Networks

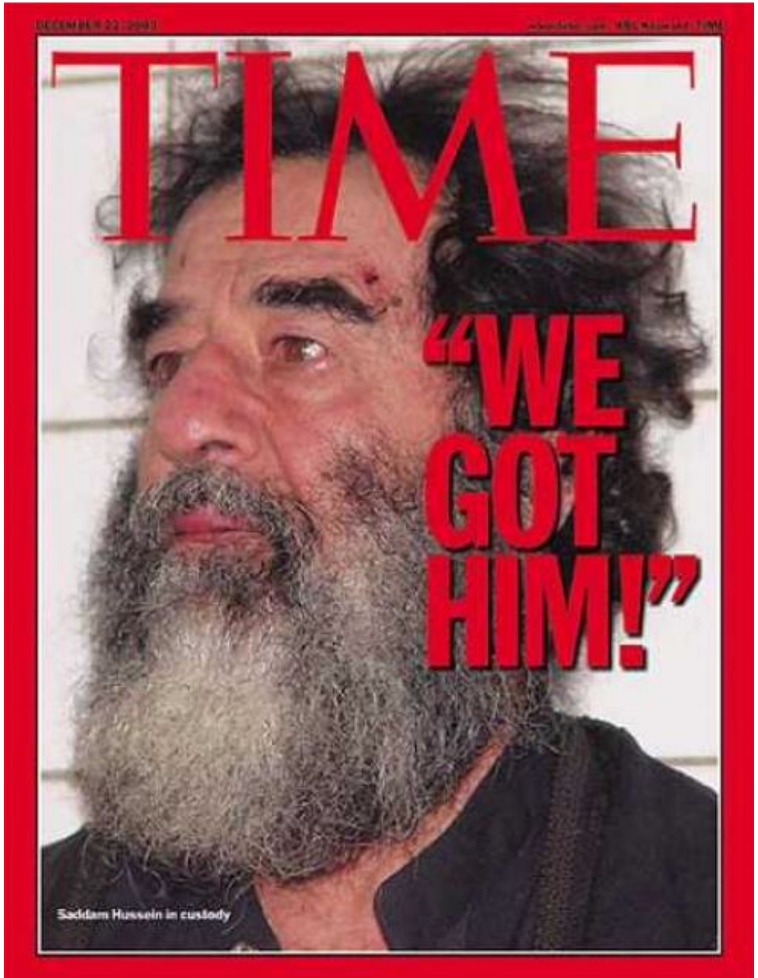
Introduction & Measures

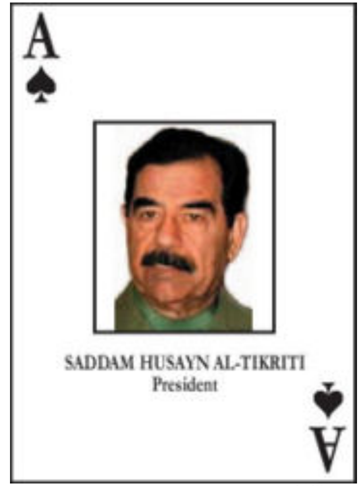
http://tiny.cc/wsc2016

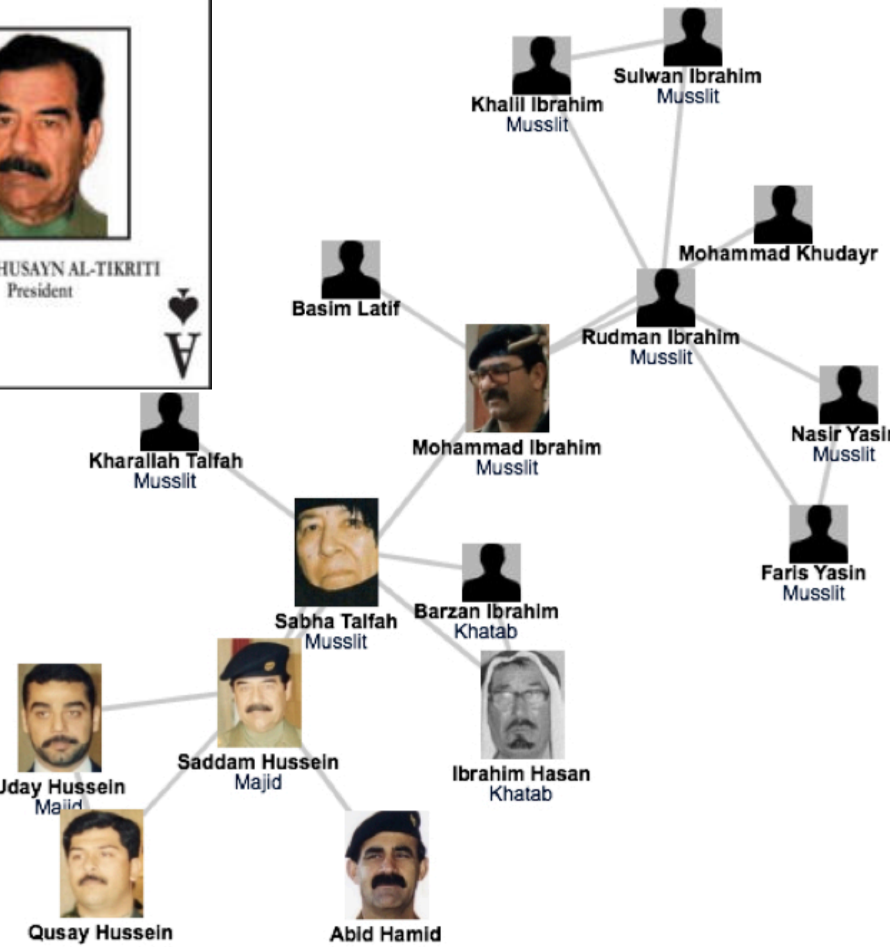
Michael Lees
m.h.lees@uva.nl

Slides adapted from Lázló Barabási
http://www.barabasilab.com/

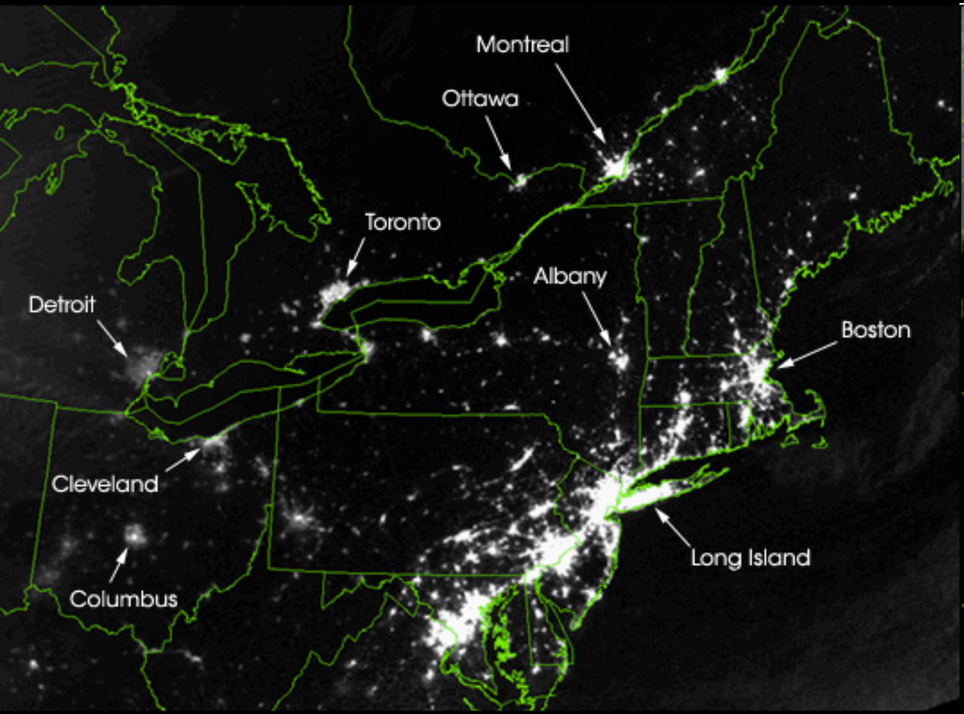
A SIMPLE STORY (1)　The fate of Saddam and network science



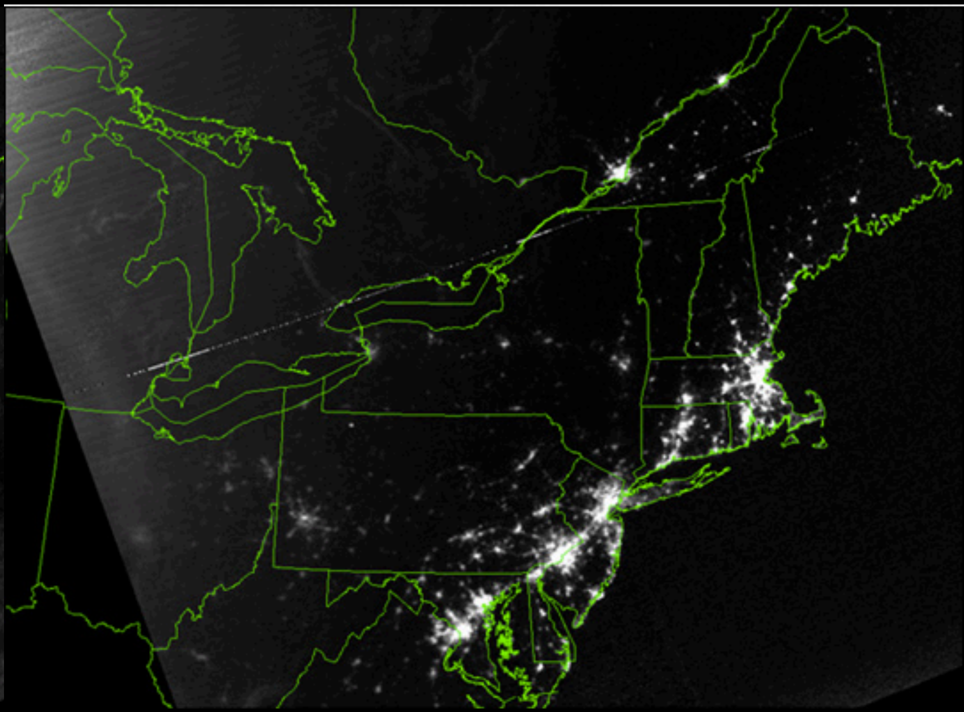




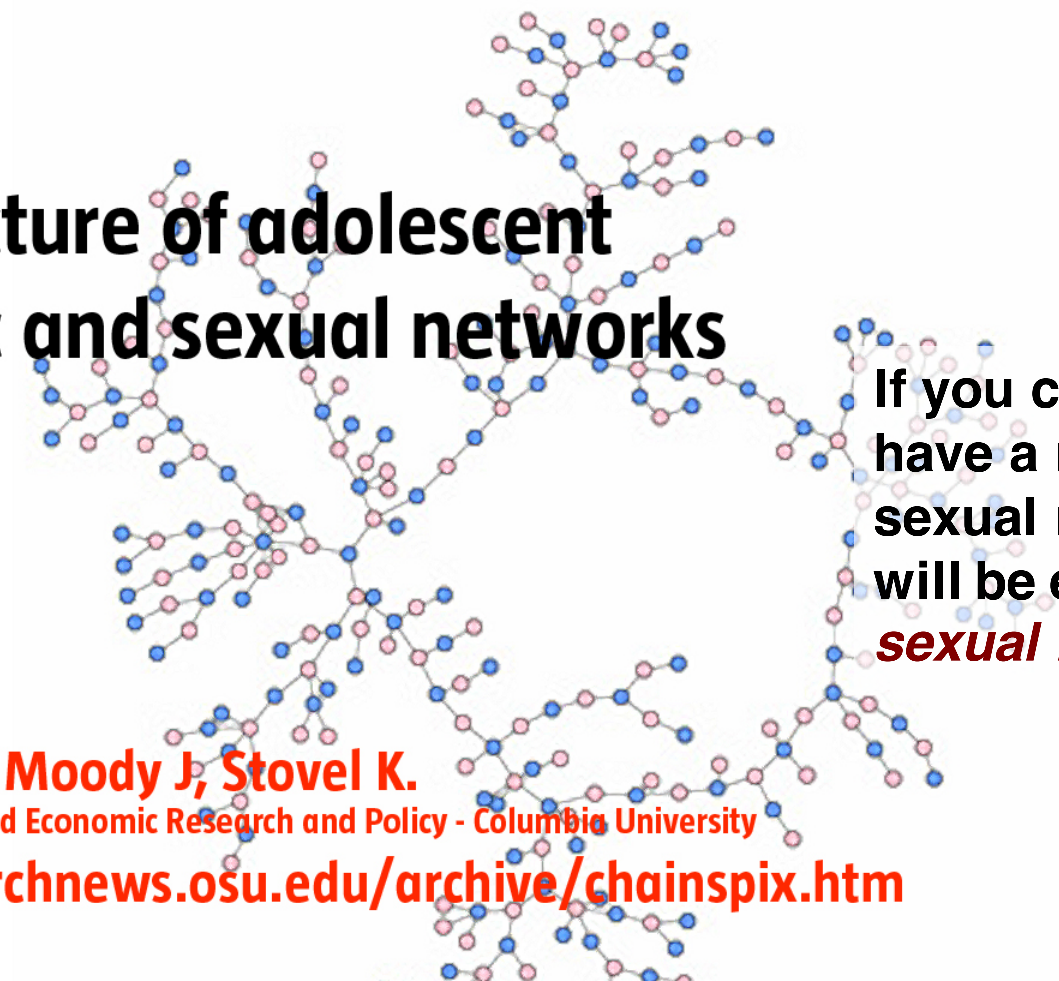
A SIMPLE STORY (2): August 15, 2003 blackout.



August 14, 2003: 9:29pm EDT
20 hours before



August 15, 2003: 9:14pm EDT
7 hours after



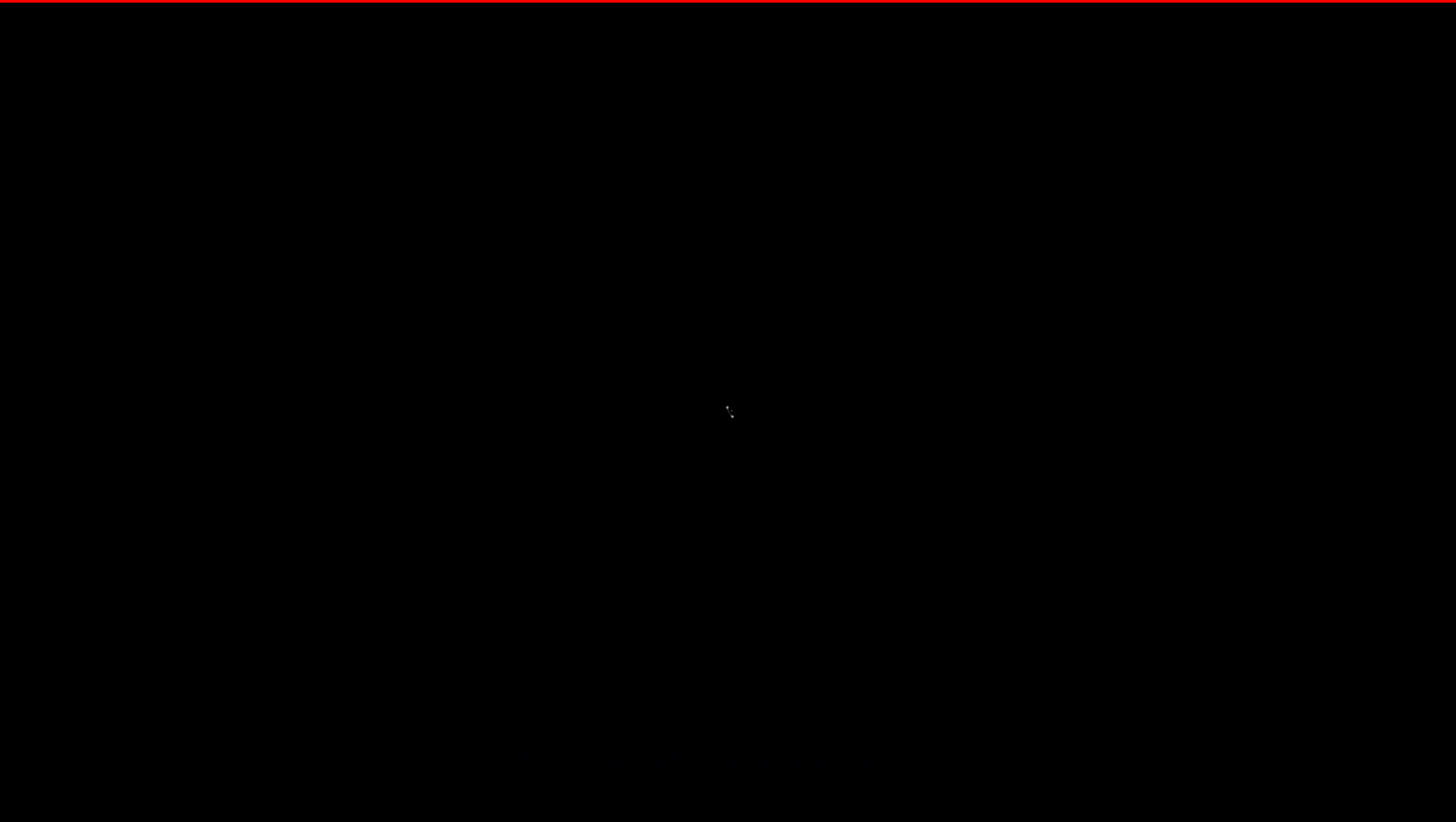
The structure of adolescent romantic and sexual networks

If you connect those that have a romantic and sexual relationship, you will be exploring the ***sexual networks***.

Bearman PS, Moody J, Stovel K.
Institute for Social and Economic Research and Policy - Columbia University

http://researchnews.osu.edu/archive/chainspix.htm

Connections



Predicting the H1N1 pandemic

NETWORKS AT THE HEART OF COMPLEX SYSTEMS

COMPLEX SYSTEMS

Complex

[adj., v. kuh m-pleks, kom-pleks; n. kom-pleks]
–adjective

1.
composed of many interconnected parts; compound; composite: a complex highway system.

2.
characterized by a very complicated or **involved arrangement** of parts, units, etc.: complex machinery.

3.
so complicated or intricate as to be hard to understand or deal with: a complex problem.

Source: Dictionary.com

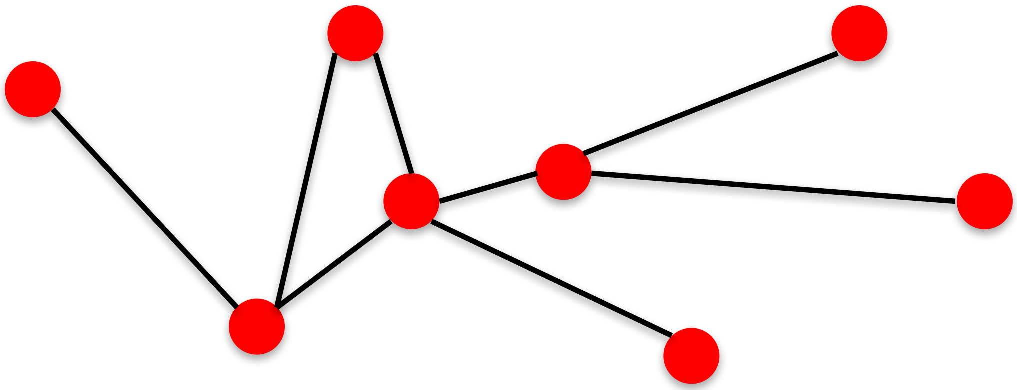
Complexity, a **scientific theory** which asserts that some systems display behavioral phenomena that are completely inexplicable by any conventional analysis of the systems' constituent parts. These phenomena, commonly referred to as emergent behaviour, seem to occur in many complex systems involving living organisms, such as a stock market or the human brain.

Source: John L. Casti, Encyclopædia Britannica

Complexity

Behind each complex system there is a **network**, that defines the interactions between the component.

COMPONENTS OF A COMPLEX SYSTEM

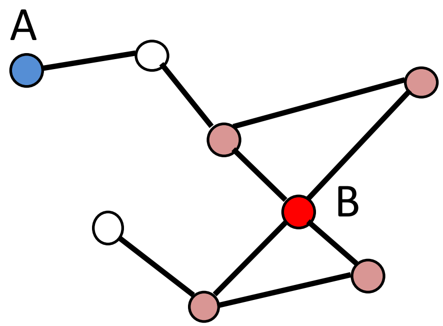


- **components**: nodes, vertices — N
- **interactions**: links, edges — L
- **system**: network, graph — (N,L)

Degree, Average Degree and Degree Distribution

NODE DEGREES

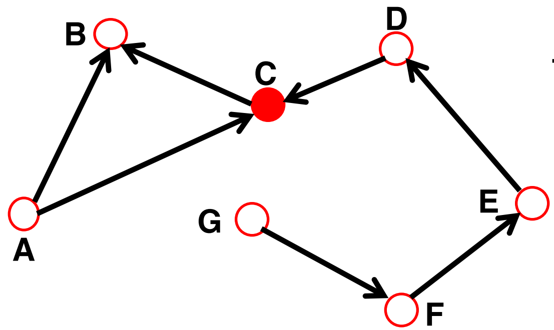
Undirected



Node degree: the number of links connected to the node.

$$k_A = 1 \qquad k_B = 4$$

Directed



In *directed networks* we can define an in-degree and out-degree.

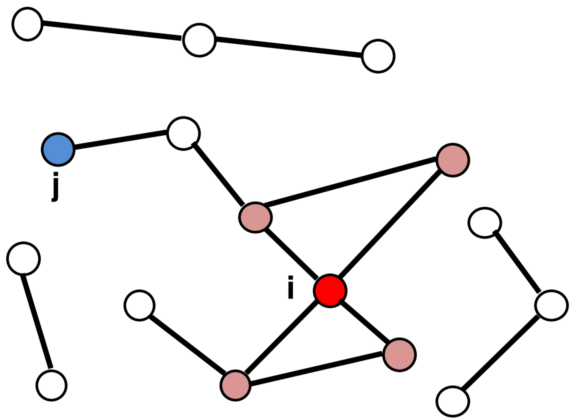
The (total) degree is the sum of in- and out-degree.

$$k_C^{in} = 2 \qquad k_C^{out} = 1 \qquad k_C = 3$$

Source: a node with $k^{in} = 0$; Sink: a node with $k^{out} = 0$.

AVERAGE DEGREE

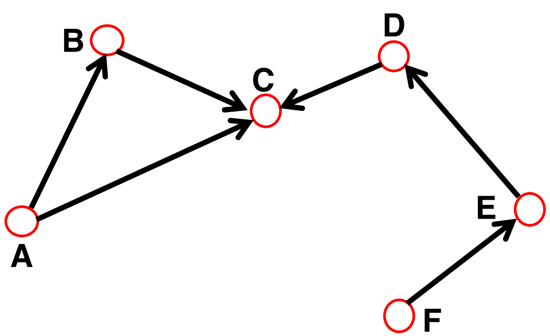
Undirected



$$\langle k \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} k_i \qquad \langle k \rangle \equiv \frac{2L}{N}$$

N – the number of nodes in the graph

Directed



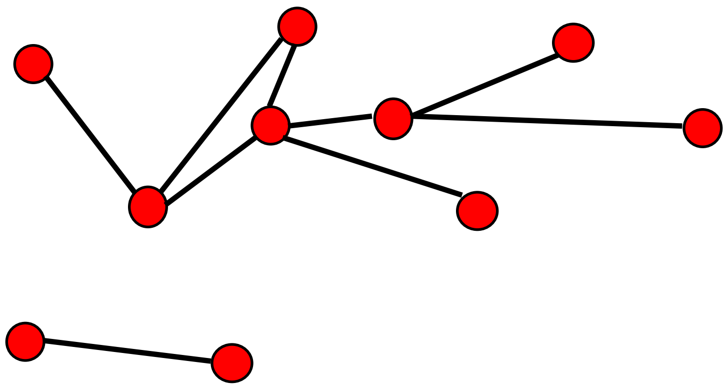
$$\langle k^{in} \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} k_i^{in} \qquad \langle k^{out} \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} k_i^{out}$$

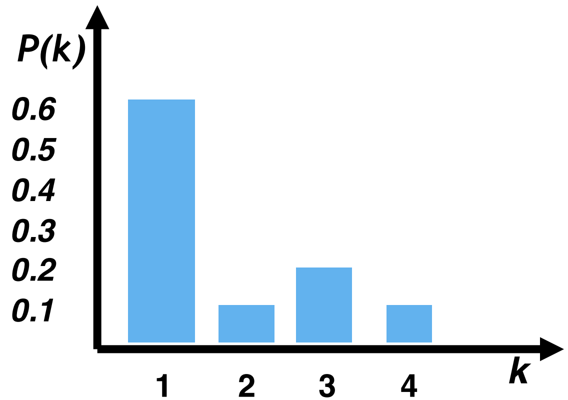
$$\langle k^{in} \rangle = \langle k^{out} \rangle$$

$$\langle k \rangle \equiv \frac{L}{N}$$

Degree distribution

P(k): probability that a randomly chosen node has degree *k*





$$P(k) = \frac{N_k}{N}$$

N_k *Number of nodes with degree k*

DEGREE DISTRIBUTION

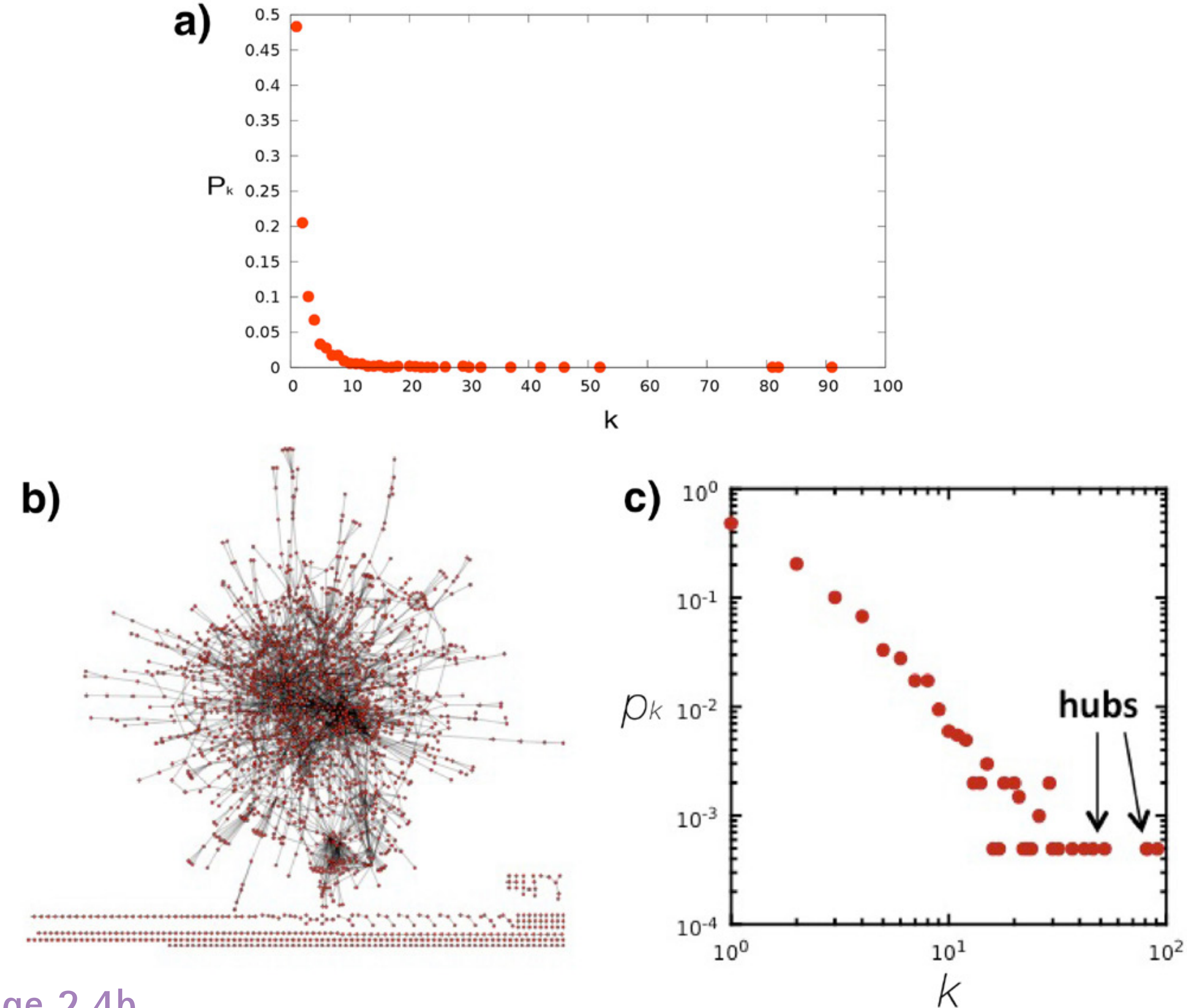
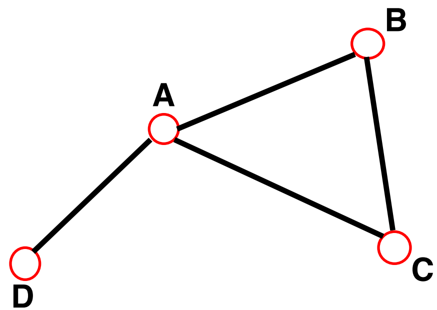


Image 2.4b

Average Path Length and Diameter

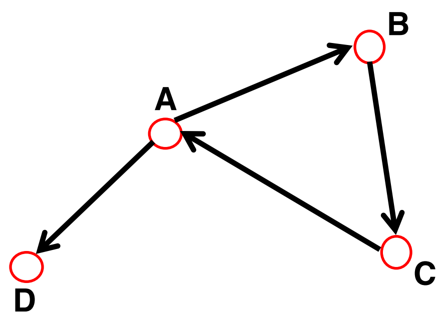
DISTANCE IN A GRAPH

Shortest Path, Geodesic Path



The *distance (shortest path, geodesic path)* between two nodes is defined as the number of edges along the shortest path connecting them.

*If the two nodes are disconnected, the distance is infinity.



In directed graphs each path needs to follow the direction of the arrows.

Thus in a digraph the distance from node A to B (on an AB path) is generally different from the distance from node B to A (on a BCA path).

NETWORK DIAMETER AND AVERAGE DISTANCE

Diameter: $\boldsymbol{d_{max}}$ the maximum distance between any pair of nodes in the graph.

Average path length/distance, <d>, for a directed connected graph:

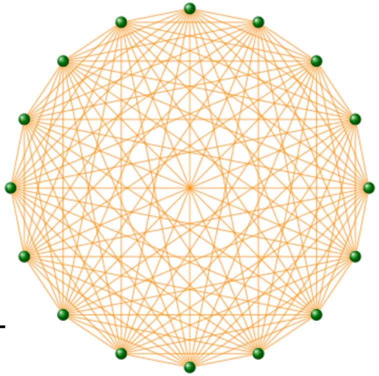
$$\langle d \rangle \equiv \frac{1}{2L_{max}} \sum_{i,j \neq i} d_{ij}$$

where d_{ij} is the distance from node i to node j

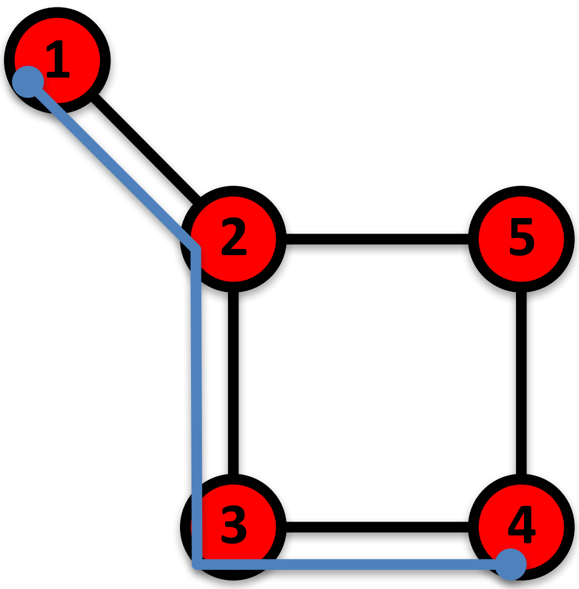
In an *undirected graph* $d_{ij} = d_{ji}$*, so* we only need to count them once:

$$\langle d \rangle \equiv \frac{1}{L_{max}} \sum_{i,j > i} d_{ij}$$

$$L_{max} = \frac{N(N-1)}{2}$$



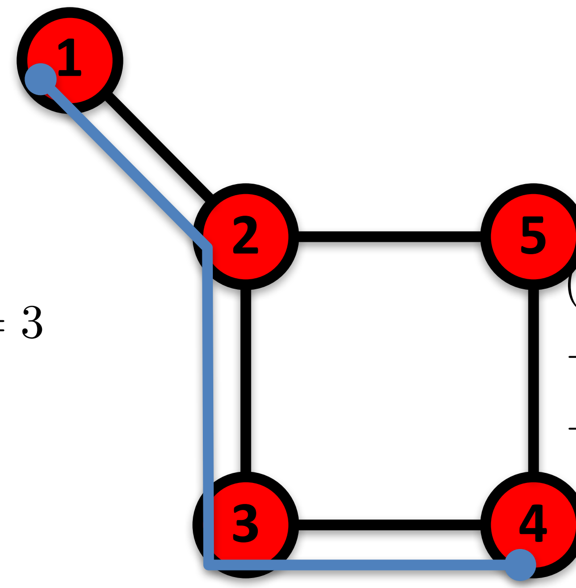
Diameter



$$l_{1\to4} = 3$$

The longest shortest path in a graph

Average Path Length



$$(l_{1\to2} + l_{1\to3} + l_{1\to4} + \\ + l_{1\to5} + l_{2\to3} + l_{2\to4} + \\ + l_{2\to5} + l_{3\to4} + l_{3\to5} + \\ + l_{4\to5})/10 = 1.6$$

The average of the shortest paths for all pairs of nodes.

Clustering coefficient

CLUSTERING COEFFICIENT

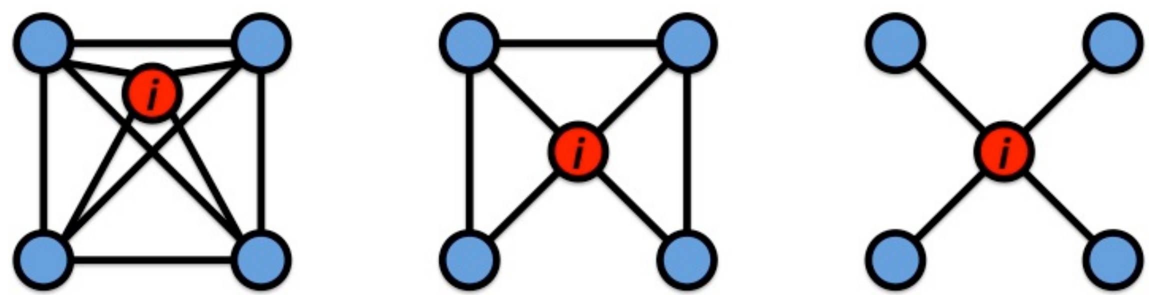
* Clustering coefficient:

what fraction of your neighbors are connected?

* Node i with degree k_i

* C_i in [0,1]

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



$C_i = 1$ $C_i = 1/2$ $C_i = 0$

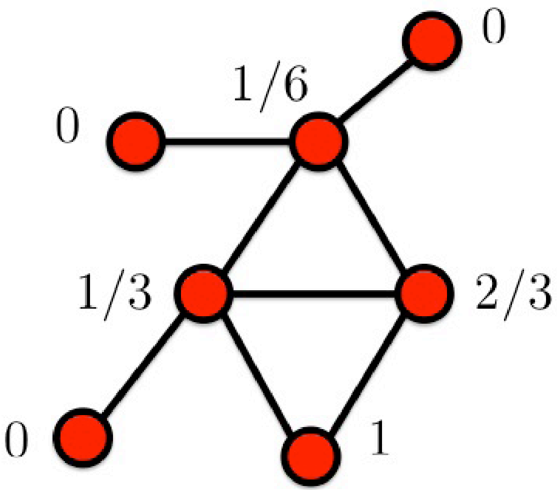
Watts & Strogatz, Nature 1998.

Clustering coefficient:

what fraction of your neighbors are connected?

- Node i with degree k_i
- C_i in [0,1]

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



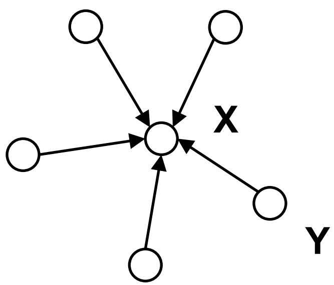
$$\langle C \rangle = \frac{13}{42} \approx 0.310$$

$$C = \frac{3}{8} = 0.375$$

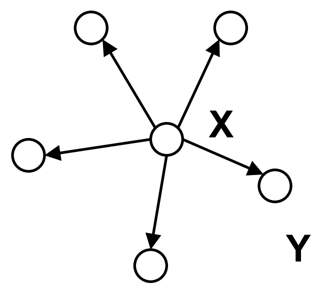
Watts & Strogatz, Nature 1998.

Centrality Measures

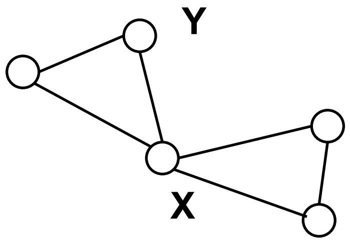
Centrality Measures



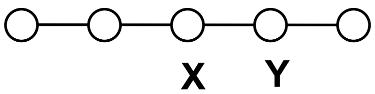
indegree



outdegree



betweenness

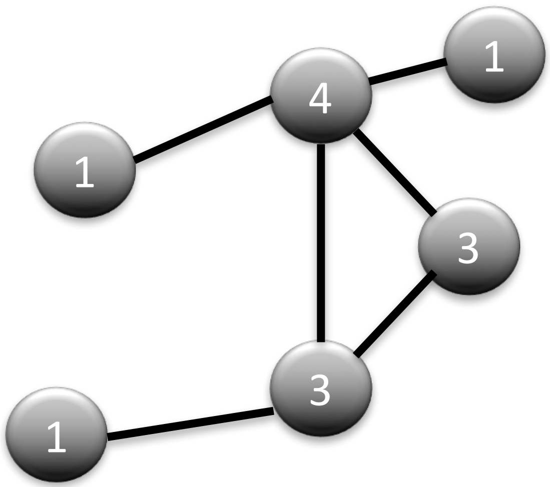


closeness

Degree Centrality

Intuition: How many friends do you have, how many contacts do you make?

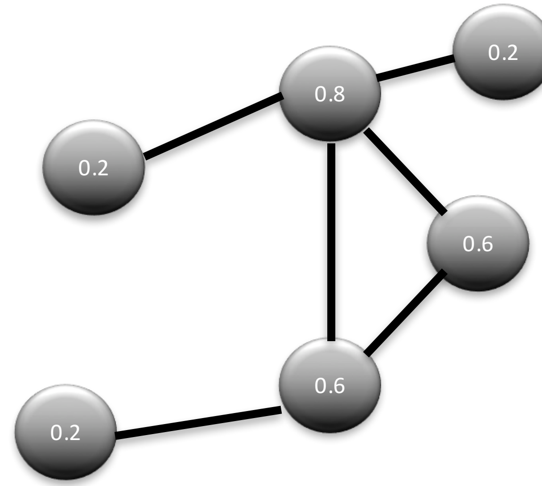
<u>NON-NORMALISED VERSION</u>



Node *i* Degree $C_D(i) = k_i$

Network Ave. Degree $\langle C_D \rangle = \sum_{i=1}^{N} k_i$

<u>NORMALISED VERSION</u>



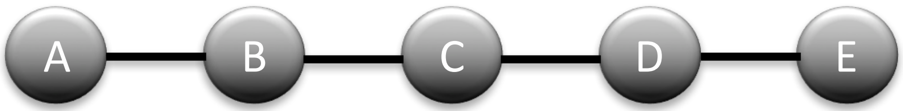
Node *i* Degree $C_D(i) = \dfrac{k_i}{N-1}$

Network Ave. Degree $\langle C_D \rangle = \sum_{i=1}^{N} \dfrac{k_i}{N-1}$

Betweenness Centrality

Intuition: How many pairs of your friends would have to communicate through you?

Calculate shortest path between all pairs of nodes…
What fraction of those paths does *i* lie on (not as end or start point)



$$C_B(i) = \sum_{h \neq i \neq j} \frac{\sigma_{hj}(i)}{\sigma_{hj}}$$

$\sigma_{hj}(i)$ = number of shortest paths *i* lies on between *h* and *j*

σ_{hj} = total number of shortest paths between *h* and *j*

Betweenness Centrality

All shortest paths?

Undirected network so shortest path A→E = E→A

A→B [A,B], A→C [A,B,C], A→D [A,B,C,D], A→E [A,B,C,D,E]
B→C [B,C], B→D [B,C,D], B→E [B,C,D,E]
C→D [C,D], C→E [C,D,E],
D→E [D,E]

$$C_B(i) = \sum_{h \neq i \neq j} \frac{\sigma_{hj}(i)}{\sigma_{hj}}$$

$C_B(A) = 0$
$C_B(B) = 3$
$C_B(C) = 4$
$C_B(D) = 3$
$C_B(E) = 0$

Normalized Version

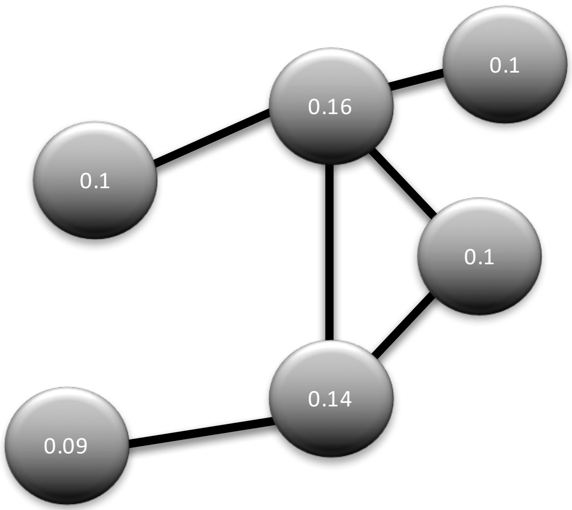
$$C'_B(i) = \frac{C_B(i)}{\frac{(N-1)(N-2)}{2}}$$

Normalised by total number of other pairs (without *i*)

Closeness Centrality

Intuition: How far am I away from the rest of the network

NON-NORMALISED VERSION



$C_C(A) = (1+2+2+2+3)^{-1}$

$C_C(B) = (1+1+1+1+2)^{-1}$

$C_C(C) = (2+1+2+2+3)^{-1}$

$C_C(D) = (2+1+2+1+2)^{-1}$

$C_C(E) = (2+1+2+1+1)^{-1}$

$C_C(F) = (3+2+3+2+1)^{-1}$

Node *i* Closeness $$C_C(i) = \frac{1}{\sum_{j=1}^{N} d(i,j)}$$

$d(i,j)$ Distance from *i* to *j*

NORMALISED VERSION

$$C'_C(i) = \frac{C_C(i)}{N-1}$$