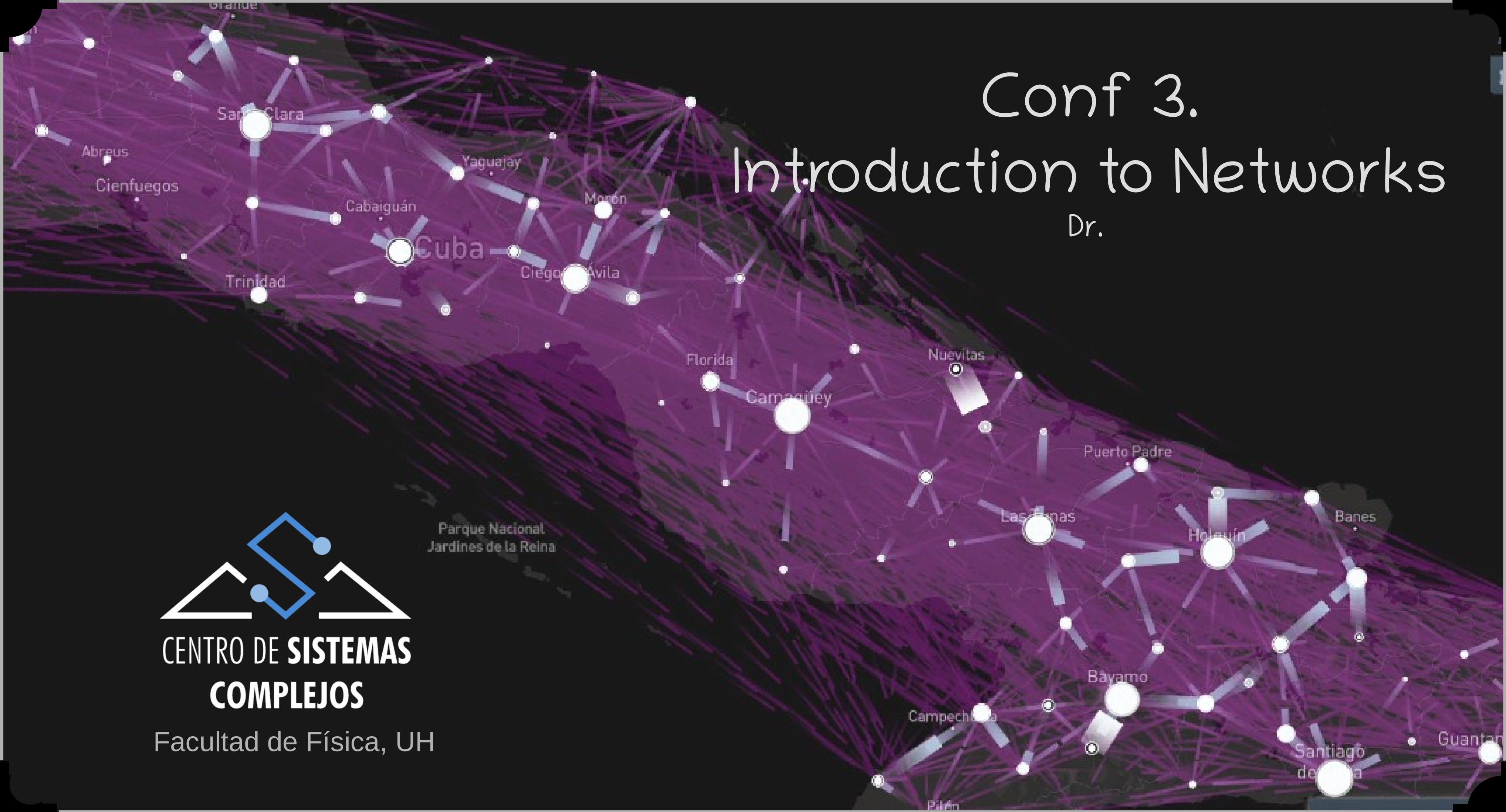


Conf 3. Introduction to Networks

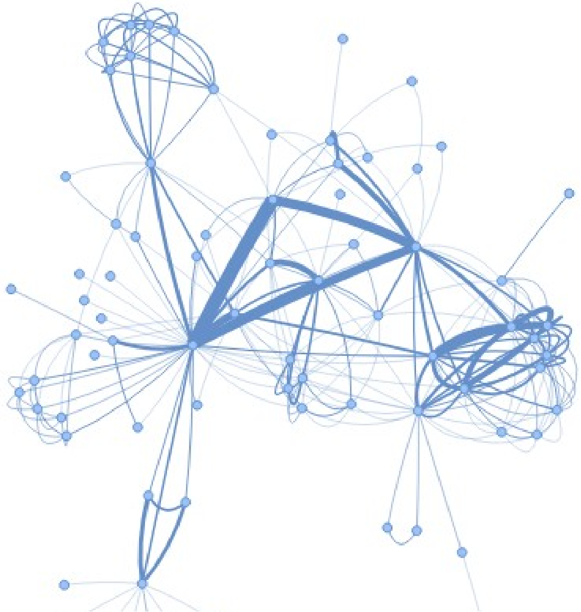
Dr.

CENTRO DE SISTEMAS COMPLEJOS

Facultad de Física, UH

Fast introduction to Networks

- Graphs, mathematical deifinition
- Which nodes are important?
- Local and global properties and parameters
- Small famous networks, Large networks
- Random graphs
- Building your own graphs: Regular, poissonian, SBM
- Visualizing graphs



Dynamical processes on complex networks

Marc Barthélemy

Lectures IPhT 2010

The structure and function of complex networks

M. E. J. Newman

Department of Physics, University of Michigan, Ann Arbor, MI 48109, U.S.A. and Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, U.S.A.

Vertex, edges and Graphs

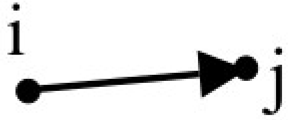
Vertex = nodes Edges = links

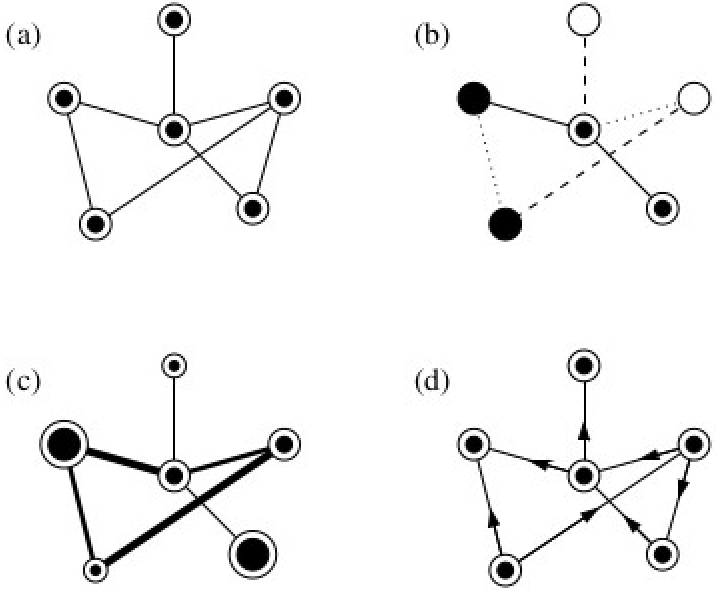
(a) Simple graph

(b) Types of nodes, edges

(c) Weighted vs Weigthed

(d) Directed



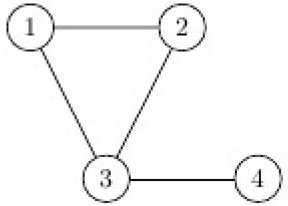


Graphs

Definition: A *graph* G is a pair (V, E), where V is a set of vertices (or nodes), and E is a set of edges, where each edge is a pair of vertices $(u, v) \in V \times V$.

- The vertices are said to be adjacent if they are endpoints of an edge.
- The set of all edges incident to a vertex v is called the *neighborhood* of v, denoted by ∂v.

$\partial_1 = \{2, 3\}$

$G =$ 

$\partial_3 = \{1, 2, 4\}$

Consider the graph $G = (V, E)$, where $V = \{1, 2, 3, 4\}$ and

$$E = \{(1,2), (1,3), (2,3), (3,4)\}$$

The *adjacency matrix* A_G of a graph G

- a square matrix of size $n \times n$, where $n = |V|$,
- if $(u, v) \in E$, then $A_G(u, v) = 1$
- else $A_G(u, v) = 0$

$$A_G = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Graphs

Definition: A *graph* G is a pair (V, E), where V is a set of vertices (or nodes), and E is a set of edges, where each edge is a pair of vertices $(u, v) \in V \times V$.

- The vertices are said to be adjacent if they are endpoints of an edge.
- The set of all edges incident to a vertex v is called the *neighborhood* of v, denoted by ∂v.

Consider the graph $G = (V, E)$, where $V = \{1, 2, 3, 4\}$ and

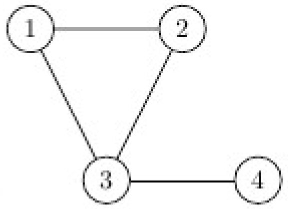
$$E = \{(1,2), (1,3), (2,3), (3,4)\}$$

Degree of a Node: The *degree* of a node v in a graph $G = (V, E)$ is the number of edges incident to v, i.e., the number of neighbors of v. The degree of v is denoted by $deg(v)$.

$\partial_1 = \{2, 3\}$

$deg(1) = 2.$

$G =$



$\partial_3 = \{1, 2, 4\}$

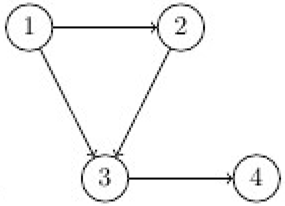
$$A_G = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Directed Graphs

Definition: A *graph* G is a pair (V, E), where V is a set of vertices (or nodes), and E is a set of edges, where each edge is a pair of vertices $(u, v) \in V \times V$.

- The vertices are said to be adjacent if they are endpoints of an edge.
- The set of all edges incident to a vertex v is called the *neighborhood* of v, denoted by ∂v.

$G =$



Consider the graph $G = (V, E)$, where $V = \{1, 2, 3, 4\}$ and

$$E = \{(1,2), (1,3), (2,3), (3,4)\}$$

Degree of a Node: The *degree* of a node v in a graph $G = (V, E)$ is the number of edges incident to v, i.e., the number of neighbors of v. The degree of v is denoted by $deg(v)$.

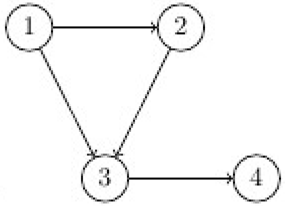
$$A_G = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Directed Graphs

Definition: A *graph* G is a pair (V, E), where V is a set of vertices (or nodes), and E is a set of edges, where each edge is a pair of vertices $(u, v) \in V \times V$.

- The vertices are said to be adjacent if they are endpoints of an edge.
- The set of all edges incident to a vertex v is called the *neighborhood* of v, denoted by ∂v.

$G =$



Consider the graph $G = (V, E)$, where $V = \{1, 2, 3, 4\}$ and

$$E = \{(1,2), (1,3), (2,3), (3,4)\}$$

- Vertex 1: in-degree = 0, out-degree = 2
- Vertex 2: in-degree = 1, out-degree = 1

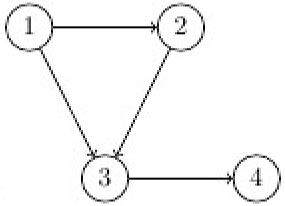
Degree of a Node: The *degree* of a node v in a graph $G = (V, E)$ is the number of edges incident to v, i.e., the number of neighbors of v. The degree of v is denoted by $deg(v)$. In a directed graph, the degree of a node is the sum of its in-degree (number of incoming edges) and out-degree (number of outgoing edges).

$$A_G = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Paths and cycles

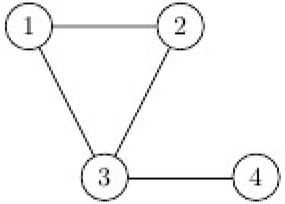
Definition: A *graph* G is a pair (V, E), where V is a set of vertices (or nodes), and E is a set of edges, where each edge is a pair of vertices $(u, v) \in V \times V$.

- The vertices are said to be adjacent if they are endpoints of an edge.
- The set of all edges incident to a vertex v is called the *neighborhood* of v, denoted by ∂v.

$G =$ 

Consider the graph $G = (V, E)$, where $V = \{1, 2, 3, 4\}$ and

$$E = \{(1, 2), (1, 3), (2, 3), (3, 4)\}$$

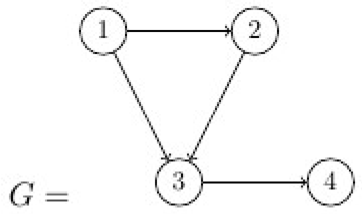
$G =$ 

Paths: A *path* in a graph $G = (V, E)$ is a sequence of distinct vertices $v_1, v_2, \ldots, v_k$ such that $(v_i, v_{i+1}) \in E$ for $i = 1, 2, \ldots, k - 1$. The length of the path is $k - 1$, and it is denoted by $|P|$. A path is said to be *simple* if all its vertices are distinct.

Paths and cycles

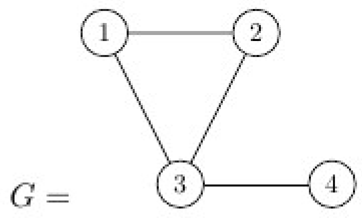
Definition: A *graph* G is a pair (V, E), where V is a set of vertices (or nodes), and E is a set of edges, where each edge is a pair of vertices $(u, v) \in V \times V$.

- The vertices are said to be adjacent if they are endpoints of an edge.
- The set of all edges incident to a vertex v is called the *neighborhood* of v, denoted by ∂v.



Consider the graph $G = (V, E)$, where $V = \{1, 2, 3, 4\}$ and

$$E = \{(1,2), (1,3), (2,3), (3,4)\}$$

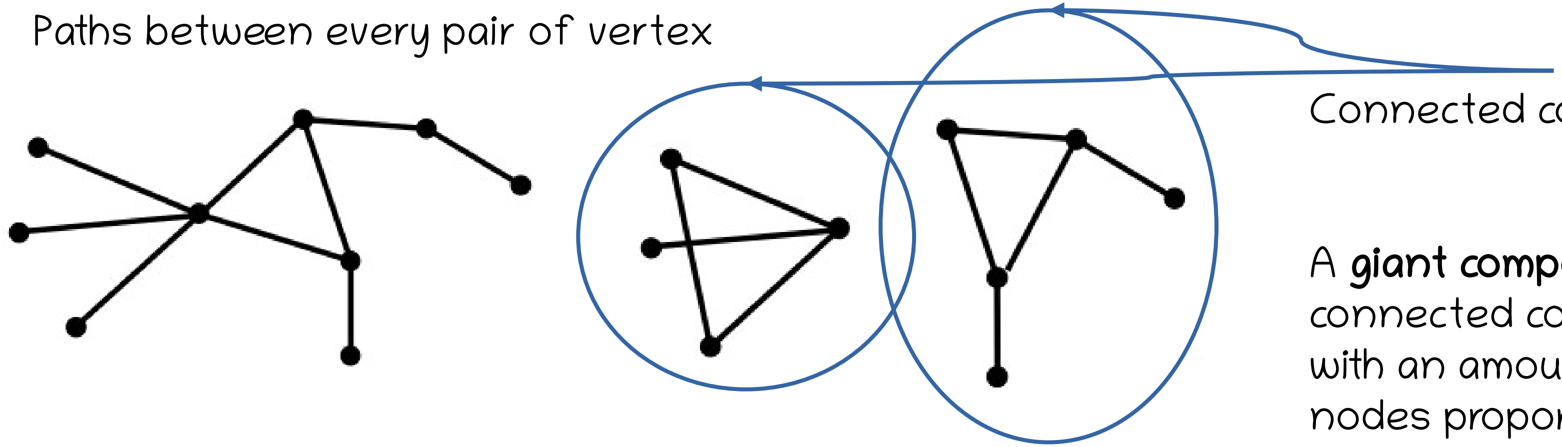


Paths: A *path* in a graph $G = (V, E)$ is a sequence of distinct vertices $v_1, v_2, \ldots, v_k$ such that $(v_i, v_{i+1}) \in E$ for $i = 1, 2, \ldots, k-1$. The length of the path is $k - 1$, and it is denoted by $|P|$. A path is said to be *simple* if all its vertices are distinct.

Cycles: A *cycle* in a graph $G = (V, E)$ is a simple path $v_1, v_2, \ldots, v_k$ such that $(v_k, v_1) \in E$. The length of the cycle is k, and it is denoted by $|C|$.

Connected components

Paths between every pair of vertex



Connected components

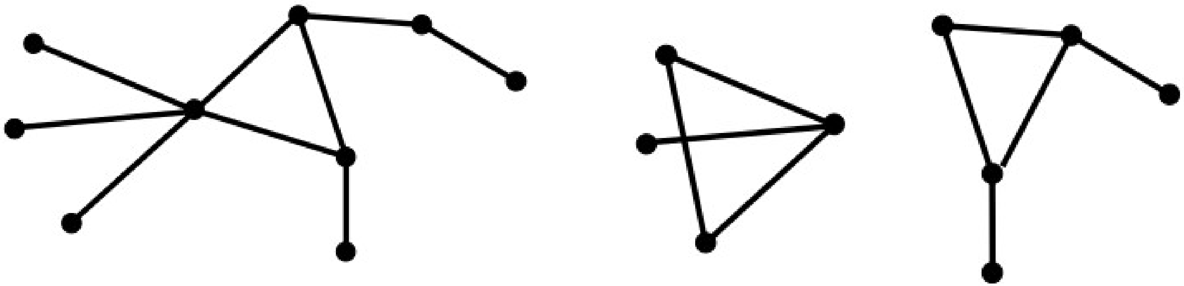
A **giant component**: a connected component with an amount of nodes proportional to N.

Paths: A *path* in a graph $G = (V, E)$ is a sequence of distinct vertices $v_1, v_2, \dots, v_k$ such that $(v_i, v_{i+1}) \in E$ for $i = 1, 2, \dots, k-1$. The length of the path is $k - 1$, and it is denoted by $|P|$. A path is said to be *simple* if all its vertices are distinct.

Cycles: A *cycle* in a graph $G = (V, E)$ is a simple path $v_1, v_2, \dots, v_k$ such that $(v_k, v_1) \in E$. The length of the cycle is k, and it is denoted by $|C|$.

Connected components

Paths between every pair of vertex



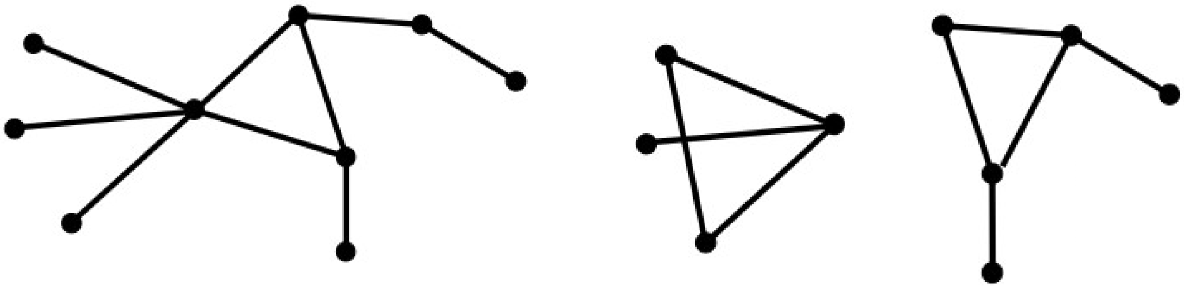
Density: For a graph $G(V, E)$, density is defined as

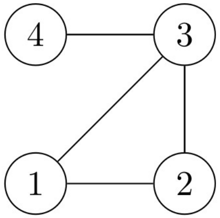
$$D = \frac{|E|}{|V|}$$

Graphs with $D \ll 1$ are named sparse.

Laplacian Matrix

Paths between every pair of vertex





$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Given a graph G with n vertices, the Laplacian matrix L is defined as:

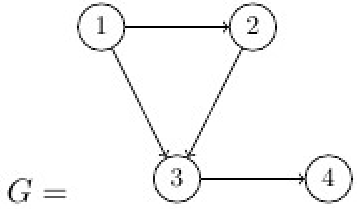
$$L = D - A$$

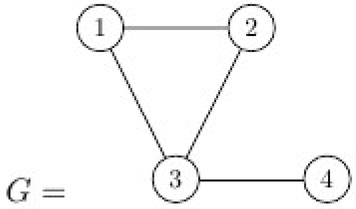
where D is the degree matrix and A is the adjacency matrix of G.

Shortests path

The set of all paths.

The shortest path: *geodesic distance*





Paths: A *path* in a graph $G = (V, E)$ is a sequence of distinct vertices $v_1, v_2, \ldots, v_k$ such that $(v_i, v_{i+1}) \in E$ for $i = 1, 2, \ldots, k-1$. The length of the path is $k - 1$, and it is denoted by $|P|$. A path is said to be *simple* if all its vertices are distinct.

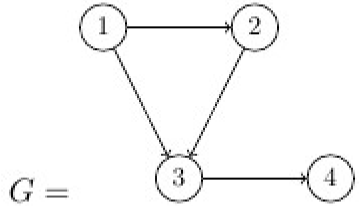
Cycles: A *cycle* in a graph $G = (V, E)$ is a simple path $v_1, v_2, \ldots, v_k$ such that $(v_k, v_1) \in E$. The length of the cycle is k, and it is denoted by $|C|$.

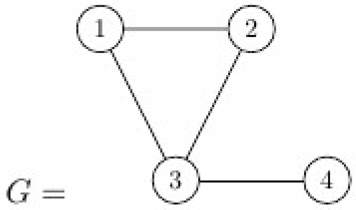
Shortests path

The set of all paths: **breadth first search**

The shortest path: *geodesic distance*

Powers of the Adjacency Matrix.





Distance metrics

Mean distance: Average of distances between all pairs of nodes in a **Connected Graph**

$$\ell = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}$$

Eccentricity of vertex v: distance to the farthest vertex

$$\epsilon(v) = \max_{u \in V} d(v, u)$$

Radius of a graph: shortest eccentricity

$$r = \min_{v \in V} \epsilon(v) = \min_{v \in V} \max_{u \in V} d(v, u)$$

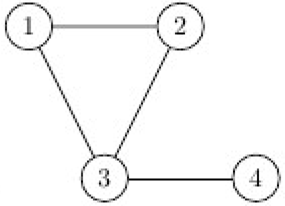
Diameter of a graph: largest eccentricity

$$d = \max_{(v \in V)} \epsilon(v) = \max_{(v \in V)} \max_{(u \in V)} d(v, u)$$

Jordan center: Nodes with minimum eccentricity

$$\{v_1, \ldots, v_n\} | \forall_{i \in 1 \ldots n} \epsilon(v_i) = r$$

$G =$



Try all in this graph

Shortests path

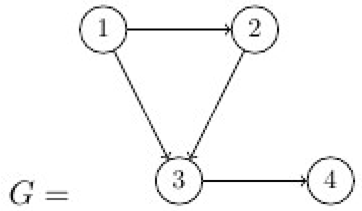
Complete graph: $\ell(i,j) = 1 \ \ \forall i, j$

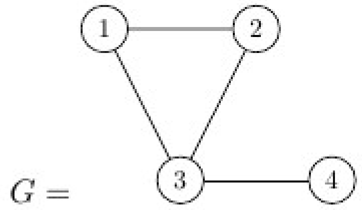
Regular lattice: $\ell(i,j) = |\vec{u}_i - \vec{u}_j|$

$$\bar{\ell} \sim N^{1/d}$$

"Small-world" network

$$\bar{\ell} \ll N^{1/d} (\text{eg. } \log N)$$





Centrality measures

Degree centrality: fraction of system connected to node i

$$C_d(i) = \frac{k_i}{n-1}$$

Betweeness: are you part of many shortest paths $C_b(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$

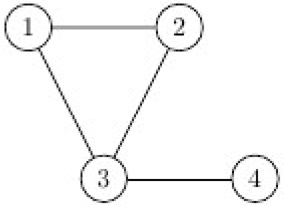
Clustering coefficient: are your neighbors, also my neighbors

$$C_i = \frac{2e_i}{k_i(k_i-1)}$$

where e_i is the number of edges between the neighbors of node i, and k_i is the degree of node i

Global clustering coefficient:

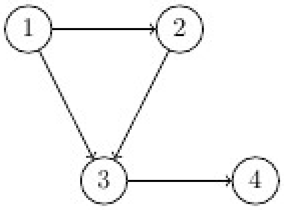
$$C = \frac{1}{n}\sum_{i=1}^{n} C_i$$

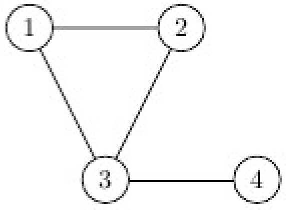
$G =$


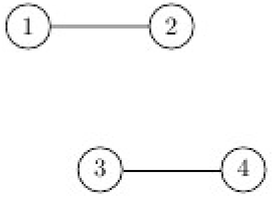
Try all in this graph

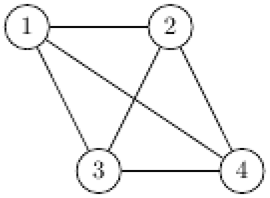
Graphs properties

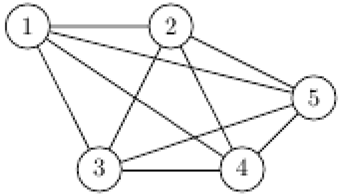
- **Directed vs undirected**
- Connected vs **Disconnected**
- **Regular** vs Non regular
- Planar vs **Non planar**
- Cyclic vs **Tree**
- **Bipartite**
- **Weighted**

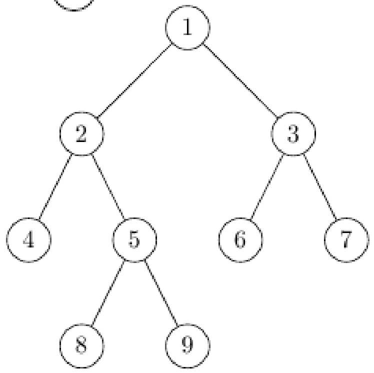


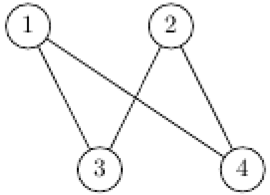


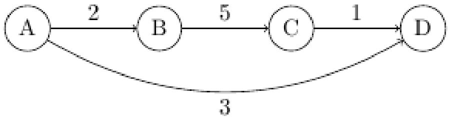












famous networks

- Useful for benchmarking
- Field specific

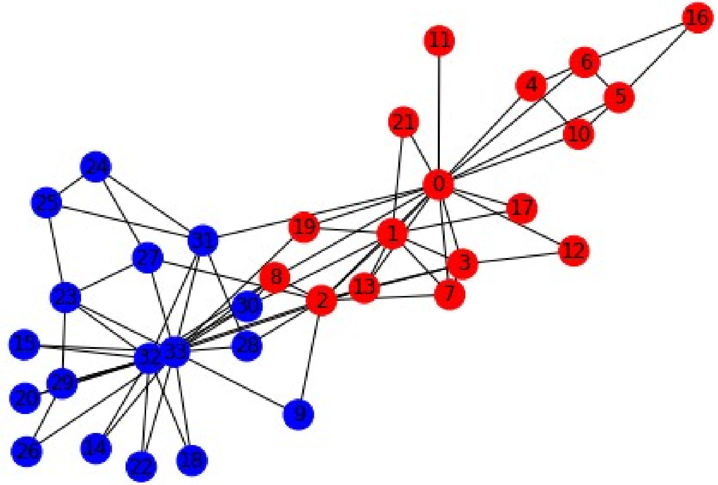
Zachary's Karate Club

Studied by Zachary, in a paper in 1977

- 34 nodes, people in a Club
- 78 edges, "friendship"
- Used for community detection

Characteristics

- Undirected
- Connected



famous networks

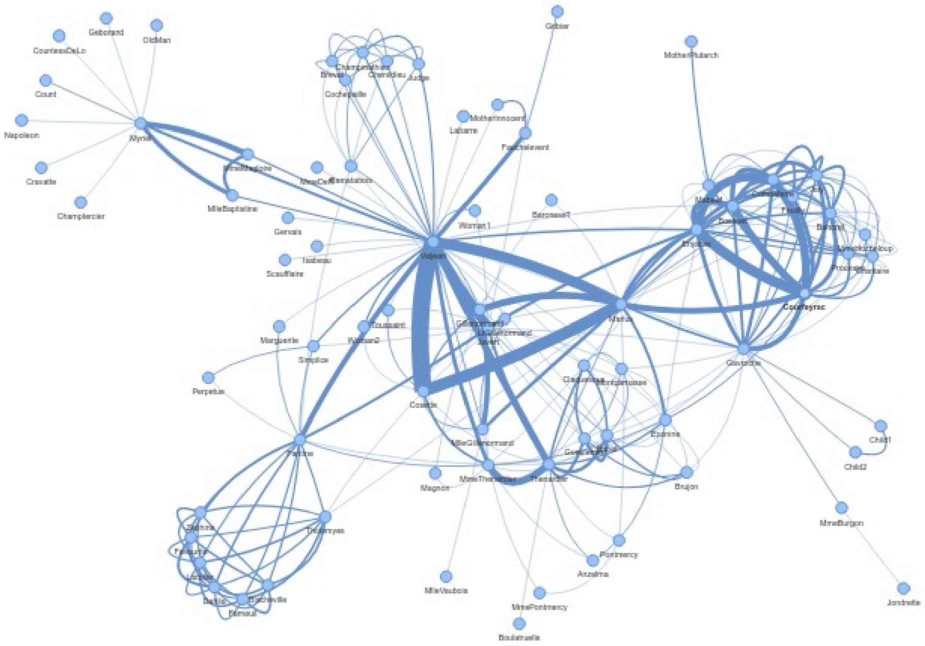
- Useful for benchmarking
- Field specific

Les Miserables (Novel by Victor Hugo)

- 77 nodes, characters in the novel
- 254 edges, present in the same chapt
- Used for community detection

Characteristics

- Undirected
- Connected
- Wheigted

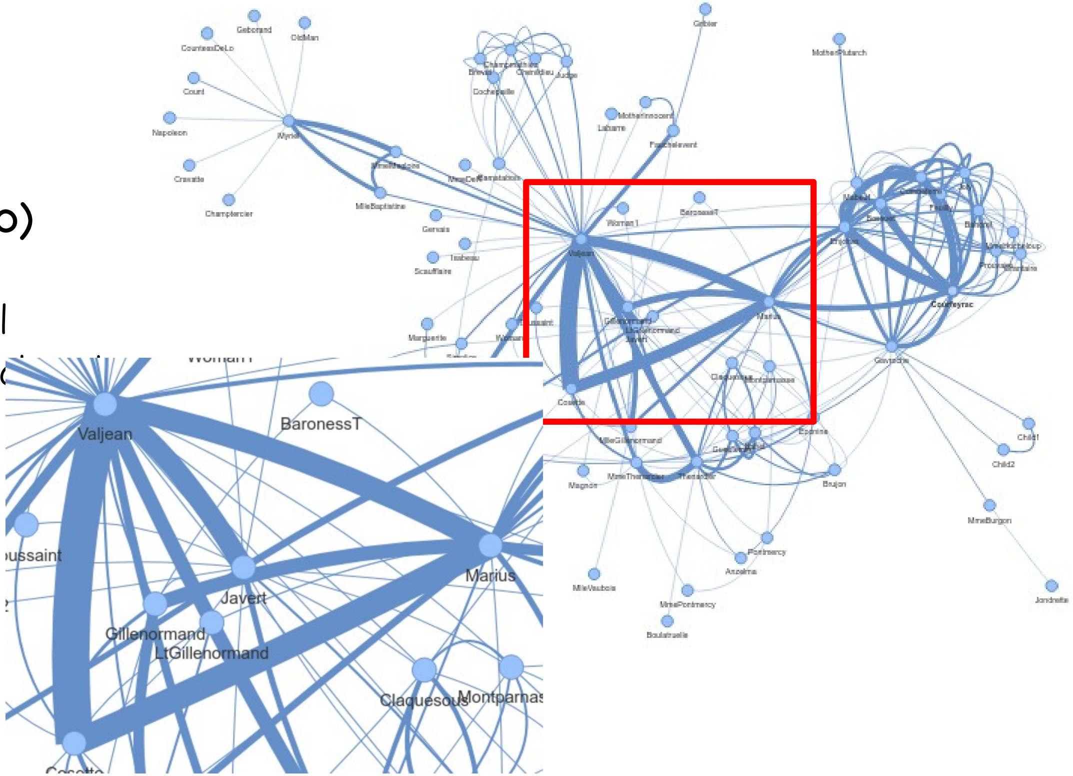


famous networks

- Useful for benchmarking
- Field specific

Les Miserables (Novel by Victor Hugo)

- 77 nodes, characters in the novel
- 254 edges, present in the same c
- Used for community detection



famous networks

- Useful for benchmarking
- Filed specific

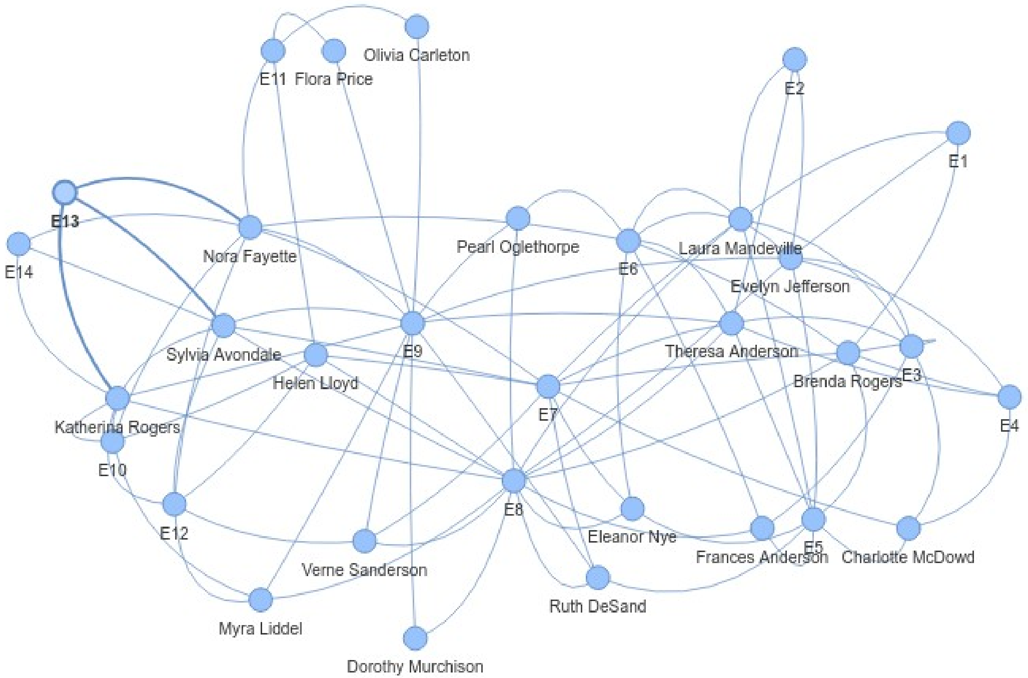
Davis Southern Women Graph

Studied by A. Davis, in a paper in 1941

- 32 nodes, women in a Social Club
- 89 edges, social events together

Characteristics

- Undirected
- Connected
- Weighted



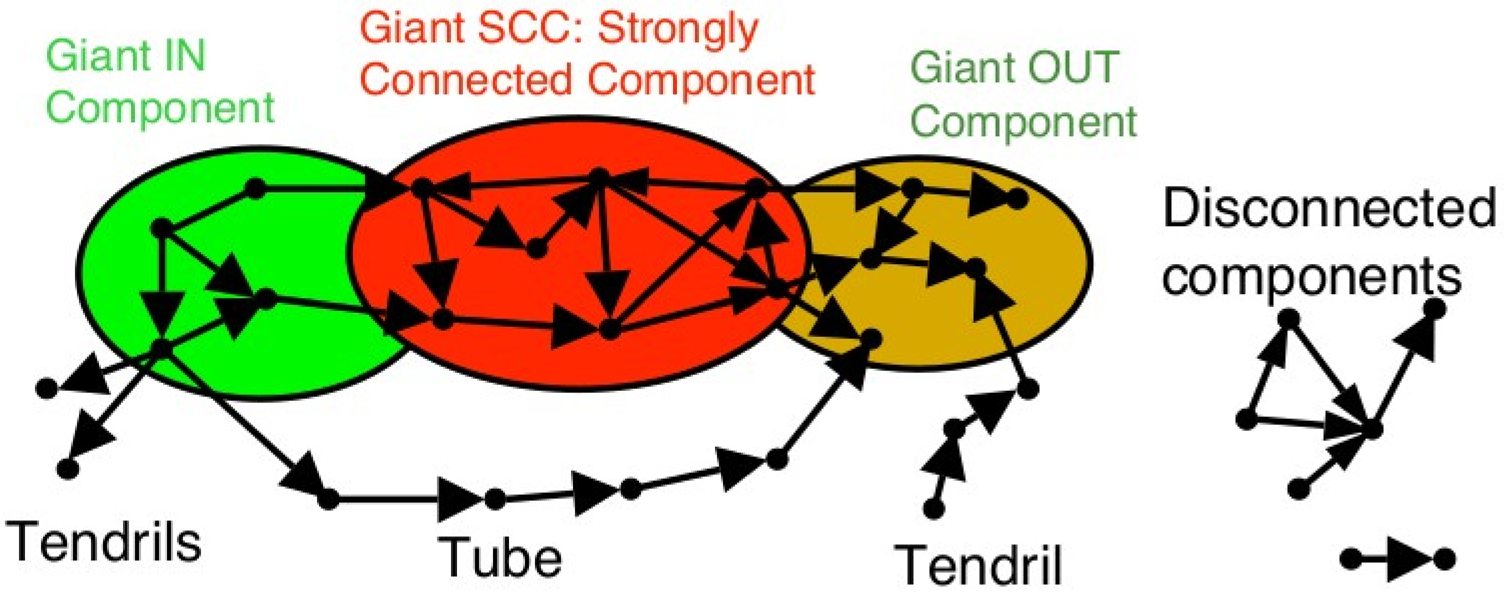
A. Davis, Gardner, B. B., Gardner, M. R., 1941. Deep South. University of Chicago Press, Chicago, IL.

famous networks

- Useful for benchmarking
- Filed specific

Internet Graph

Broder et al, 2000



Characteristics

- Directed
- Disconnected

famous networks

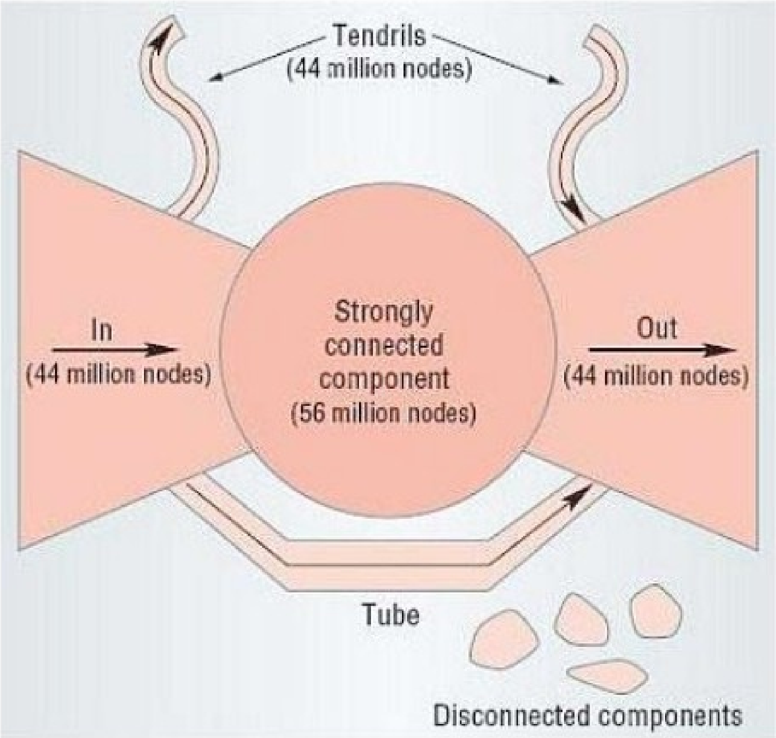
- Useful for benchmarking
- Filed specific

Internet Graph

Broder et al, 2000

Characteristics

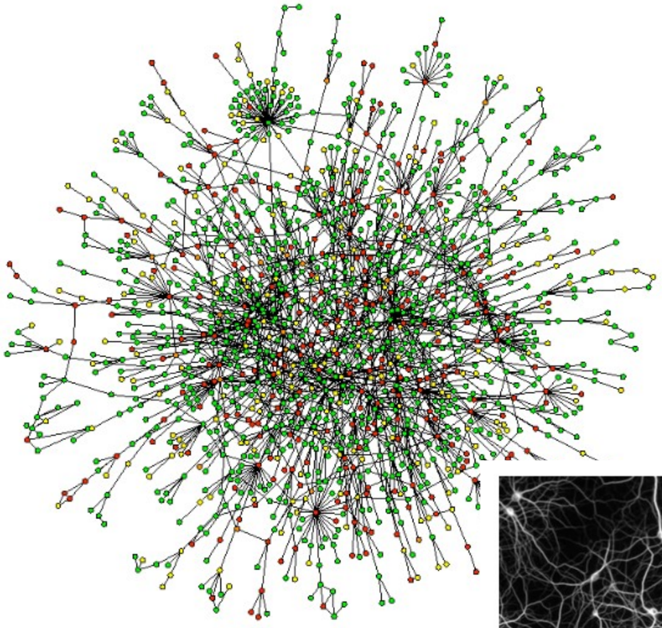
- Directed
- Disconnected

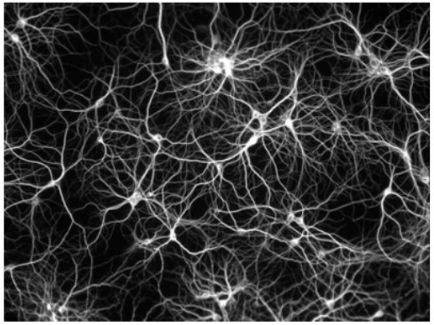


famous networks

In biology

- Trophic chain
- Protein-protein interaction network
- Gene regulatory network
- Metabolic networks
- Brain connectivity
- Human circulatory network
- Phylogenetic trees

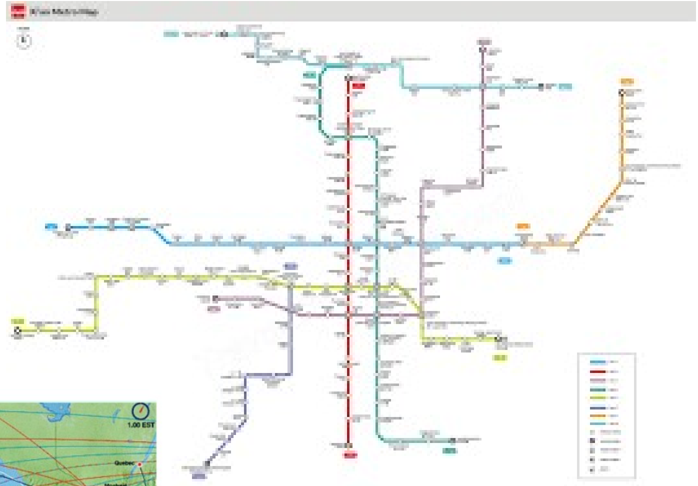


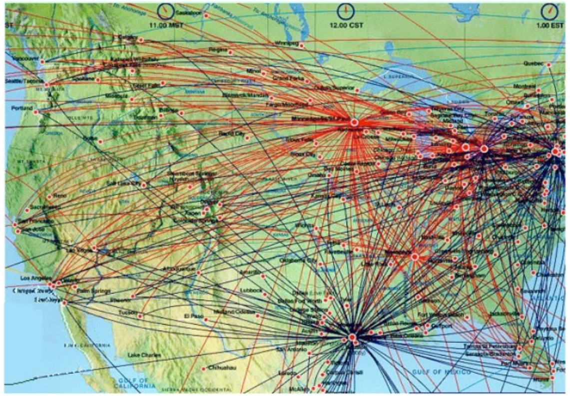


famous networks

Technological networks

- Electric grid
- Water distribution system
- City Transport networks
- Airtravel networks

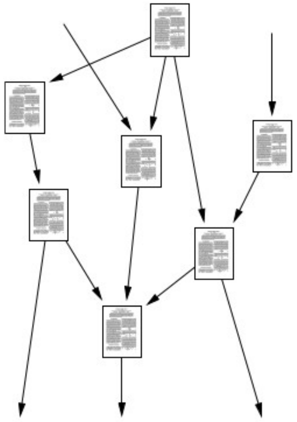




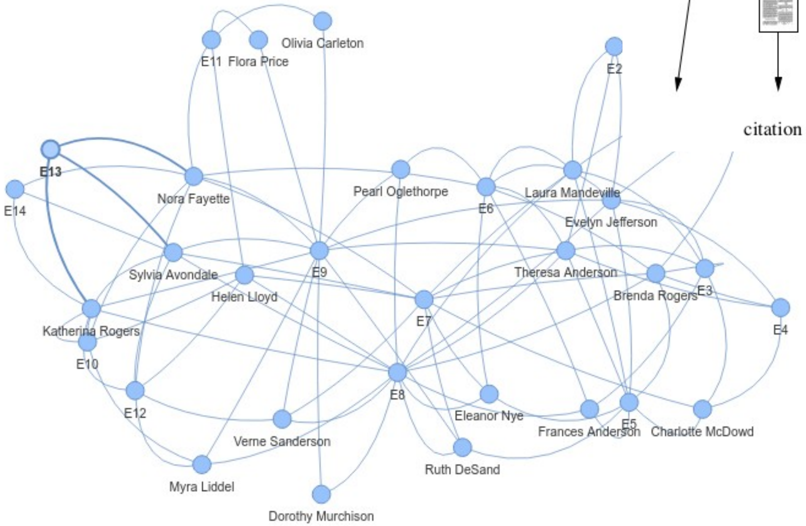
famous networks

Social Networks

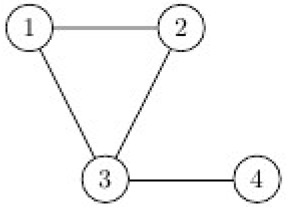
- Citation network
- Scientific collaboration network
- Twitter, instagram, facebook
- Sex couples
- Road networks
- Human mobility network
- Air trafic network
- Genealogic trees



citation network



Statistical description



Full description G(V,E). When are two graphs equal?

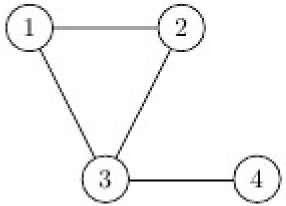
Graph ensembles

Degree list: $k_1, k_2, \ldots, k_N$ extensive in N.

Degree histogram: N_k number of nodes with a certain degree k.

Degree distribution: $P(k)=N_k/N$

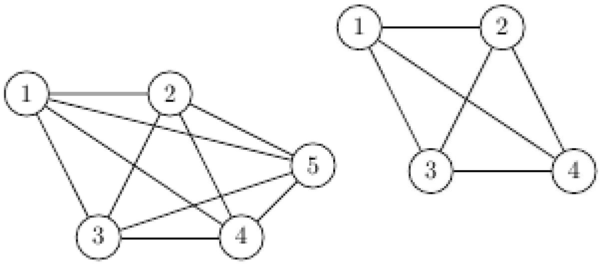
Statistical description



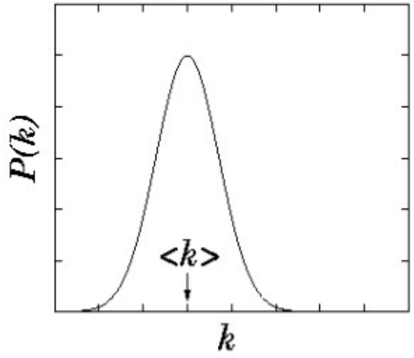
Full description G(V,E). When are two graphs equal?

Graph ensembles

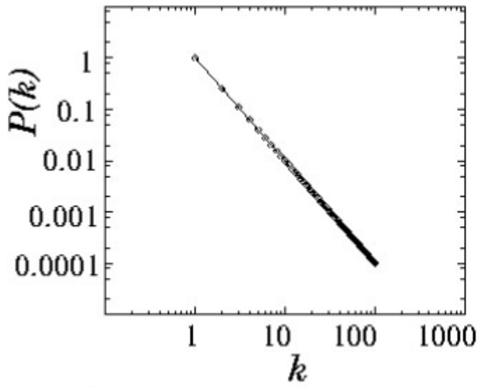
- Degree list: $k_1, k_2, \ldots, k_N$ extensive in N.
- Degree histogram: N_k number of nodes with a certain degree k.
- Degree distribution: $P(k)=N_k/N$



Regular graph



Poisson/Gaussian distribution



Power-law distribution

Statistical description

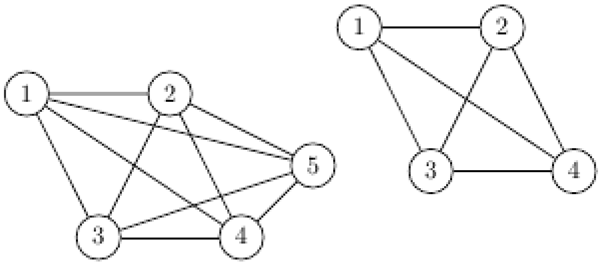
Average degree:

$$\langle k \rangle = \sum_k k P(k) = \frac{2|E|}{|V|} = 2D$$

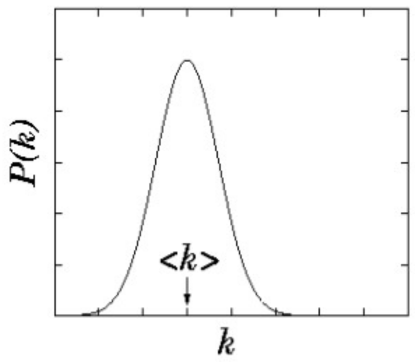
Variance:

$$\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2 = \sum_k (k - \langle k \rangle)^2 P(k)$$

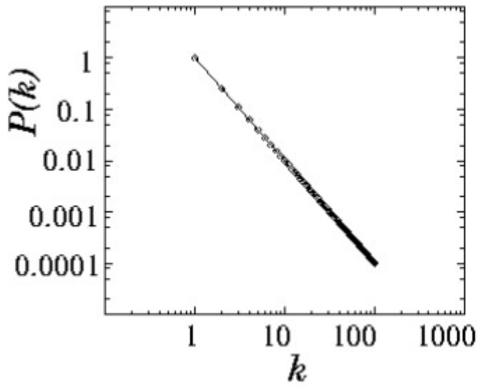
Heterogeneous graph $\frac{\sigma_k}{\langle k \rangle} \gg 1$



Regular graph



Poisson/Gaussian distribution



Power-law distribution

Random graphs

Random means:

- An algorithm to create graphs
- That is non-deterministic.

The algorithm define a **probability** over the set of all possibles graphs.

Random Regular graphs

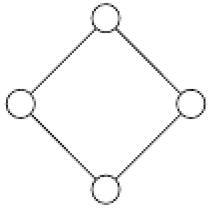
Random means:

- An algorithm to create graphs
- That is non-deterministic.

The algorithm define a **probability** over the set of all possibles graphs.

Example.

Random regular graph RRG(n=4,d=2), RRG(n=4,d=1)



Random Uncorrelated graphs

Random means:

- An algorithm to create graphs
- That is non-deterministic.

The algorithm define a **probability** over the set of all possibles graphs.

Remaining degree distribution: $$q_k = \frac{(k+1)p_{k+1}}{\sum_{j\geq 1} j p_j}$$

Uncorrelated networks:

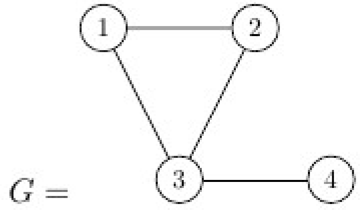
A graph where the remaining degree distribution of a node is independent of the degree of its neighbors

Example. Random **Poissonian** graph with mean connectivity

Practical Class

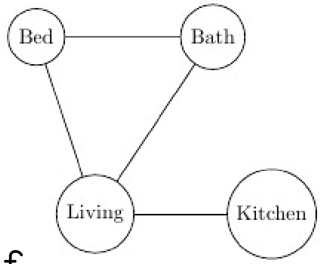
Implement this graph in Networkx:

NetworkX
Network Analysis in Python



Check documentation in:
https://networkx.org/
https://networkx.org/documentation/stable/tutorial.html
https://networkx.org/documentation/stable/reference/functions.html

Change nodes names to correpond to the drunken man problem:

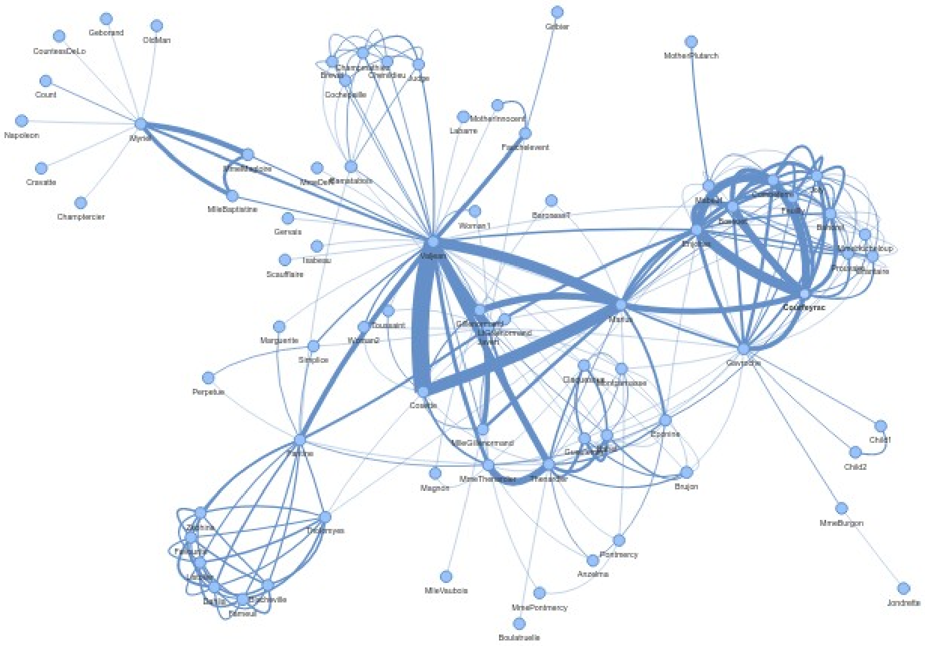


- Extract the **adjacency matrix.**
- For each node, compute: **eccentricity, clustering coefficient, betweenness.**
- Find the **Jordan Center** of this graph
- Check that **adjacency matrix powers** can be used to compute the number of pahts between nodes.

Homework

Study the properties of the Les Miserables graph

- Is this graph connected? Weighted? A tree?
- Fin degree distribution
- For each node, compute: **centrality, clustering coefficient, betweenness.**
- Which characters have the highest betweenness? Clustering?
- Which chacracters are the center (**Jordan Center**) of this graph?
- Plot the **assortativity** of this graph



你好

CENTRO DE **SISTEMAS COMPLEJOS**

Facultad de Física
Universidad de La Habana