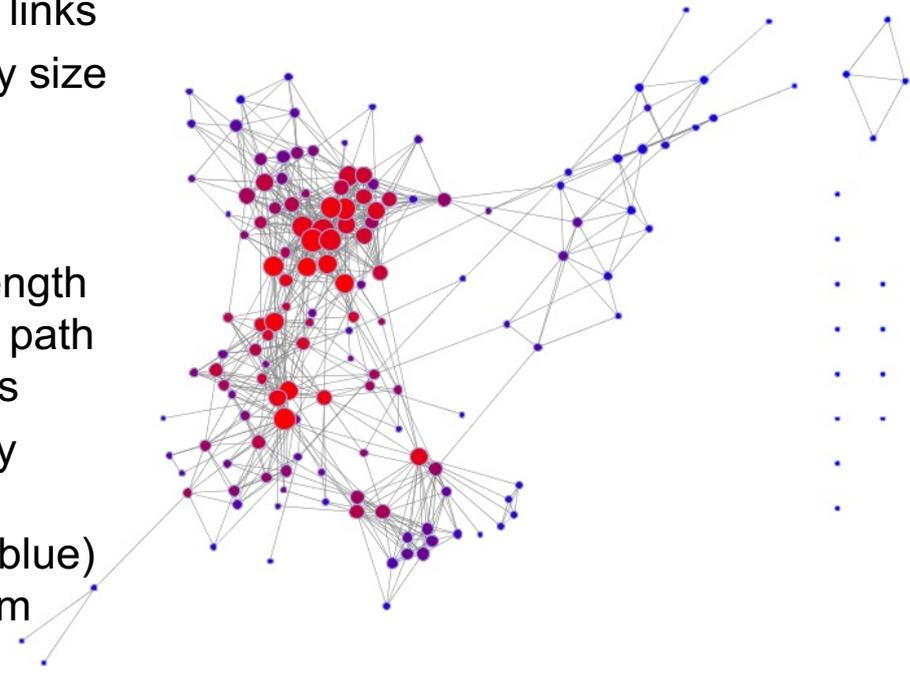
$$C'_C(C) = \frac{N - 1}{\sum_{j=1, j \neq C}^N d(C, j)} = \frac{4}{1 + 2 + 1 + 2} = \frac{4}{6} = 0.67$$

## Closeness centrality: examples



# How closely does degree correspond to closeness?

- Degree
  - Number of links
  - Denoted by size
- Closeness
  - Average length of shortest path to all others
  - Denoted by color: from minimum (blue) to maximum (red)

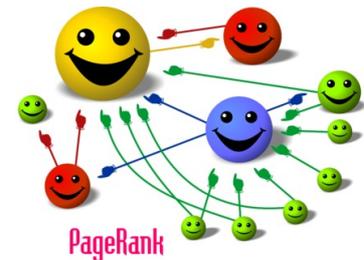


Valeria Cardellini - ADS 2023/24

64

## PageRank and centrality

- PageRank idea: *a node is important if it is linked from other important or if it is highly linked*
- PageRank can also be used to measure centrality
  - Differently from the other centrality measures, it accounts for **link direction**
    - Each node in a network is assigned a score based on its number of incoming links; these links are also weighted depending on the relative score of its originating node
    - The result is that nodes with many incoming links are influential, and nodes to which they are connected share some of that influence



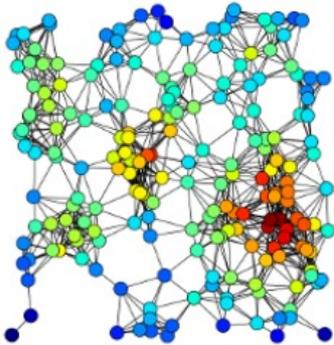
Valeria Cardellini - ADS 2023/24

65

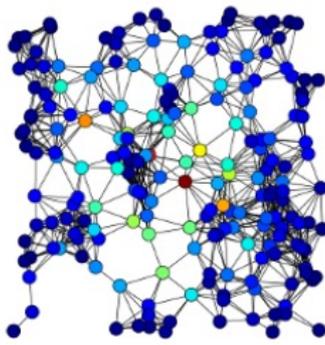
## Wrap-up on centrality

- We have looked at some popular centrality measures

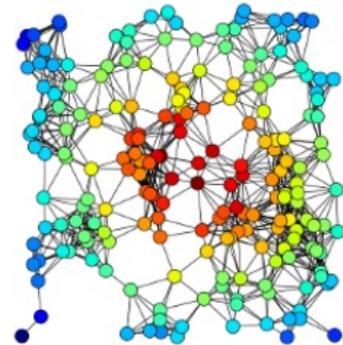
"There is certainly no unanimity on exactly what centrality is or on its conceptual foundations, and there is little agreement on the proper procedure for its measurement." (Freeman, 1979)



Degree



Betweenness

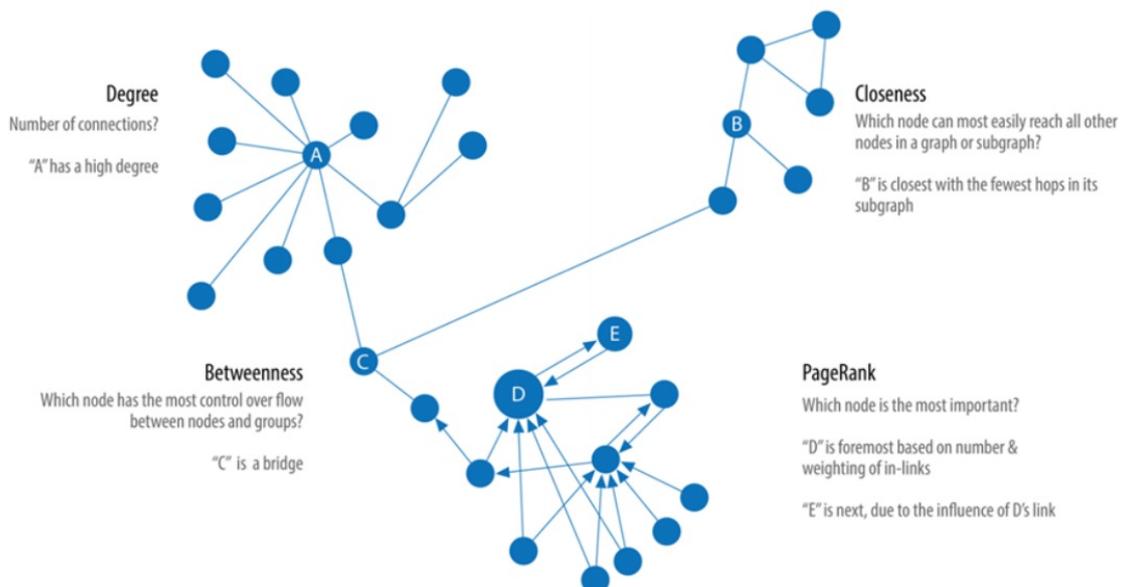


Closeness

Color from minimum (blue) to maximum (red)

## Wrap-up on centrality

- Different centrality measures can produce significantly different results based on what they were created to measure



## Wrap-up on centrality

---

- Degree centrality
  - *What it tells us:* number of links each node has to other nodes in the network
  - *When to use it:* to find very connected individuals, popular individuals, individuals who are likely to hold most information or individuals who can quickly connect with the wider network
- Betweenness centrality
  - *What it tells us:* which nodes are “bridges” between nodes in a network; measures the number of times a node lies on the shortest path between other nodes
  - *When to use it:* to find individuals who influence the flow around a network

## Wrap-up on centrality

---

- Closeness centrality
  - *What it tells us:* scores each node based on its “closeness” to all other nodes in the network; calculates the shortest paths between all nodes, then assigns each node a score based on its sum of shortest path lengths
  - *When to use it:* to find individuals who are best placed to influence the entire network most quickly
- PageRank
  - *What it tells us:* node’s importance from its linked neighbors and their neighbors; uncovers nodes whose influence extends beyond their direct links into the wider network
  - *When to use it:* to understand authority and citations

# What is a community?

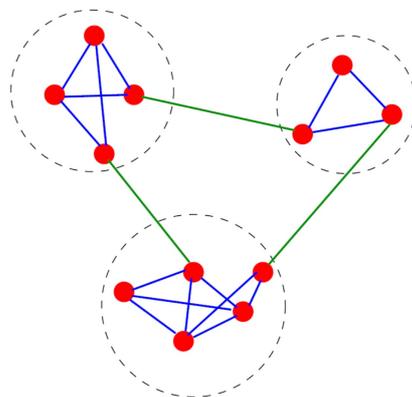
---

- Set of nodes that share something:
  - Affiliation (friends, colleagues, club, ...)
  - Similar interests (tagging systems, ...)
  - Similar contents (movies, books, products, ...)
- What is the connection with the network structure?
- Intuition: **more densely connected inside than outside**

## SNA: Community detection

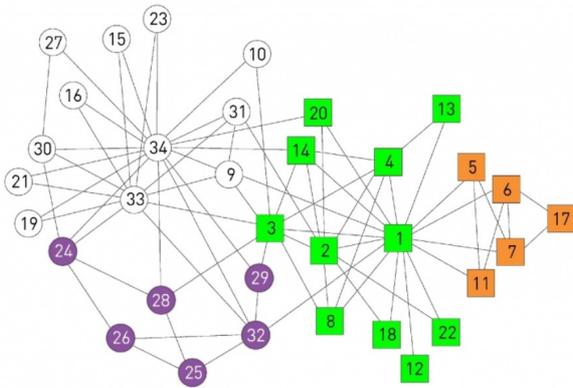
---

- Goal: identify automatically communities with statistically significantly more links between nodes in the same group than nodes in different groups



# Community detection: example

---



- Zachary observed 34 members of a karate club over 2 years. Links connect individuals who were observed to interact outside the club activities

During the course of observation, the club members split into 2 groups because of the disagreement between the club administrator and the club instructor (nodes 1 and 34), and the members of one group left to start their own club

[en.wikipedia.org/wiki/Zachary%27s\\_karate\\_club](https://en.wikipedia.org/wiki/Zachary%27s_karate_club)

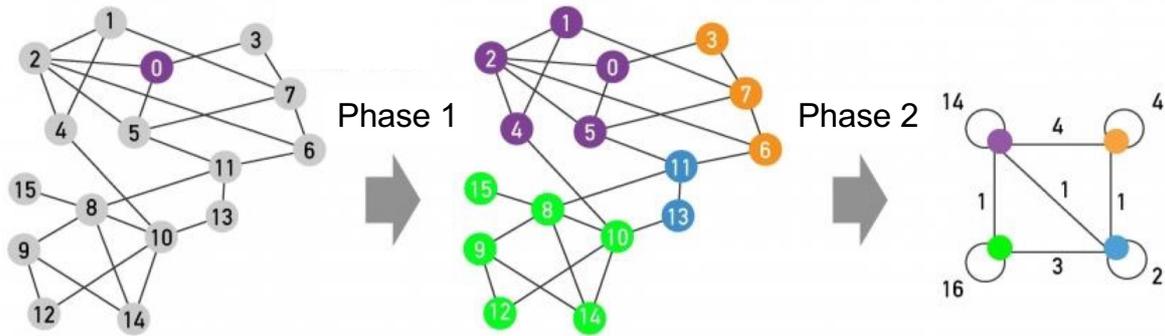
# Community detection: Louvain algorithm

---

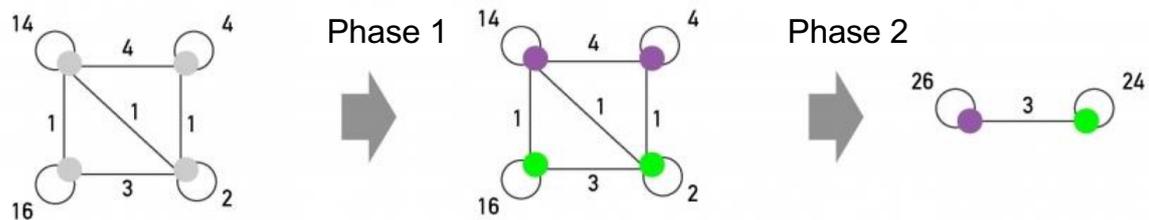
- **Louvain algorithm:** greedy optimization method that aims to maximize modularity
  - **Modularity:** value in  $[-1, 1]$  that measures the **density** of links inside communities compared to links between communities
  - Idea: 2 phases that are repeated iteratively
    - In phase 1, each node is first assigned to its own community. Then, for each node  $i$  the change in modularity is calculated by removing  $i$  from its own community and placing it into the community of each neighbor  $j$ .  $i$  is placed into the community that resulted in the greatest modularity increase
    - In phase 2, the algorithm groups all of the nodes in the same community and builds a new network whose nodes are the communities from previous phase. Then, phase 1 is re-applied to the new network until a maximum of modularity is attained
  - Running time:  $O(N \log N)$ , being  $N$  the number of nodes

# Community detection: Louvain algorithm

## 1<sup>ST</sup> PASS



## 2<sup>ND</sup> PASS



## References

- D. Easley and J. Kleinberg, [Networks, Crowds, and Markets: Reasoning About a Highly Connected World](#), chapters 1 and 2, Cambridge University Press, 2010.
- A.-L. Barabási, [Network Science](#), chapters 1, 2, 3, and 9, Cambridge University Press, 2016.