Network theory: 2

Liubov Tupikina, Marc Santolini (CRI)

Today

1. Network science: new concepts
2. Notebook Hands-on: Python and Gephi

 Break

3. Questions & Answers
4. Projects discussions and Reversed classrooms

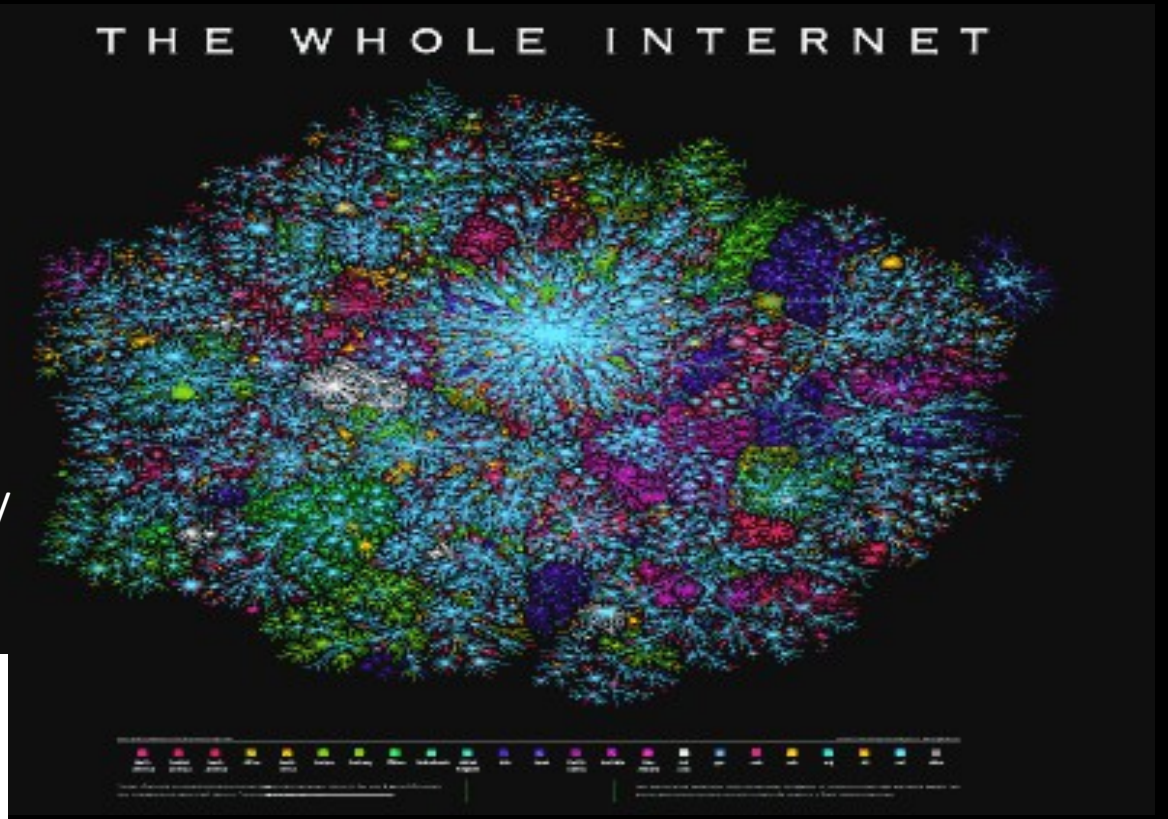
If you have questions?

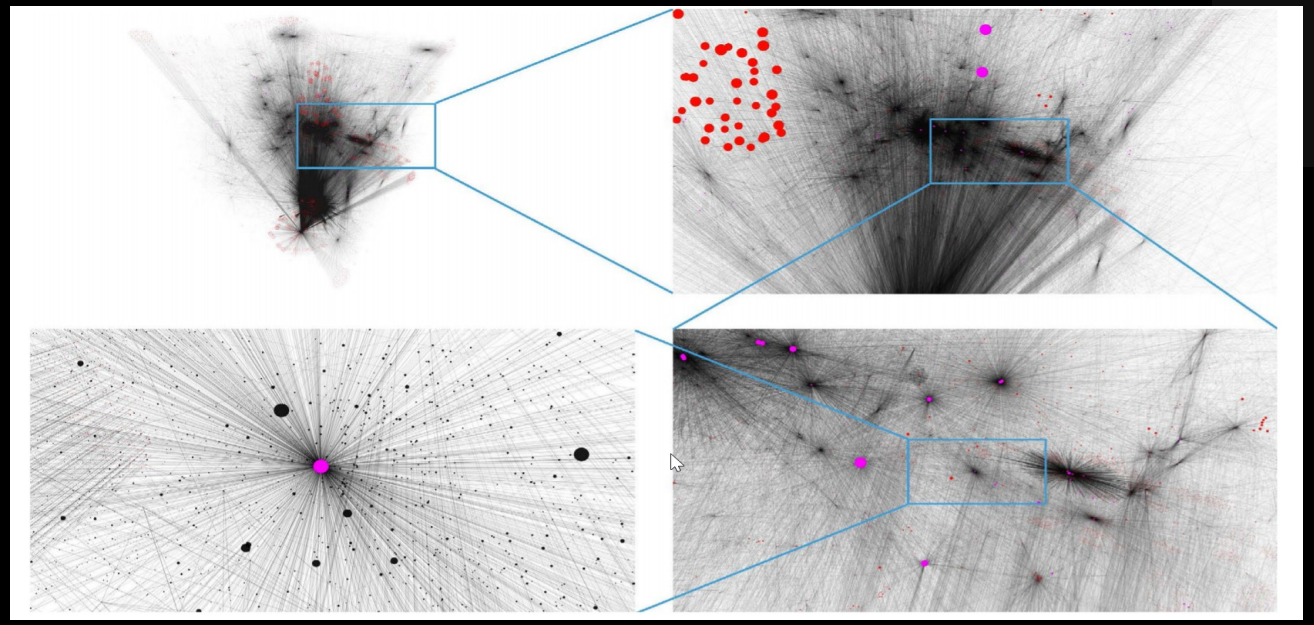
http://networksciencebook.com/

https://www.barabasilab.com

http://networkrepository.com/graph-vis.php

http://www.complexity-explorables.org/explorables/neighbors/

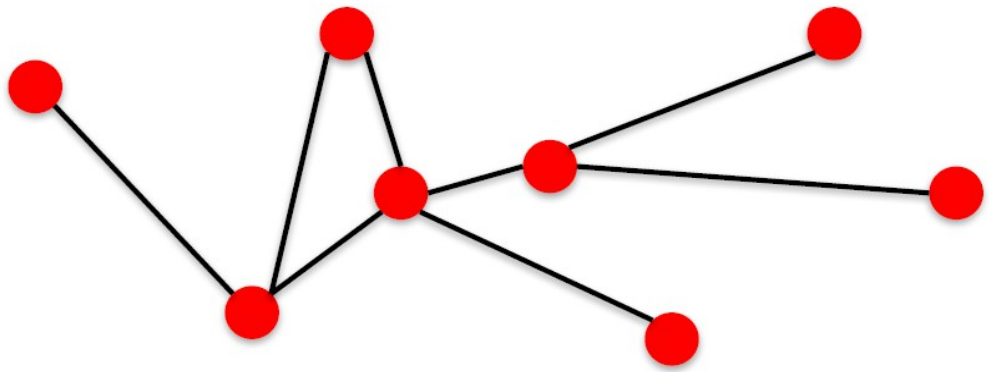




https://snap.stanford.edu/data

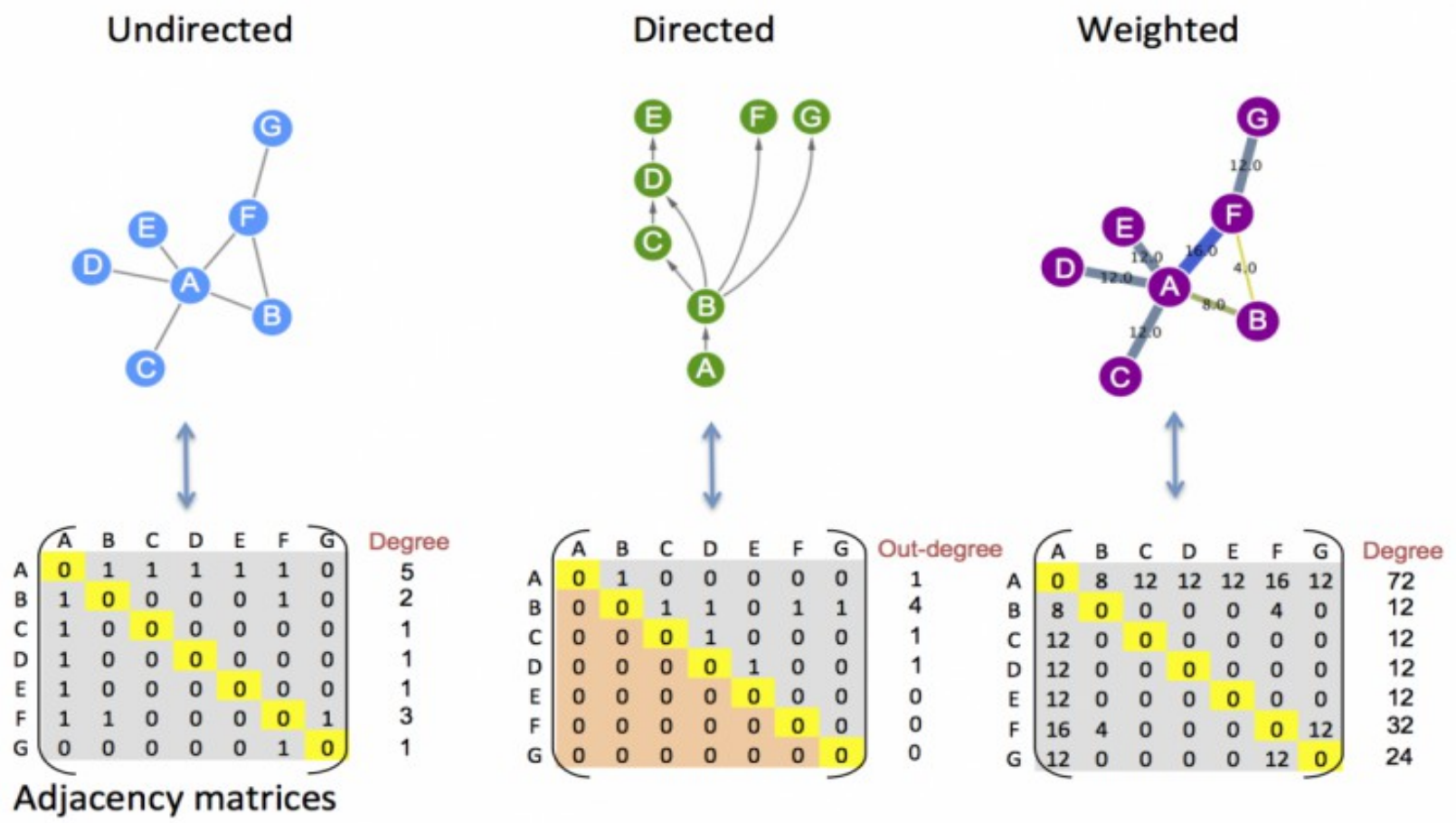
https://sites.google.com/a/binghamton.edu/netscied/teaching-learning/network-concepts

What is the network itself?



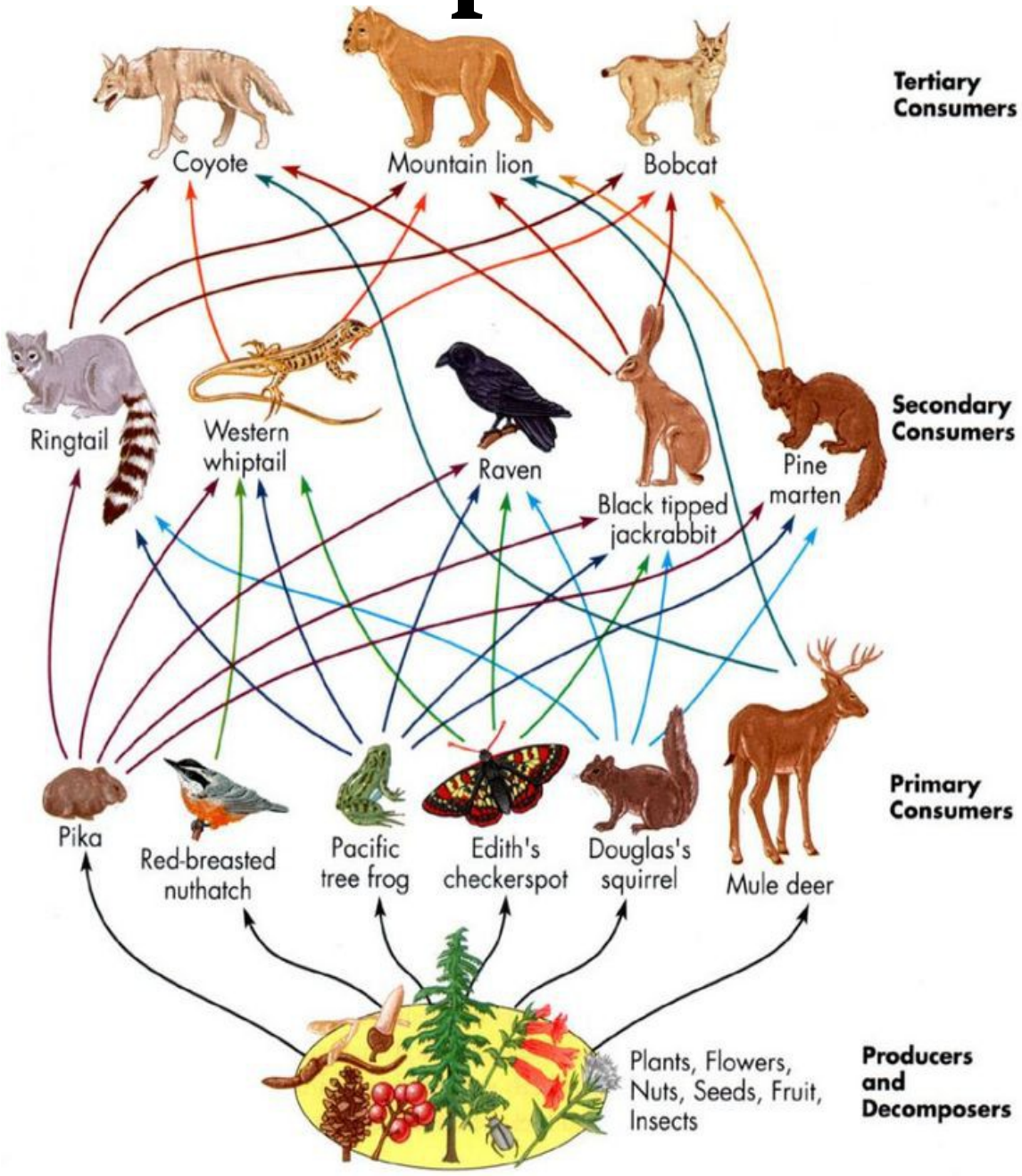
- **components**: nodes, vertices — N
- **interactions**: links, edges — L
- **system**: network, graph — (N,L)

How to describe a network?



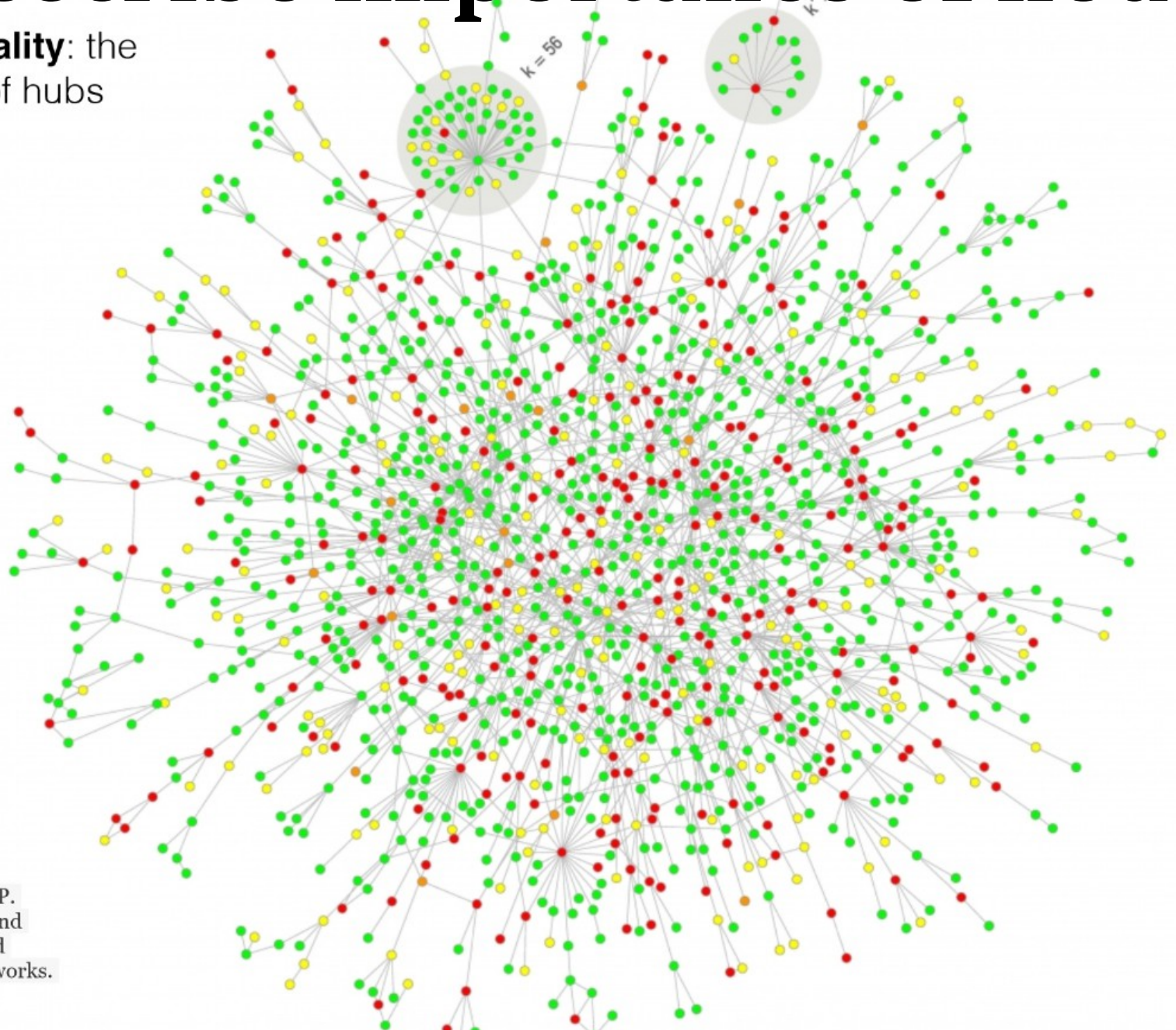
Barabasi book on network science

How to describe importance of nodes?



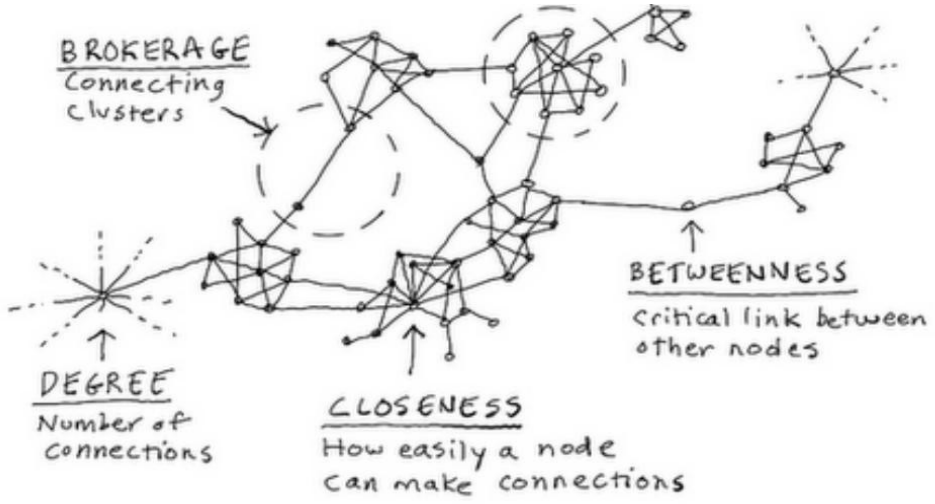
How to describe importance of nodes?

centrality-lethality: the importance of hubs



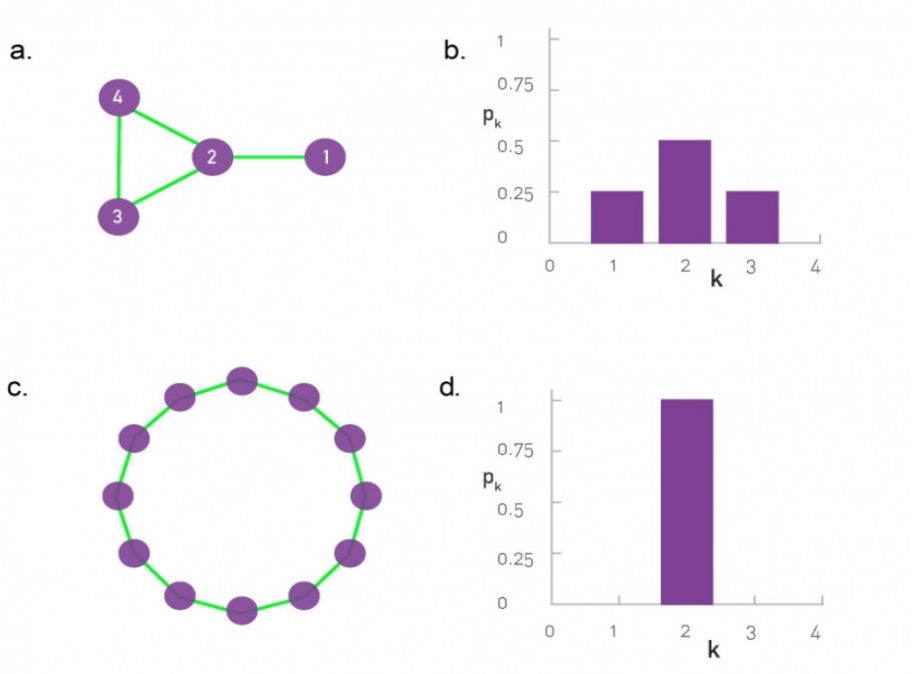
H. Jeong, B. Tombor, S. P. Mason, A.-L. Barabási, and Z.N. Oltvai. Lethality and centrality in protein networks. Nature 411: 41-42, 2001.

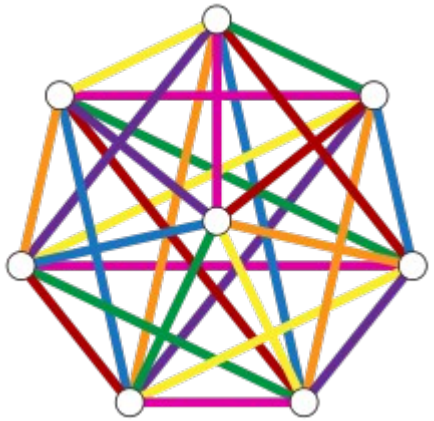
Network measures:
How to describe a network?

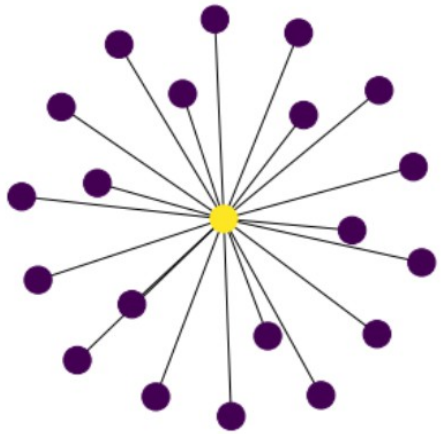


How to describe a network?

Degree of node A is the number of nodes adjacent with the node A.
Plotting degree distribution for a network can give important insight on its properties.





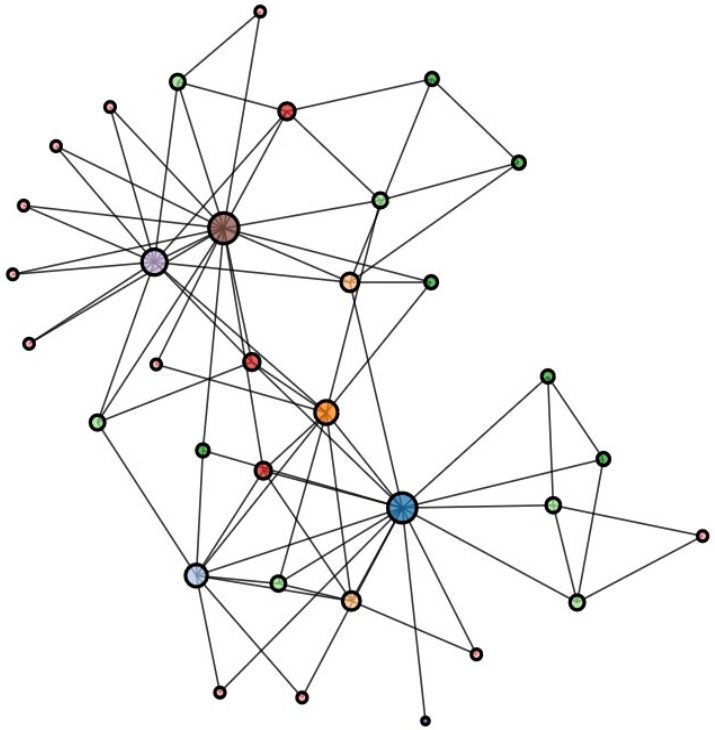


?

Barabasi book on network science

How to describe a real network?

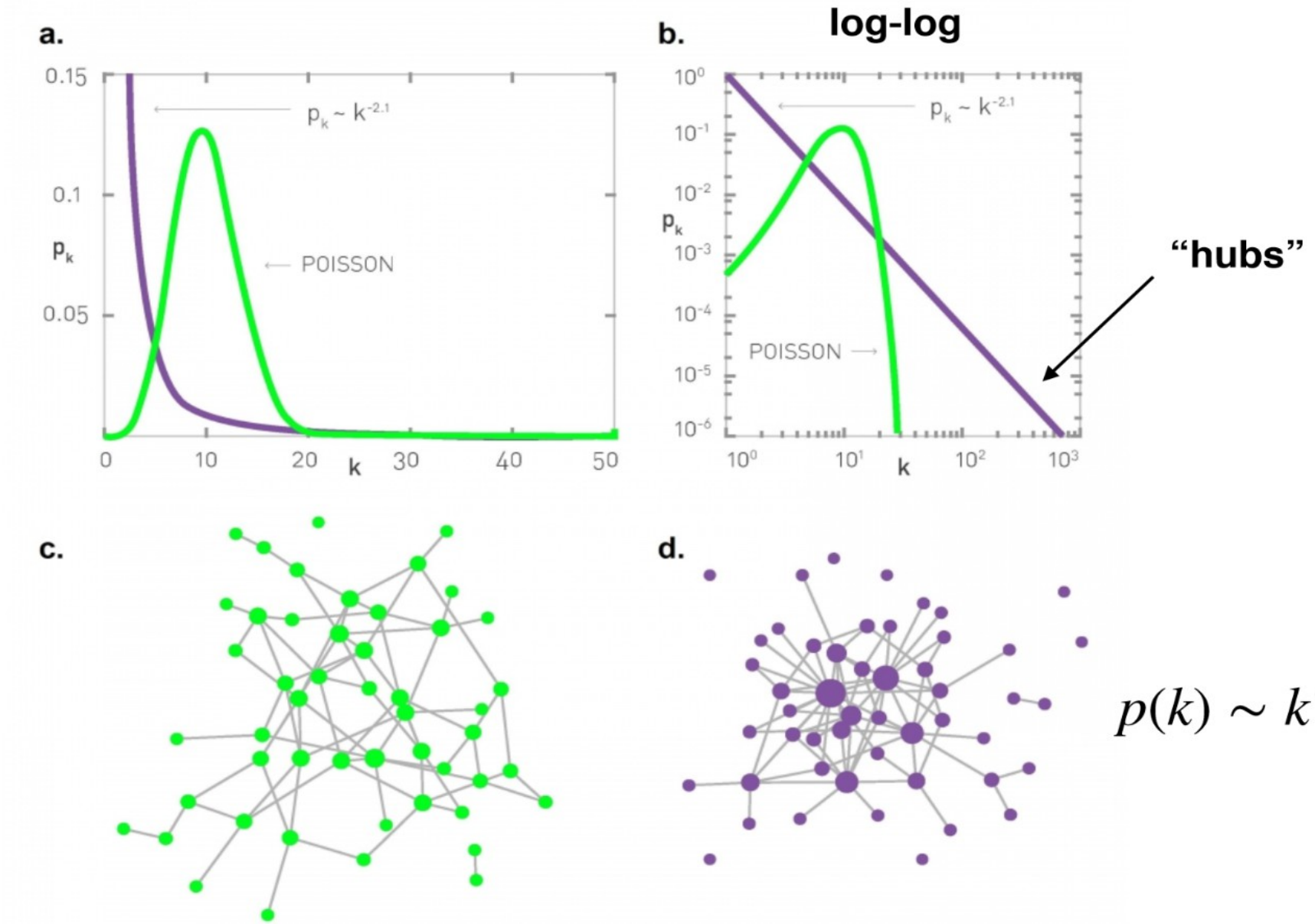
What is degree distribution of larger network? (Liuba -hands on session)
How to visualise a network in a nice way? (Marc - visualisation)



http://networkrepository.com/soc-karate.php

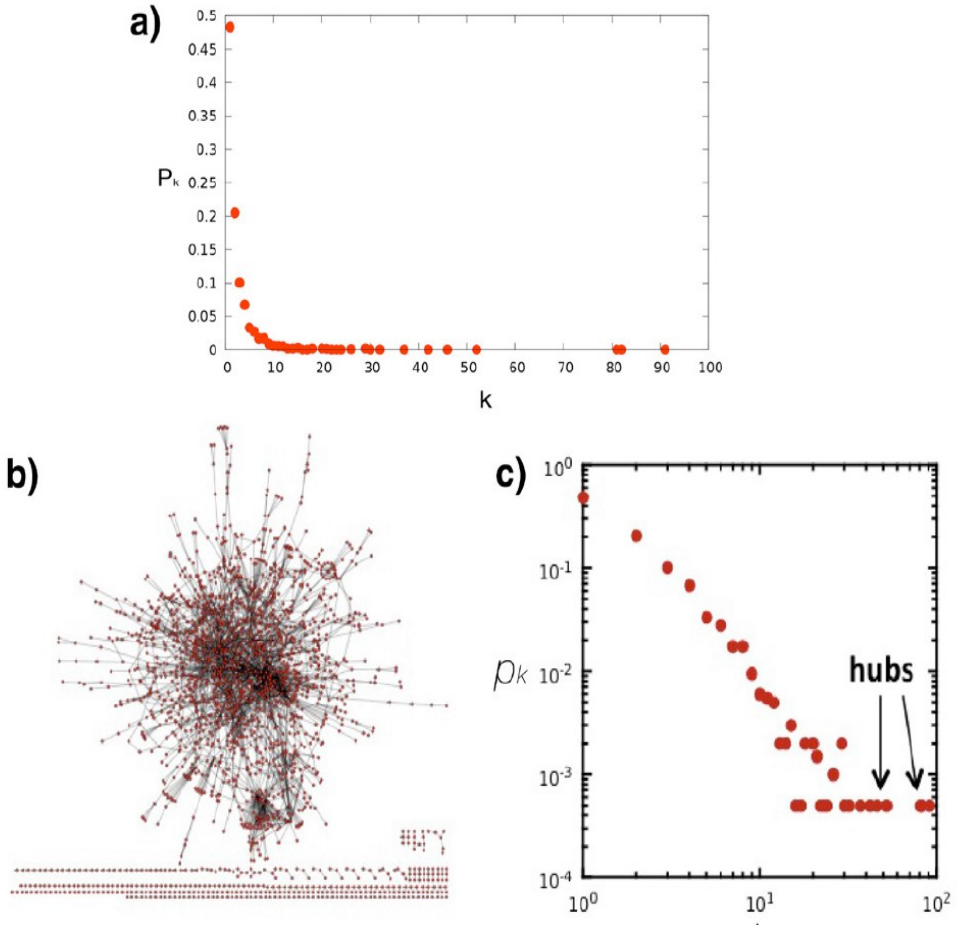
Network Data Statistics	
Nodes	34
Edges	78
Density	0.139037
Maximum degree	17
Minimum degree	1
Average degree	4
Assortativity	-0.475613
Number of triangles	135
Average number of triangles	3
Maximum number of triangles	18
Average clustering coefficient	0.570638
Fraction of closed triangles	0.255682
Maximum k-core	5
Lower bound of Maximum Clique	5

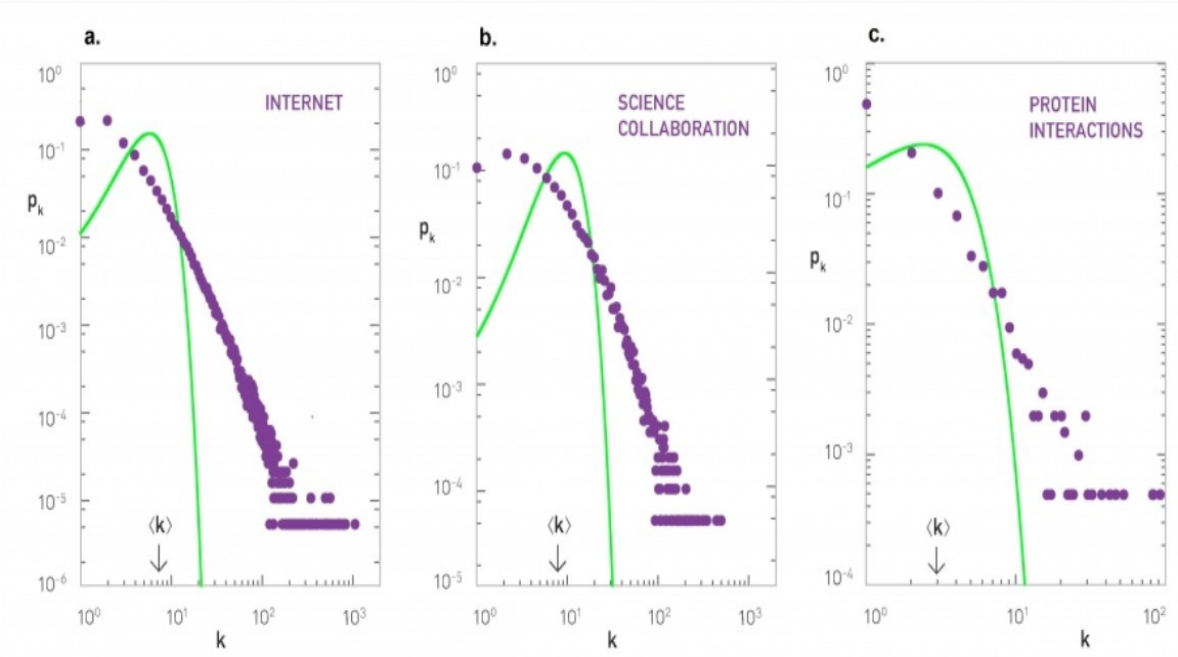
Scale-free networks (lecture 1)



$$p(k) \sim k^{-\gamma}$$

Scale-free networks (lecture 1)

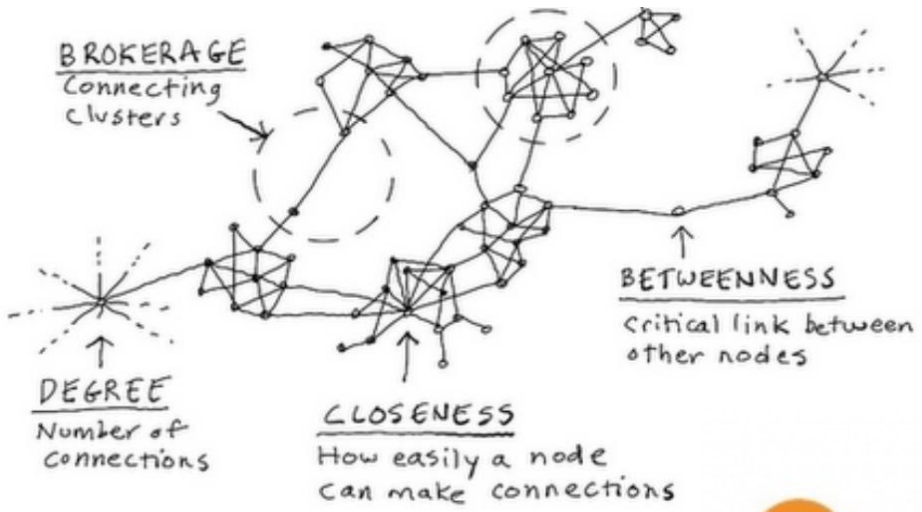


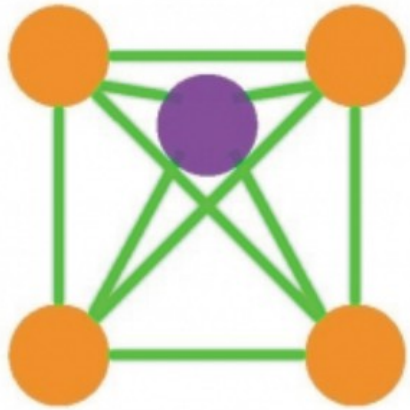


"Scale free networks can be found in many places"... at the same time
"Be careful with power-laws"

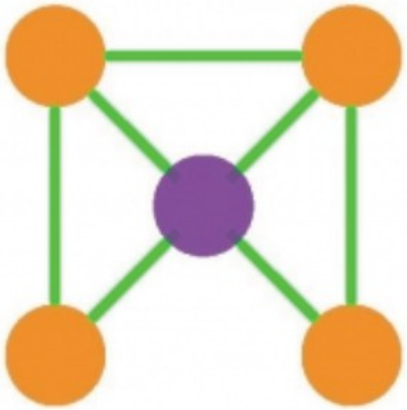
Cytoscape software alternative to Gephi for biological networks
Barabasi book

Network measures: clustering

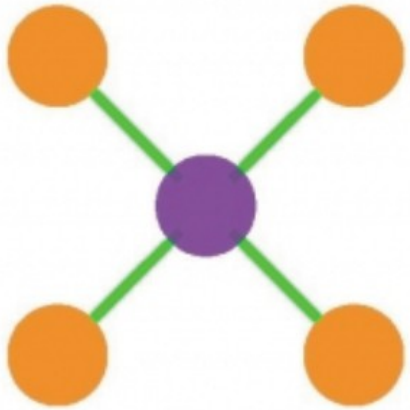




$C_i=1$

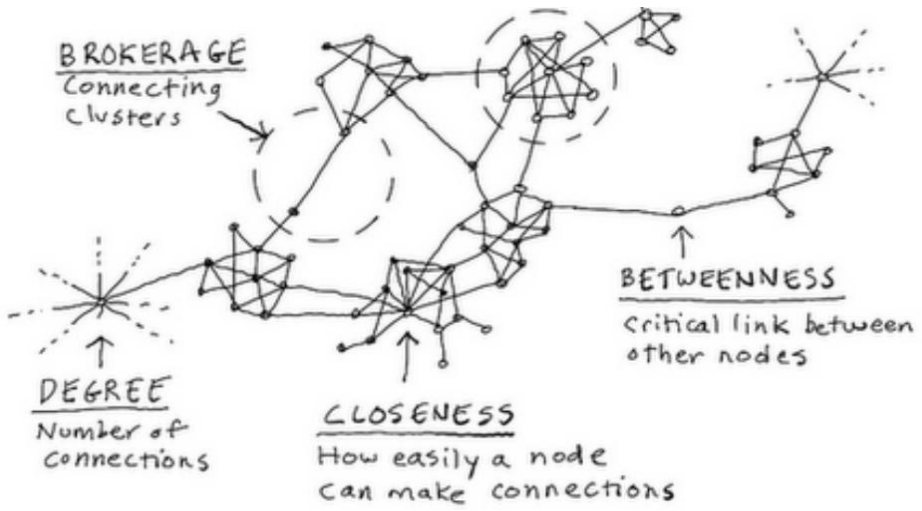


$C_i=1/2$



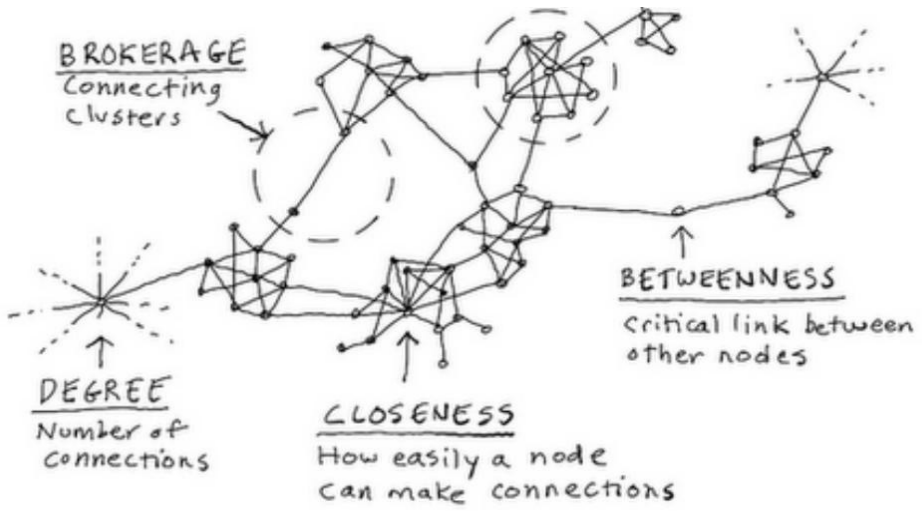
$C_i=0$

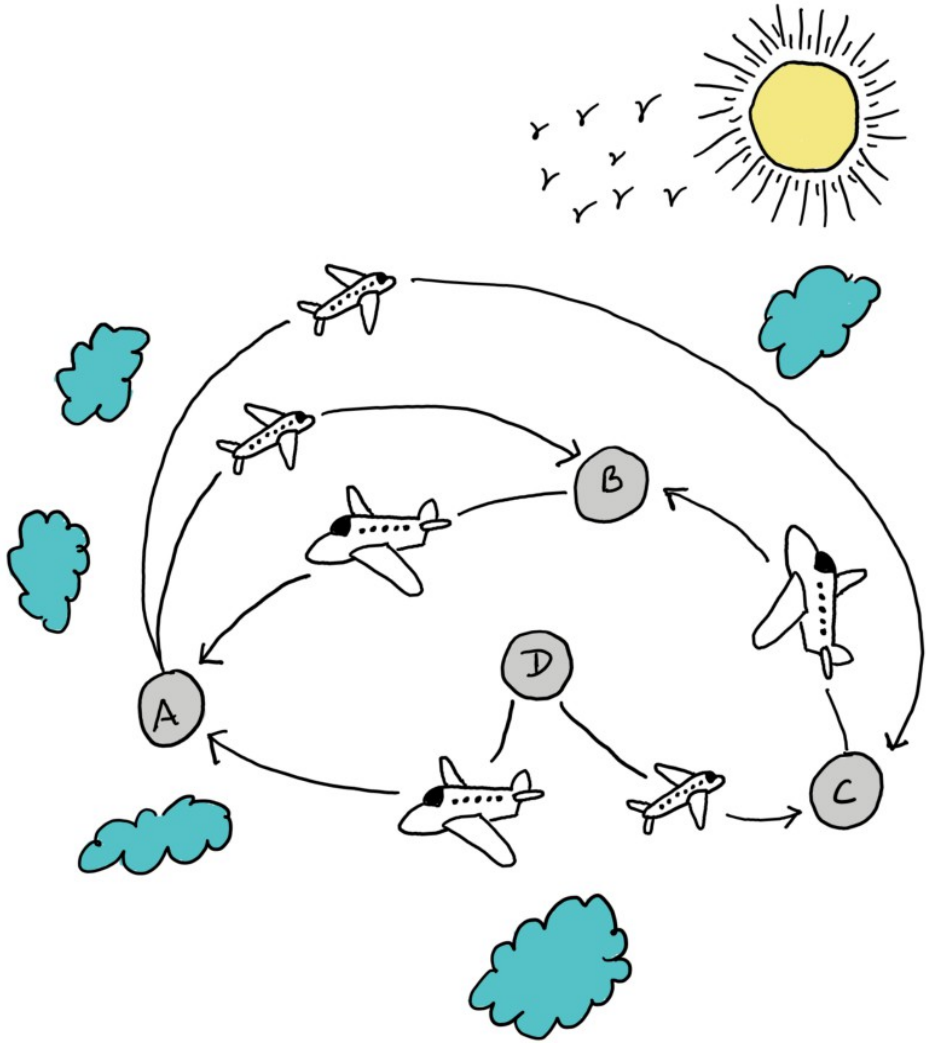
Network measures: clustering



$$C_i = \frac{|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}$$

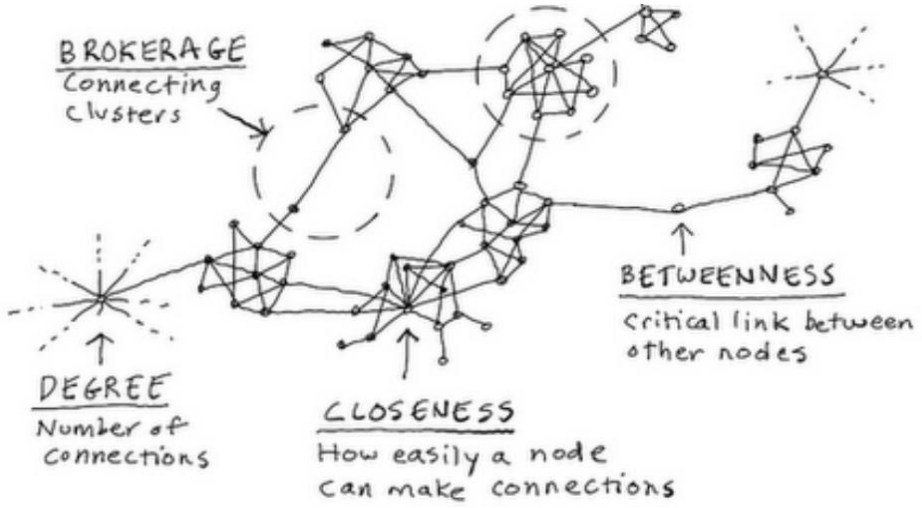
Network measures: betweenness centrality





Additional notions:
Paths in a network
Shortest paths
Random walks on a network

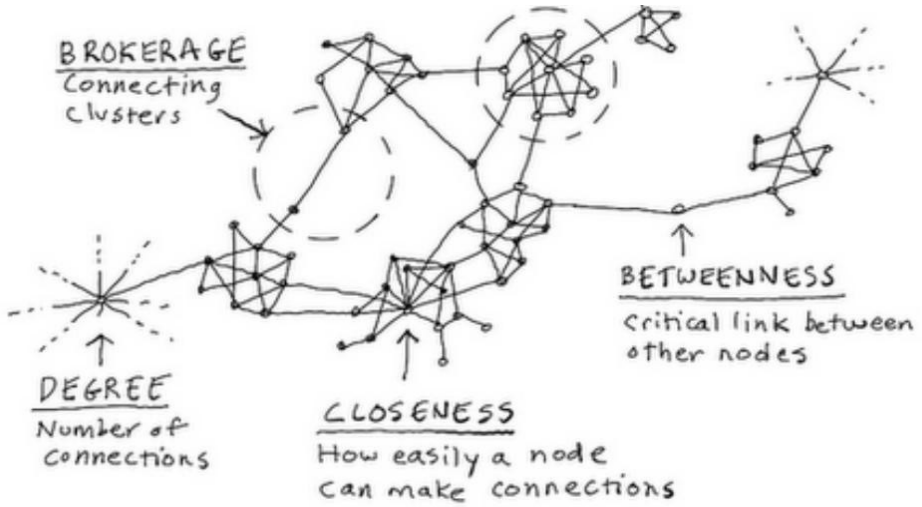
Network measures: betweenness centrality



$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

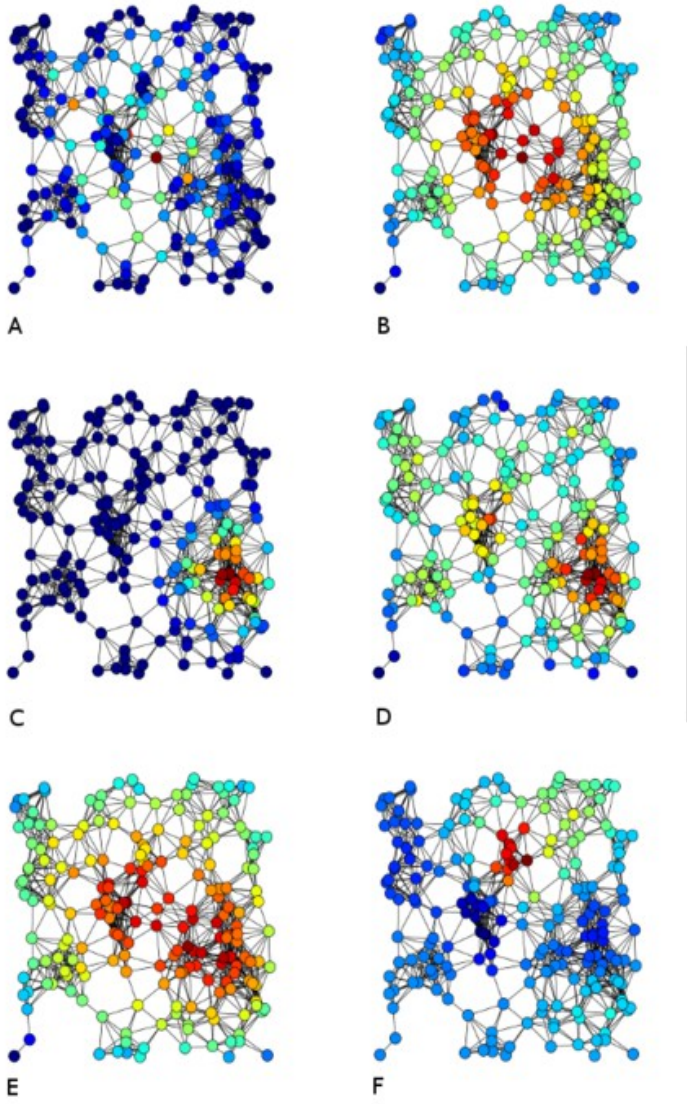
Additional notions:
Paths in a network
Shortest paths
Random walks on a network

Network measures: closseness centrality

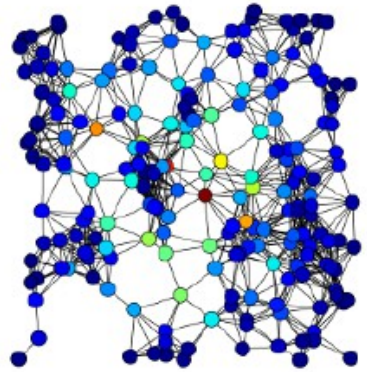


Main idea:
the more central a node is, the closer it is to all other nodes

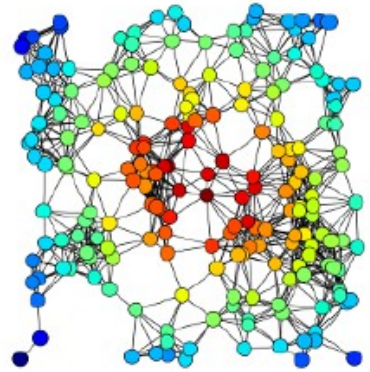
Network measures
Which is what?



Betweenness centrality



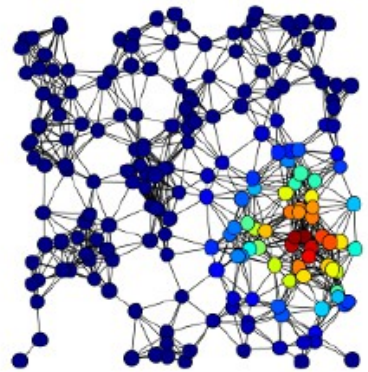
A



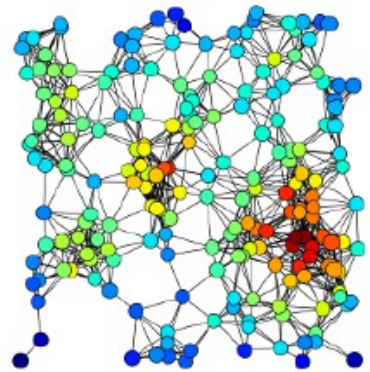
B

Closeness centrality

Eigenvector centrality



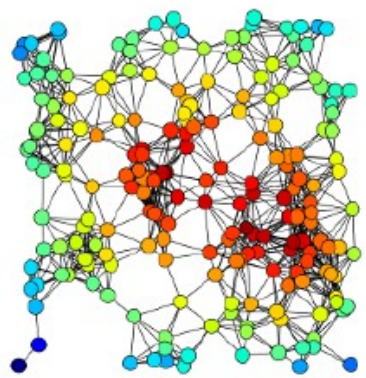
C



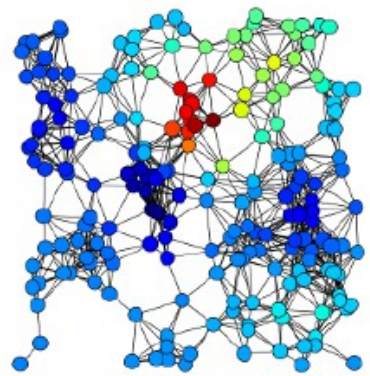
D

Degree centrality

Harmonic centrality



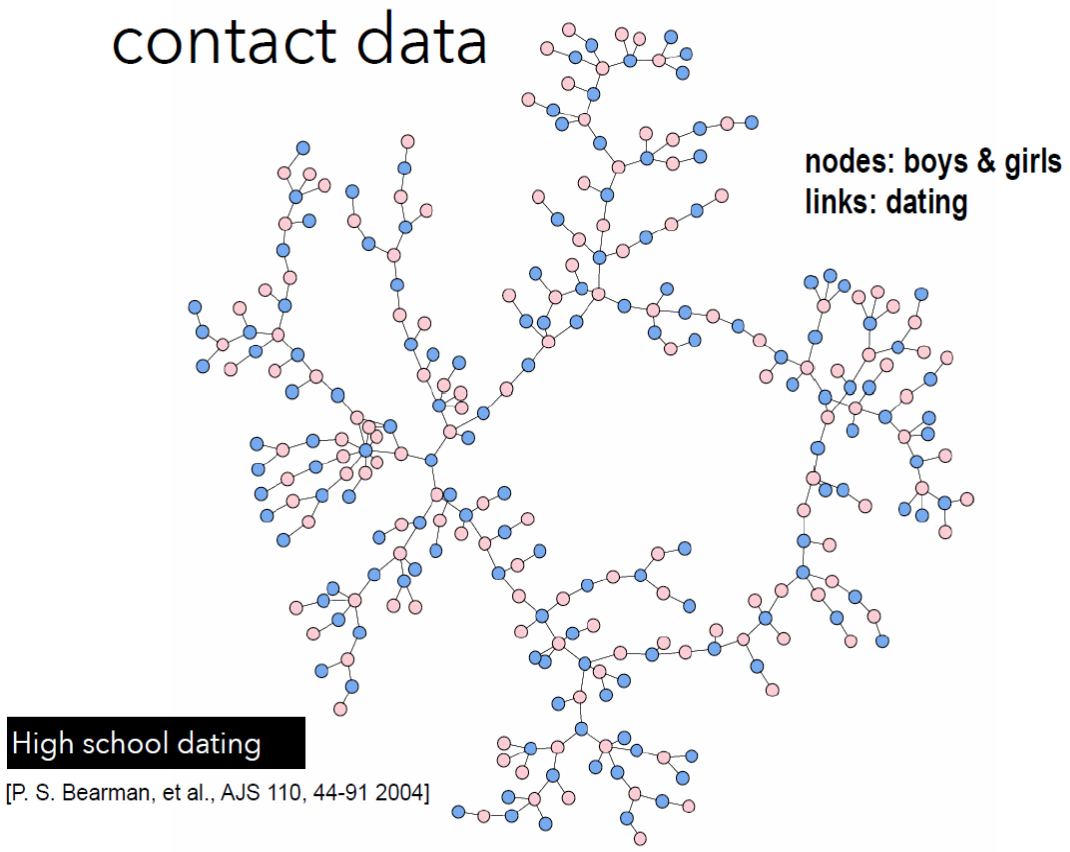
E



F

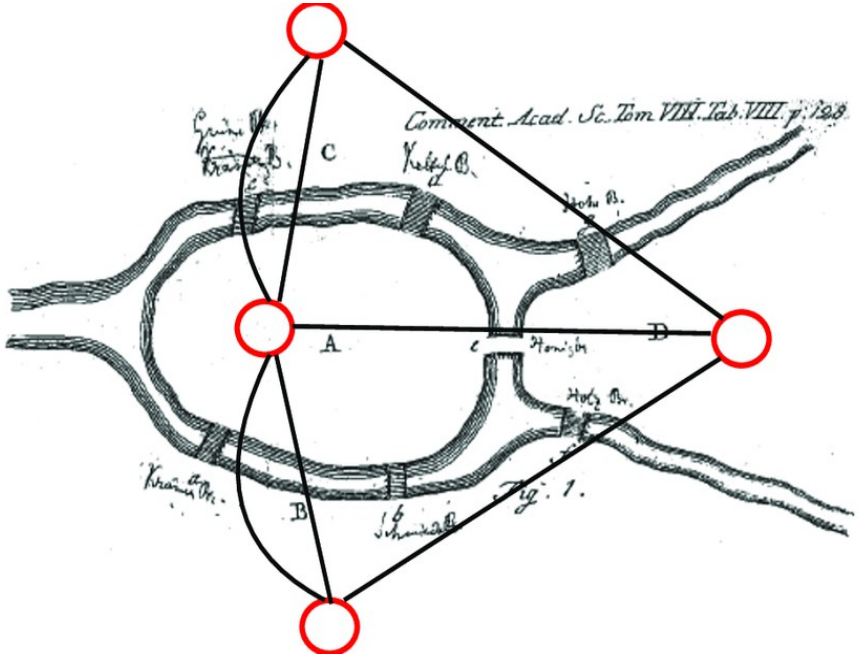
Katz centrality

For your projects:
How to analyze networks with network measures?



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

For your projects:
How to find interesting questions?



Classical example Euler (1736):
Research question:

The problem was to devise a walk through the city that would cross each of those bridges once and only once.

Data analysis or algorithm of solution:

Representation of the path and problem as a graph and proof that it is impossible.

Conclussions:

...

Hands-on part

Python
Gephi

Q&A Group discussions
Projects and reversed classrooms

Hands-on part

Idea:
Each group gets one projects

Internet and WWW:
Data collected in 1999. Ref: Albert, R., Jeong, H., & Barabasi, A. L. (1999). Internet: Diameter of the world-wide web. Nature, 401(6749), 130-131

Citation network:
Leskovec, J., Kleinberg, J., & Faloutsos, C. (2007)

Neural networks:
Classes of small-world networks L. A. N. Amaral (2000)

Ecological network
Ecological networks Montoya Sole (2000)

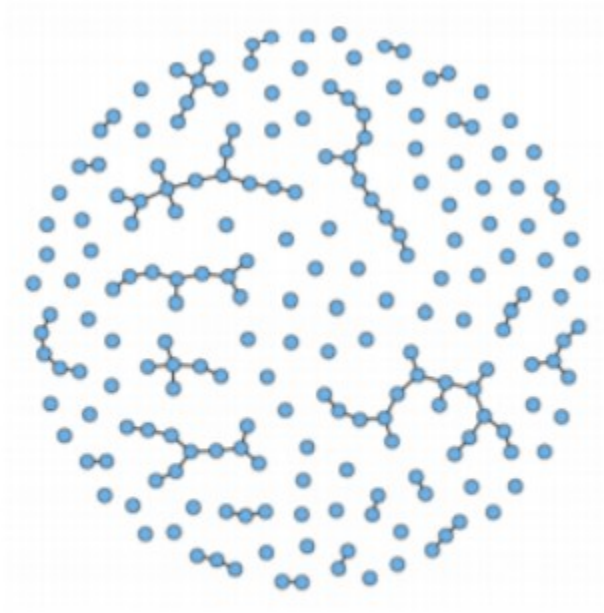
Transportational networks
Barthelemy (2011)

Additional slides:
Random networks

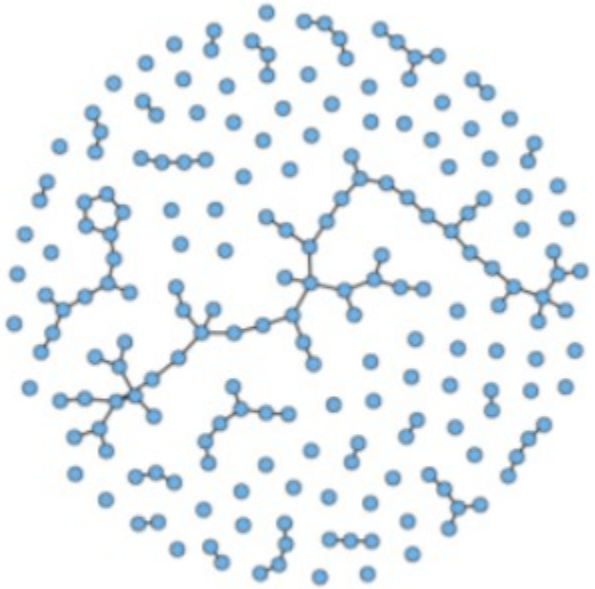
Erdos-Renyi random network

Erdos-Renyi networks

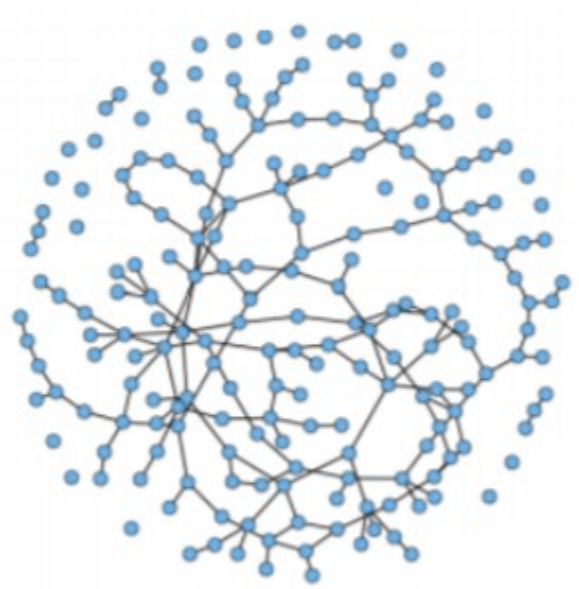
G(N,p), N number of nodes
p is probability to have link between nodes



$p < p_c$

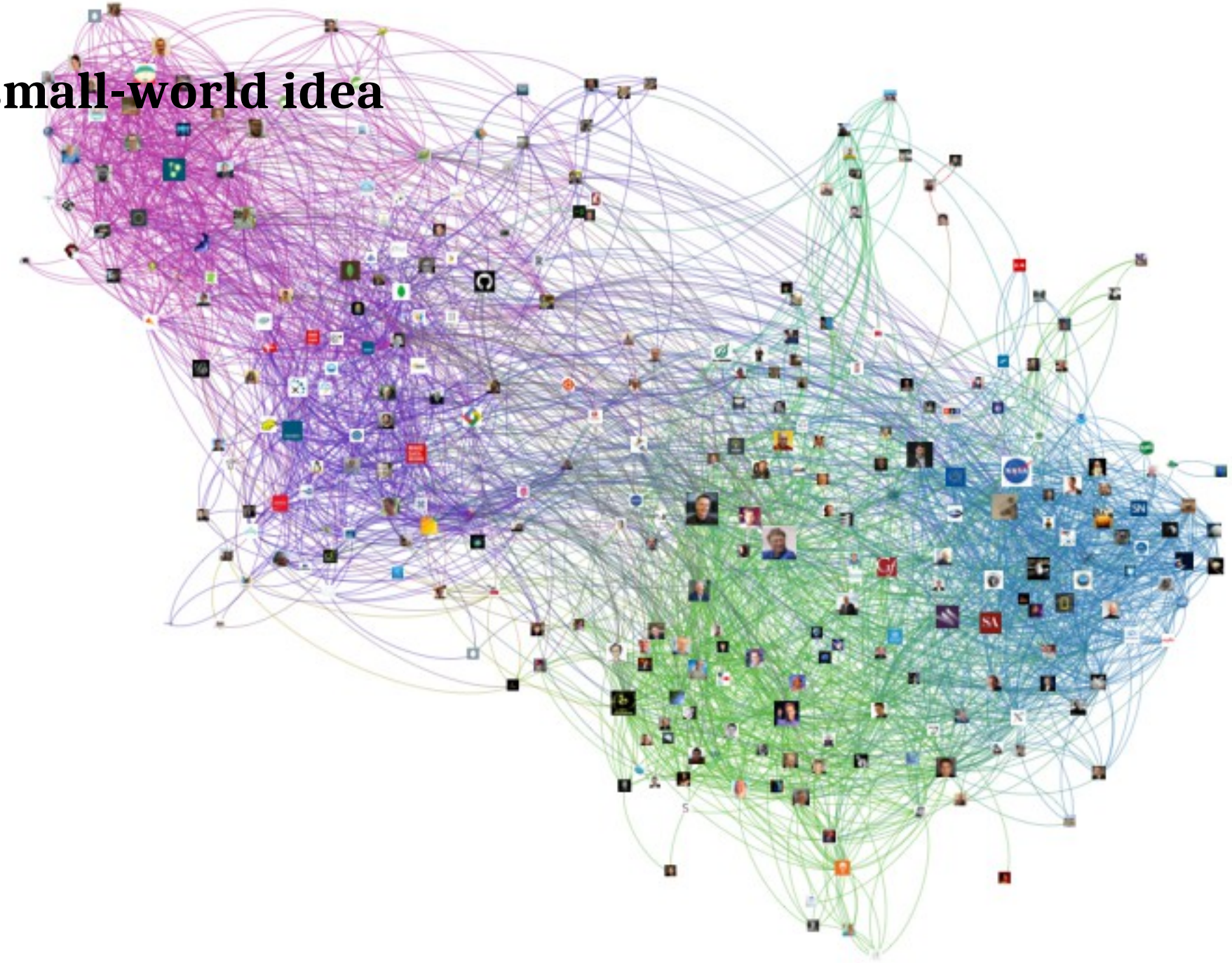


$p = p_c$



$p > p_c$

Watts-Strogarz small-world idea



Watts-Strogarz small-world model