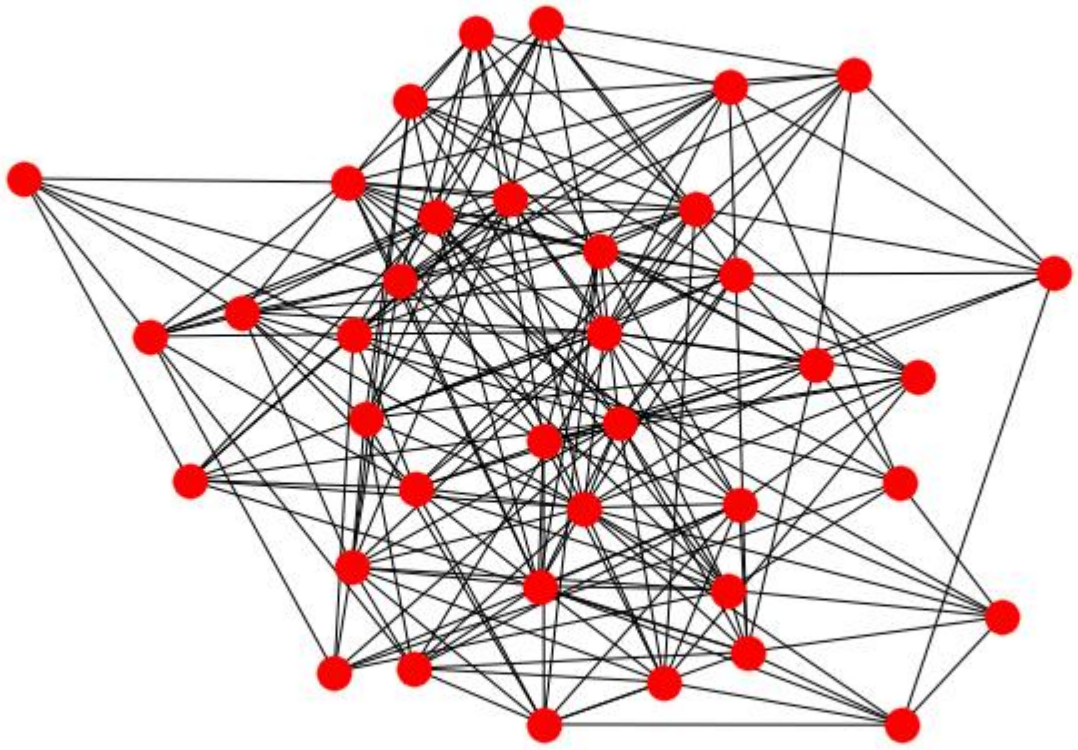
CCNSS 2018 Module 5
Tutorial 1:

Introduction to
Complex Network Theory

What is a network?

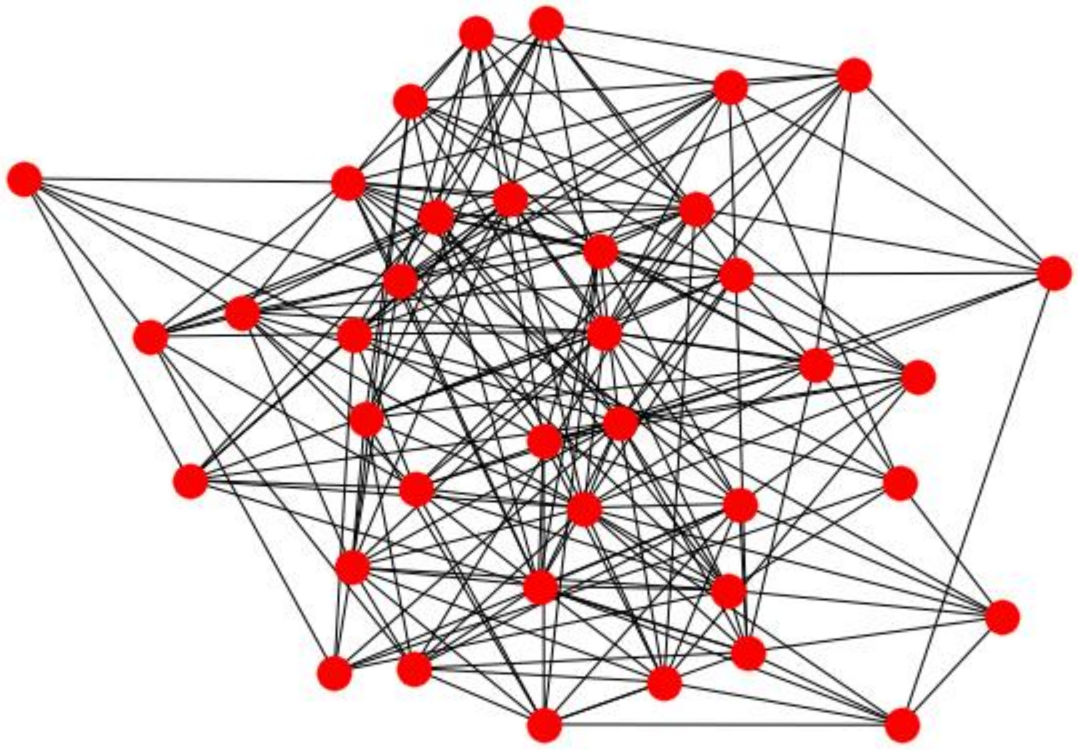
What is a network?

A network is set of elements with connections between them.



What is a network?

A network is set of elements with connections between them.



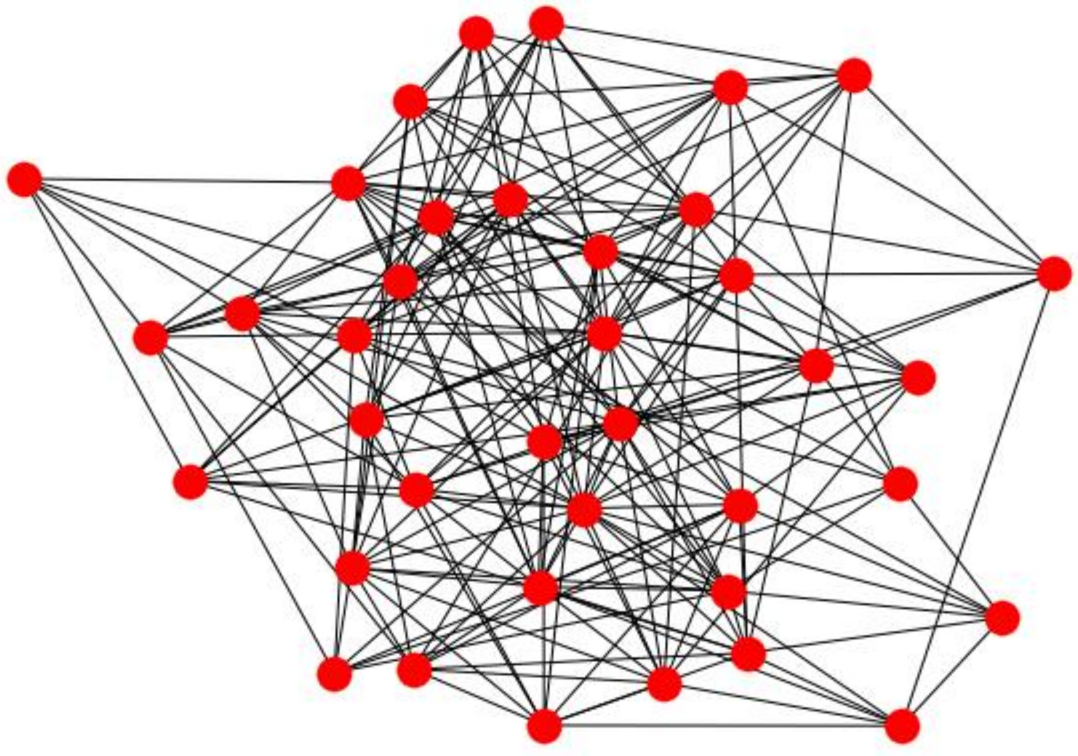
A network (graph) G consists of:

set of nodes (vertices)
$N = \{n_1, n_2, n_3, \ldots\}$,

set of links (edges)
$l = \{l_1, l_2, l_3, \ldots\}$.

What is a network?

A network is set of elements with connections between them.



System: G

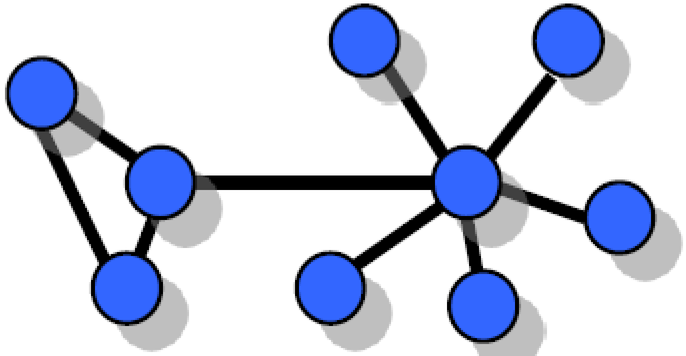
Components: $N = \{n_1, n_2, n_3, \ldots\}$,

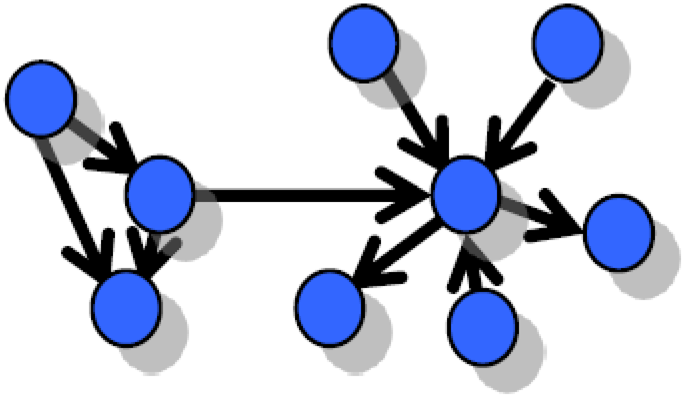
Interactions: $l = \{l_1, l_2, l_3, \ldots\}$.

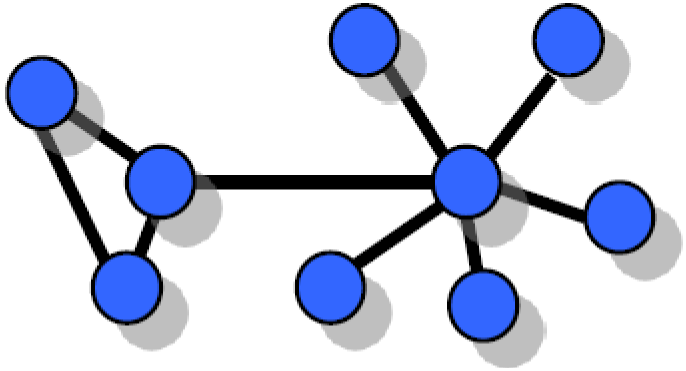
Types of networks

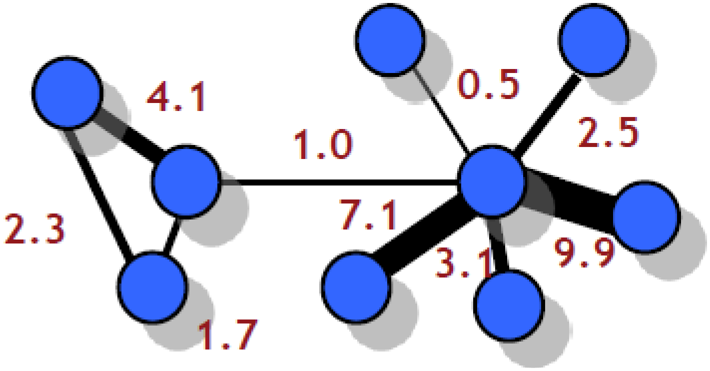
There exists various classifications of networks.

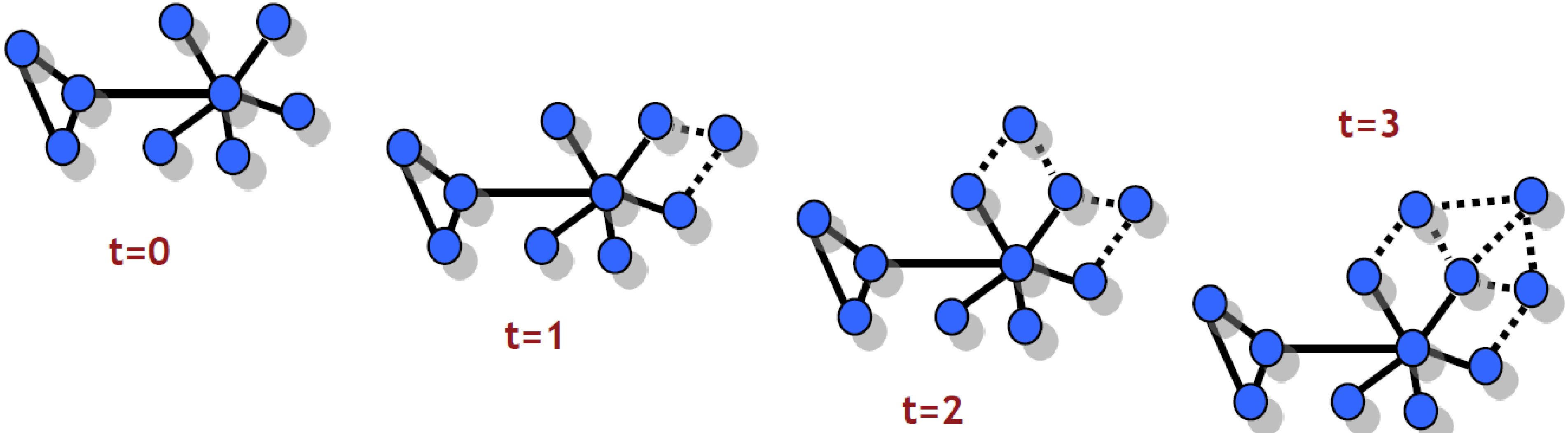
- Do the links have direction? Directed/Undirected
- Do the links have weight? Weighted/Unweighted
- Does the topology of the network change? Static/Evolving
- Does the nodes have dynamics? With/Without Dynamics





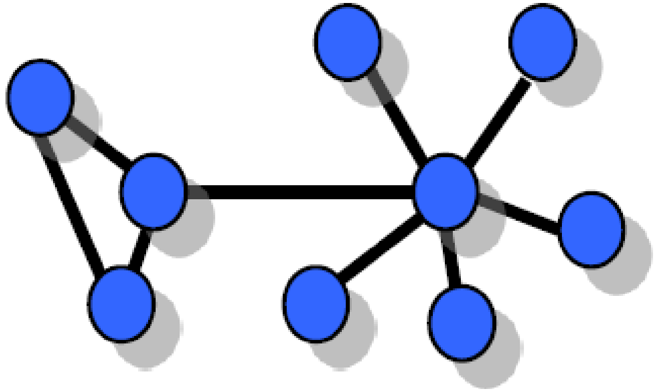






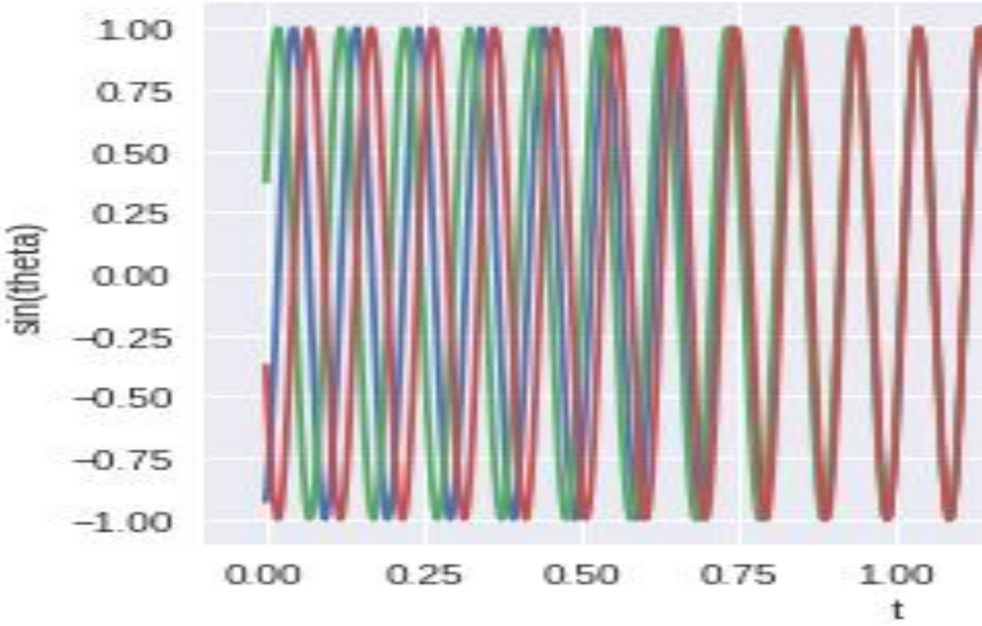
Javier Buldu

Nodes can be (coupled) dynamic systems.



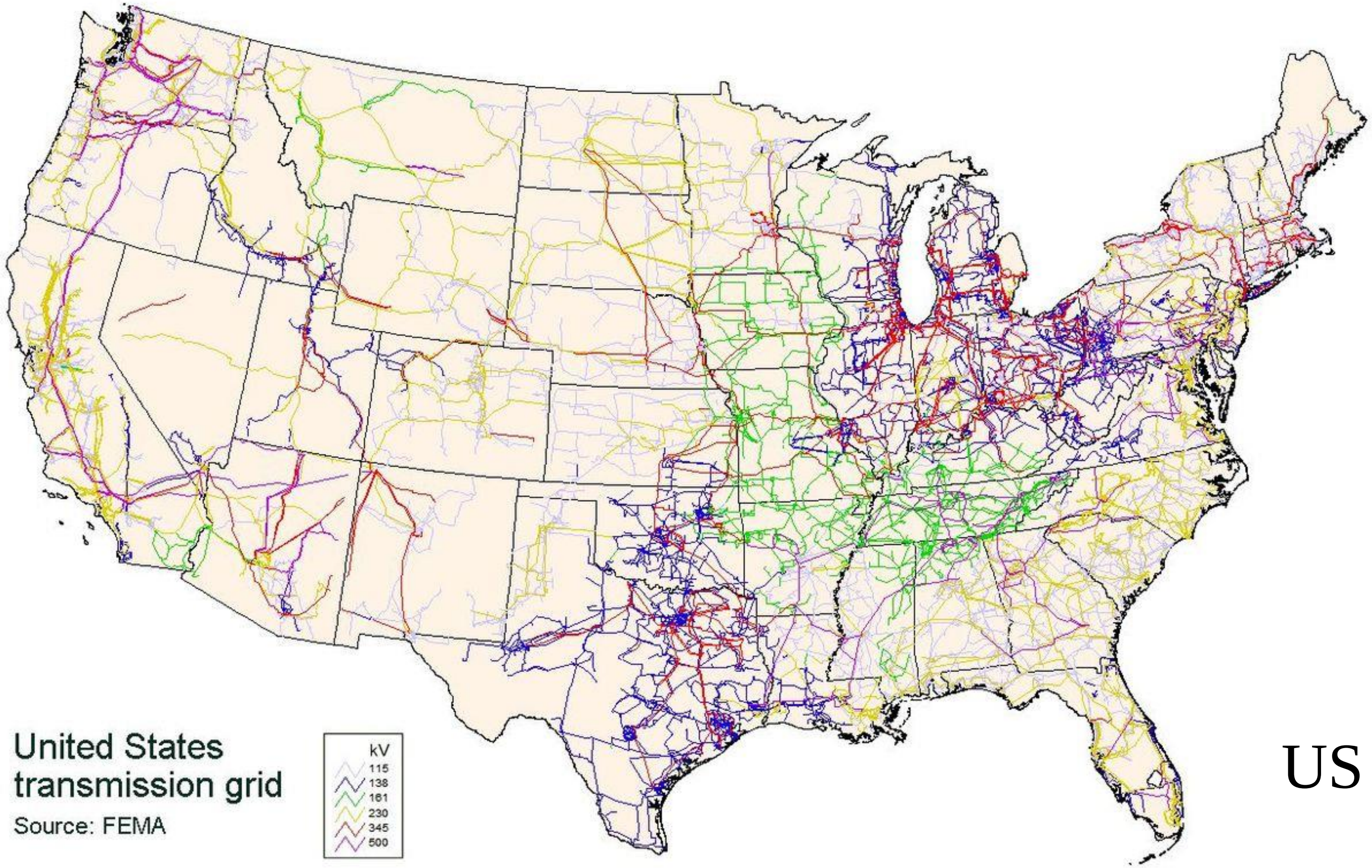
$$\dot{\theta}_j(t) = \omega_j + R^z S \sum_{k=1}^{N} K_{jk} \sin\left(\theta_k(t-\tau) - \theta_j(t)\right), \qquad j = 1,2,\dots,N.$$

Each node can have dynamics (state which changes in time), which is influenced by neighbors connected to the node.



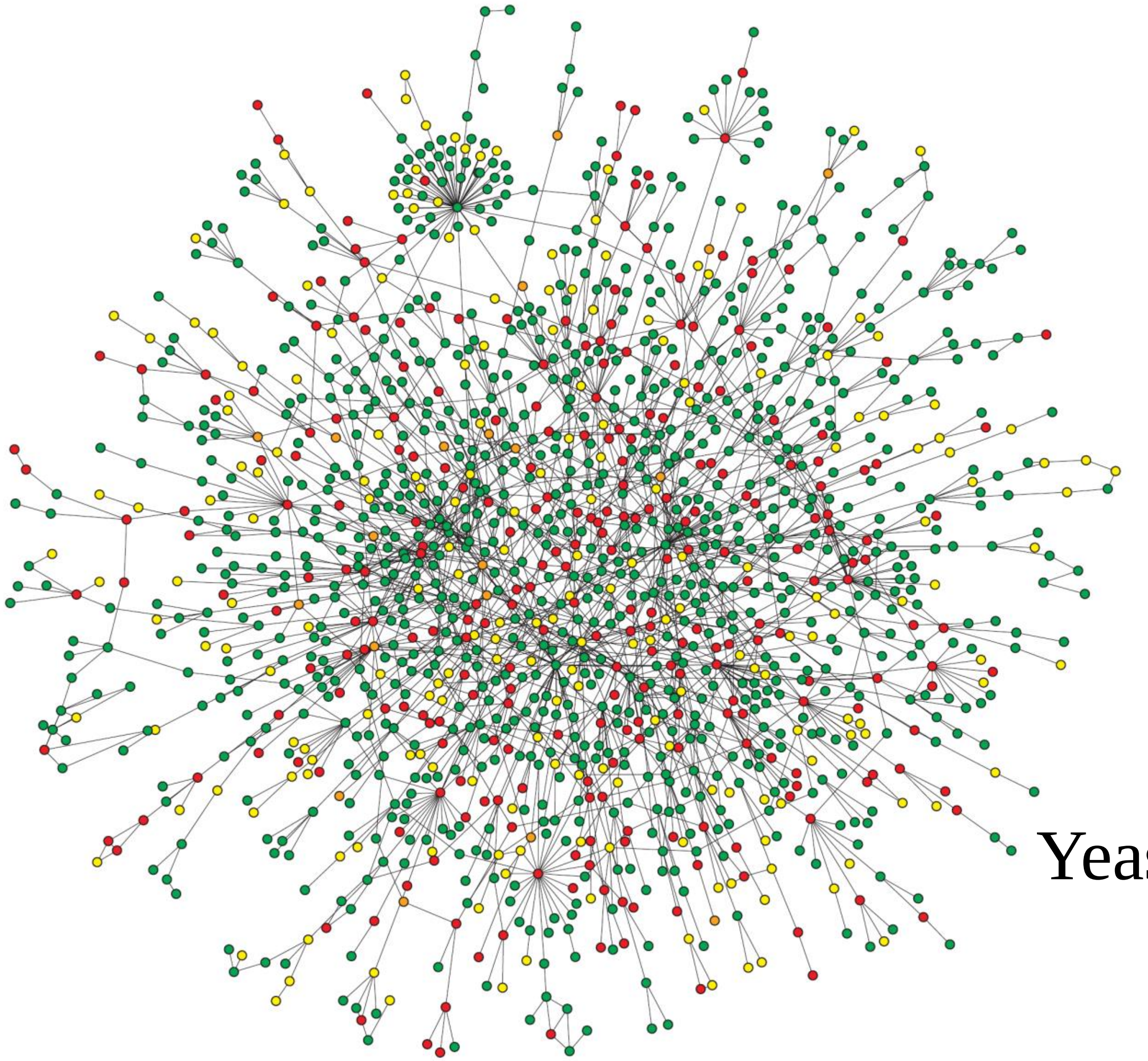
Examples of networks

Power grid network



US power grid

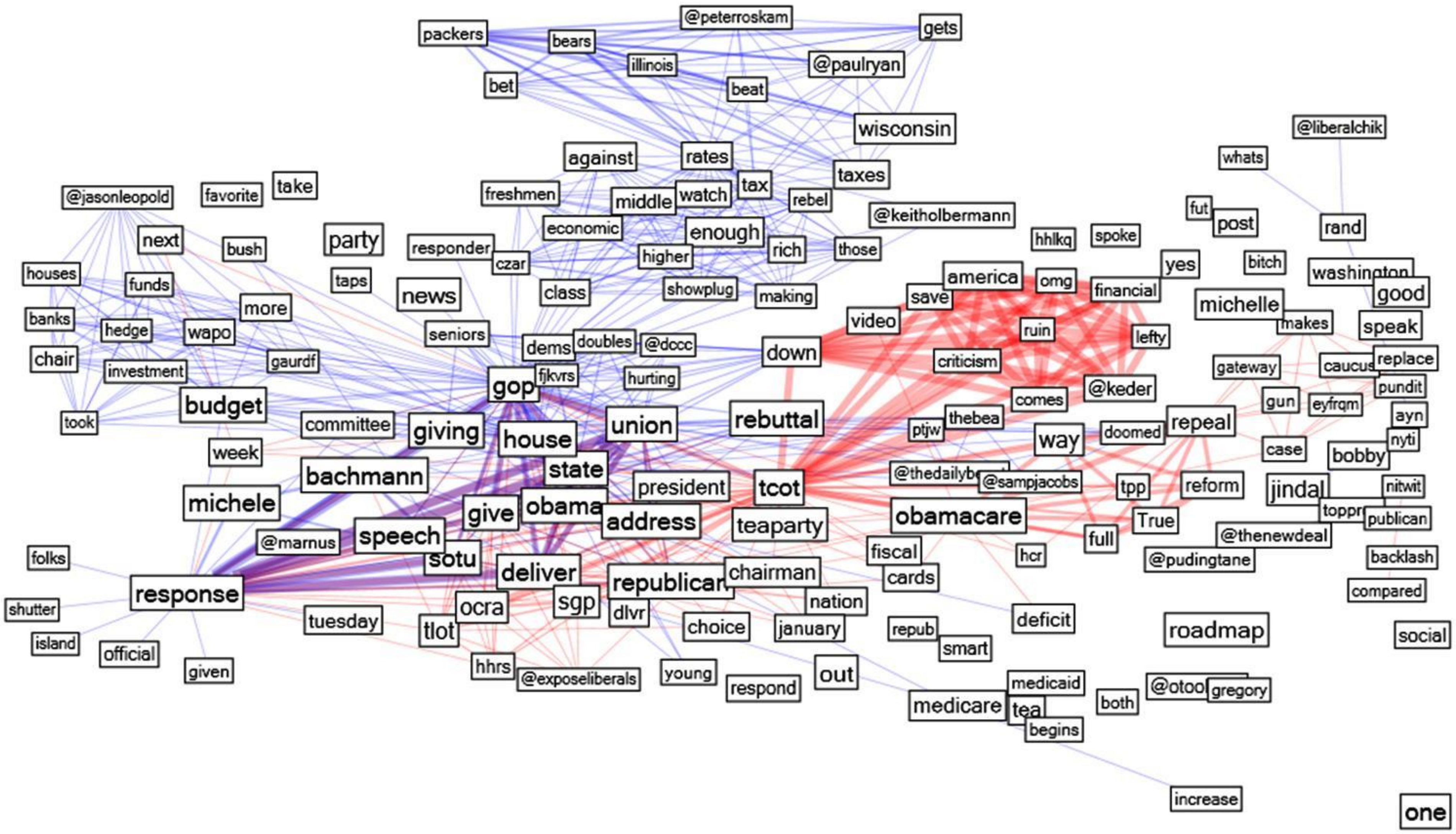
Protein-protein interaction network



Yeast p-p int. network

H. Jeong, S. P. Mason, A.-L. Barabási & Z. N. Oltvai, *Nature*, 411, 41–42 (2001)

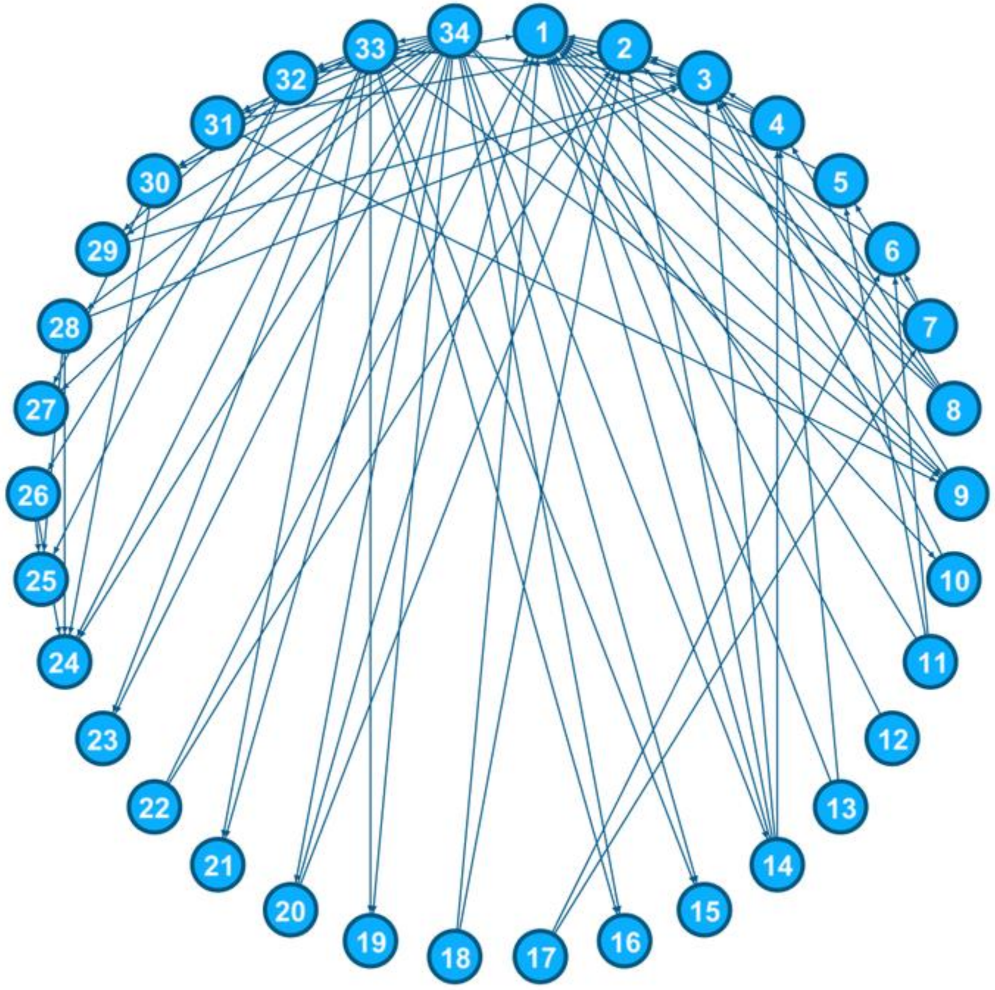
Word association network

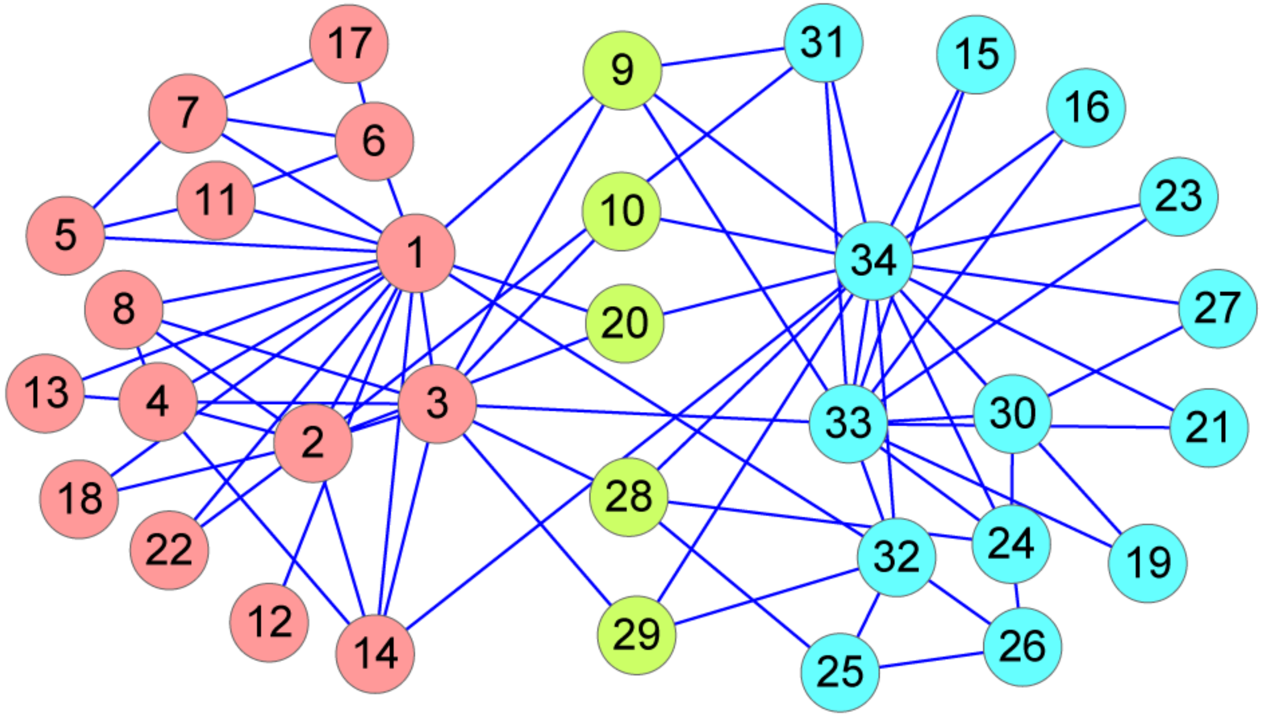


Crawled from tweeters, following a congress address by Republican Paul Ryan

https://www.connectedaction.net/keyword-networks-create-word-association-networks-from-text-with-nodexl-with-a-macro/

Social network

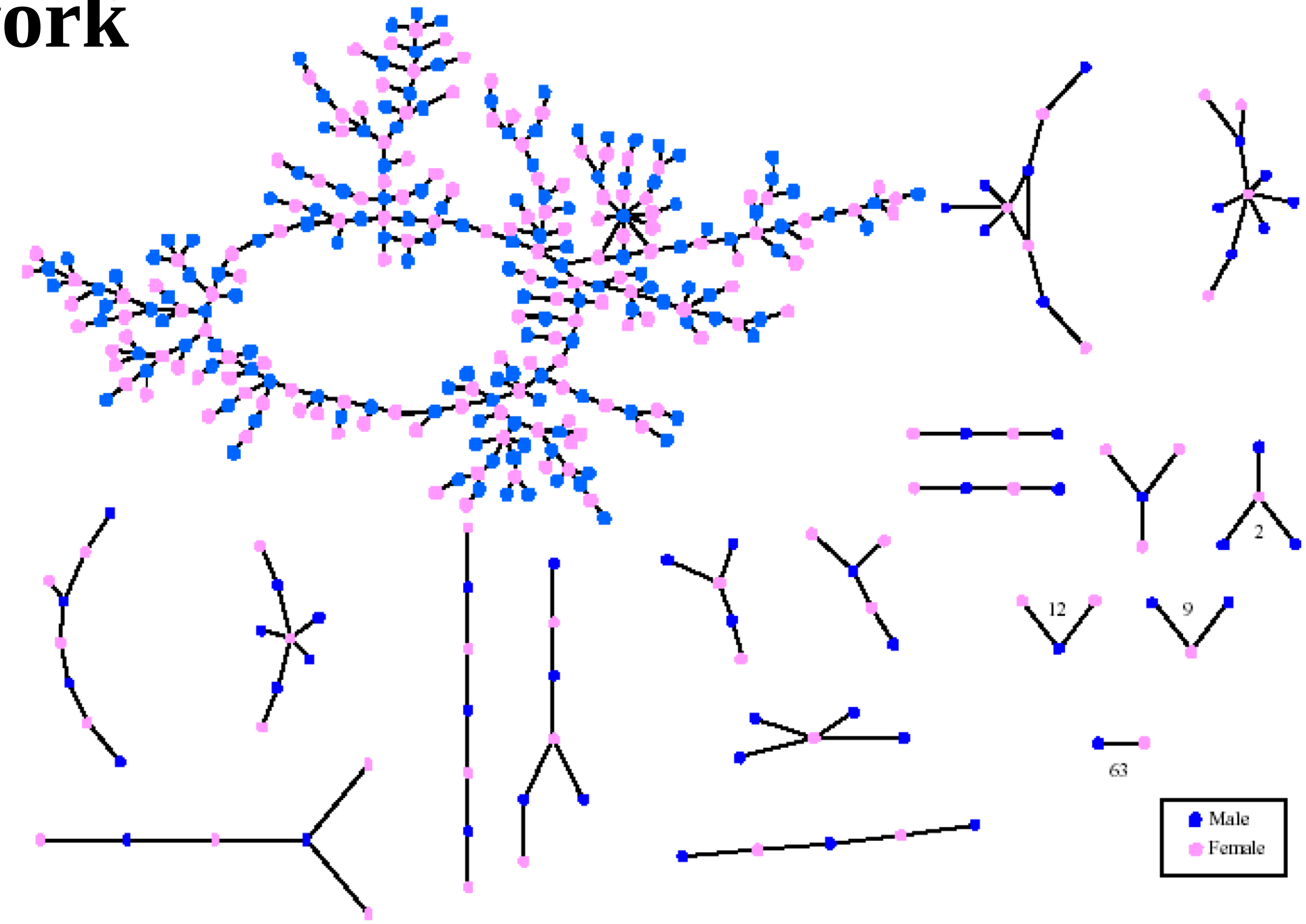




A social network of a karate club was studied by Wayne W. Zachary 1970-1972

W. Hao L. Gao J. Dong and X Yang, PLoS ONE 9(3):e91856 · March 2014
M. Girvan and Newman M. E. J., PNAS 99, 7821–7826 (2002)

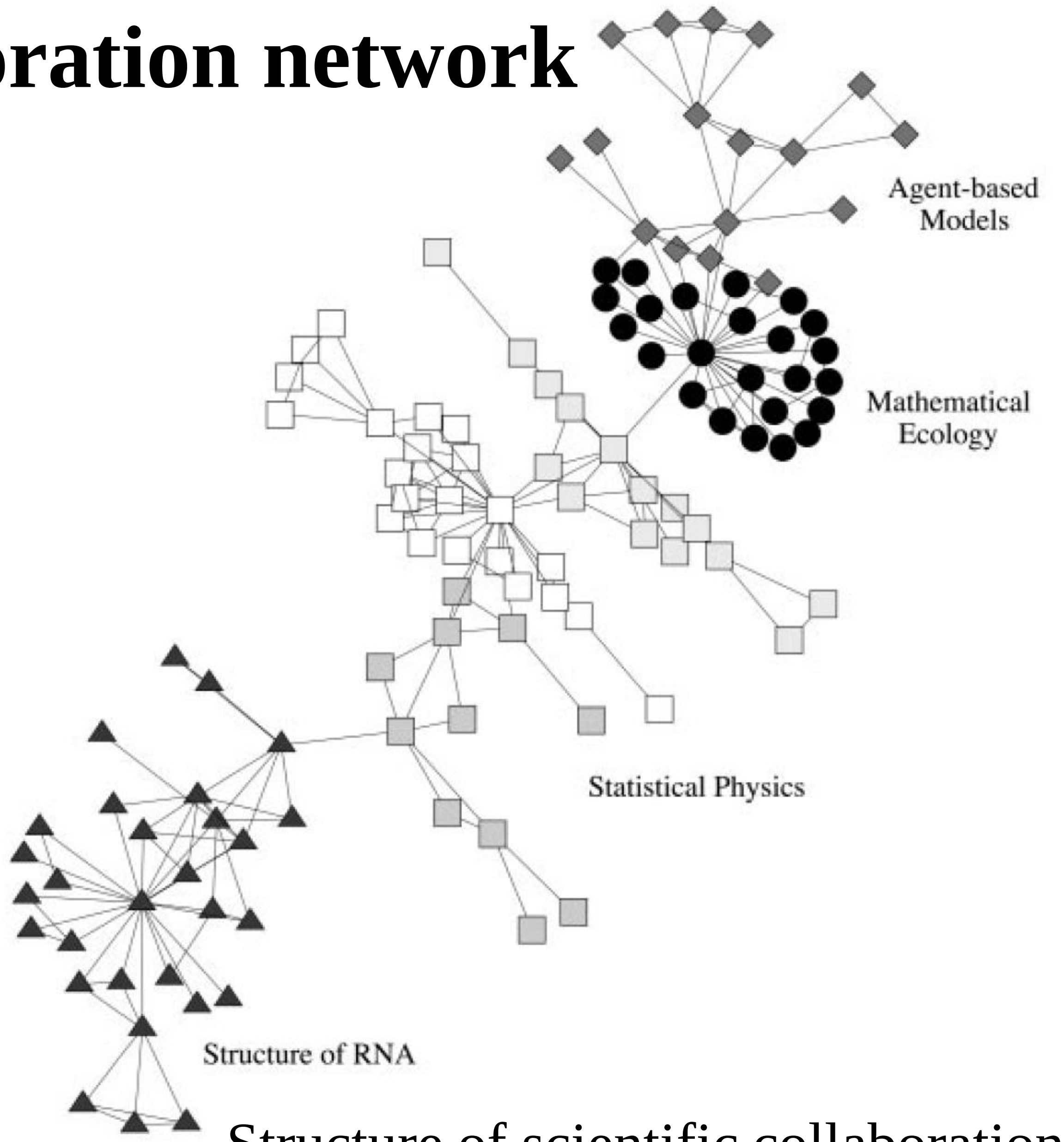
Social network



Structure of romances at Jefferson High School, 2004

P.S. Bearman, J. Moody, and K. Stovel, Am. J. Sociology, 110, 44 (2004)

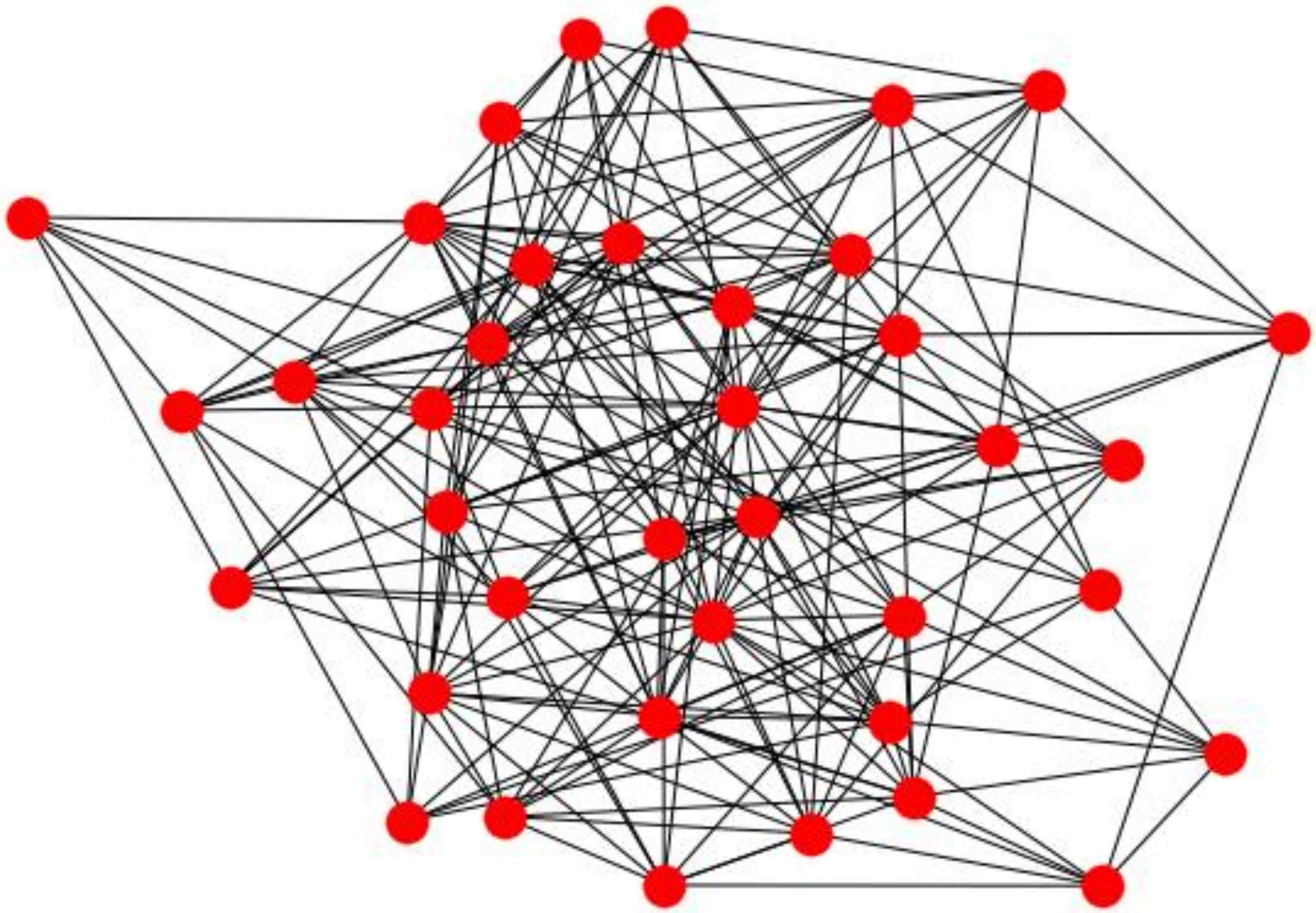
Scientific collaboration network



Structure of scientific collaboration at Santa Fe Inst., 2002

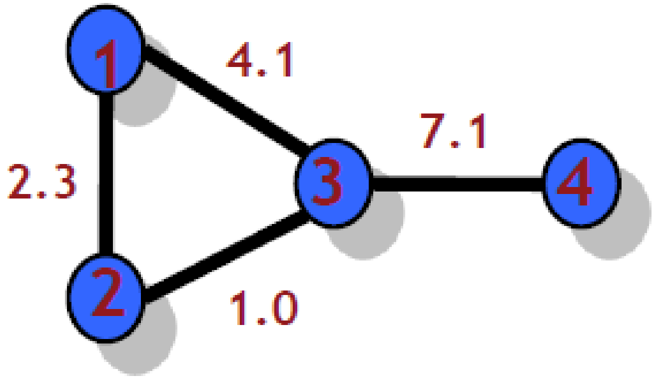
M. Girvan and M. E. J. Newman, PNAS, 12, 7821-7826 (2002)

How do we analyze complex networks?



Network tools

Connectivity matrix/ Adjacency matrix



Connectivity matrix: each element C_{ij} is the weight of the link between i and j.

$$\begin{pmatrix} 0.0 & 2.3 & 4.1 & 0.0 \\ 2.3 & 0.0 & 1.0 & 0.0 \\ 4.1 & 1.0 & 0.0 & 7.1 \\ 0.0 & 0.0 & 7.1 & 0.0 \end{pmatrix}$$

Adjacency matrix: A_{ij} is either 1 or 0 depending on the existence of the link between i and j.

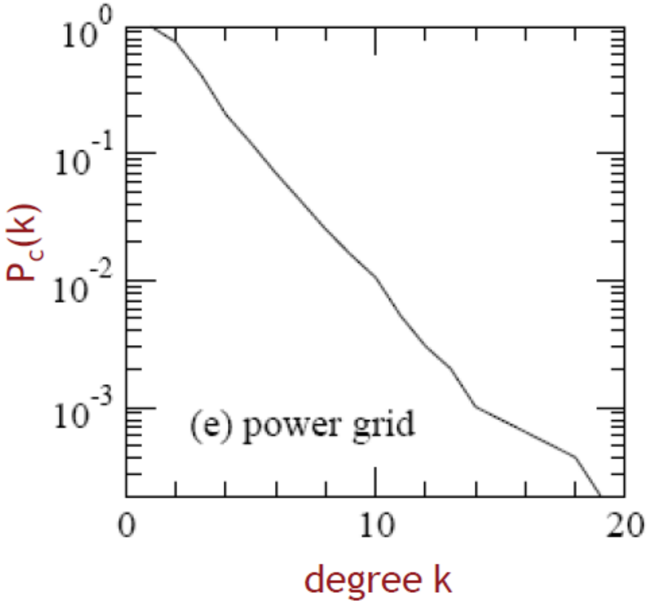
$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Degree and degree distribution

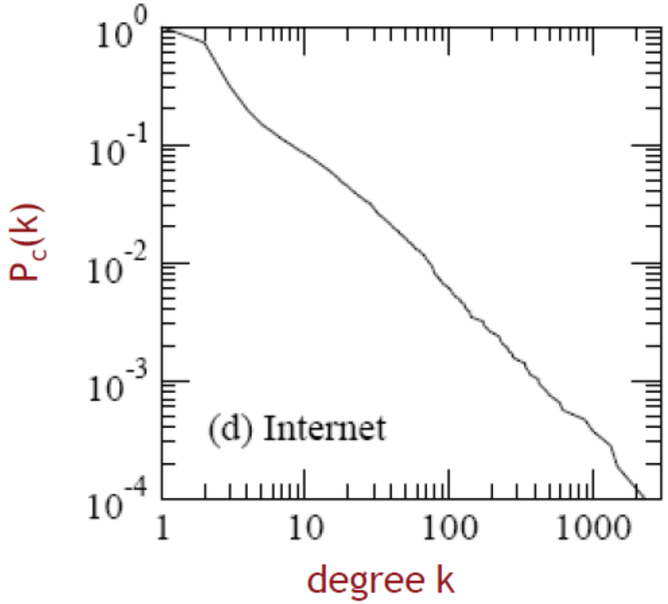
Degree k_i is the number of connections a node i has.

(Cumulative) degree distribution P(k) is the probability distribution of the degrees for the nodes in a network.

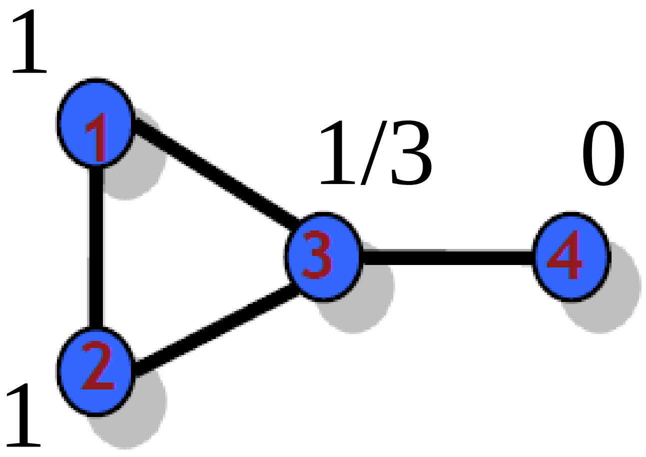
Exponential: $P(k) \sim e^{-\alpha k}$



Power law: $P(k) \sim k^{-\gamma}$



Clustering coefficient



$C_{1,2,3,4}$={1,1,1/3,0}

C ~ 0.58

Clustering coefficient of a node C_i gives you information on how clustered the neighbors are.

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

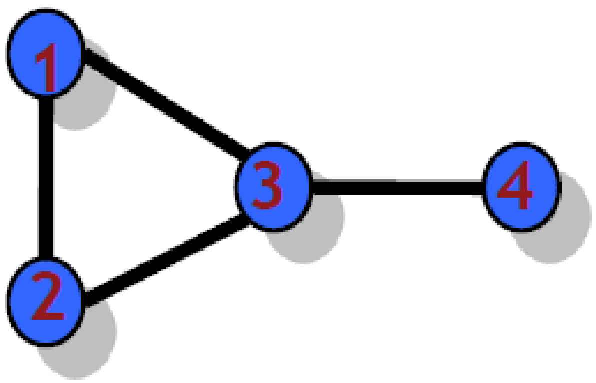
It is defined as the ratio between the number of links between neighbors (ln_i) and all possible connections between them $\binom{k_i}{2}$.

Also, we can define mean C:

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

Clustering coefficient

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



Example code for computing C_i ,with i=1:

i) Find all the neighbours of 1: $U_1=\{2,3\}$

ii) Find the neighbours of 2 in U_1: $W_2=\{3\}$

iii) Find the neighbours of 3 in U_1: $W_3=\{2\}$

iv) Sum the neighbours we found /2 : $ln_1 = (1+1)/2 = 1$

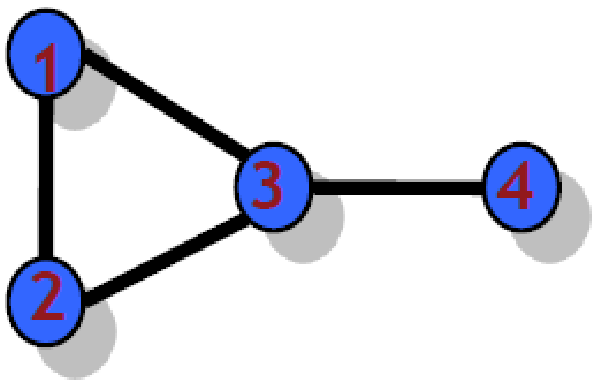
v) Divide above number ln_1 by $\binom{2}{2}=1$

We get $C_1 = 1$

$C_{1,2,3,4}=\{1,1,1/3,0\}$

$C \sim 0.58$

Distance, characteristic path length



Distance between two nodes d_{ij} are the smallest number of links that connects *i* to *j*.

Characteristic path length is the average of the distances of all pair of nodes:

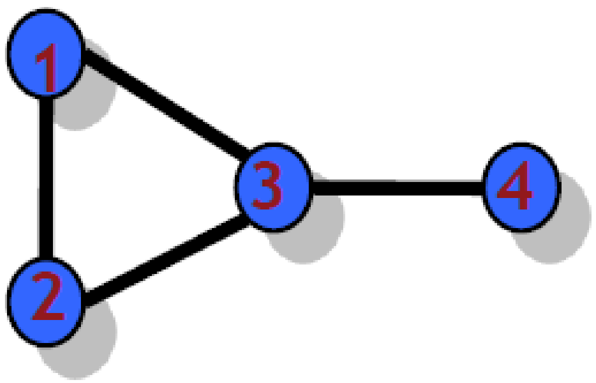
$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij}$$

$D_{1,2}=D_{1,3}=D_{2,3}=D_{3,4}=1$

$D_{1,4}=D_{2,4}=2$

L = 2/3

Distance, characteristic path length



Distance between two nodes d_{ij} are the smallest number of links that connects i to j.

How do we calculate the distance matrix?

We can use the following nice property of Connectivity matrix A:

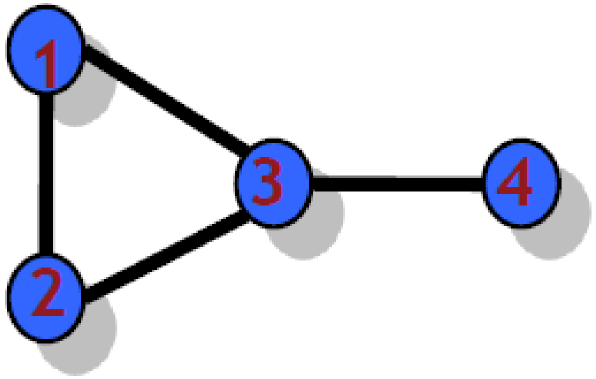
The elements in the matrix product of A, ${A_{ij}}^n$ gives you the number of paths of length n from i to j.

$D_{1,2}=D_{1,3}=D_{2,3}=D_{3,4}=1$

$D_{1,4}=D_{2,4}=2$

L = 2/3

Betweenness Centrality



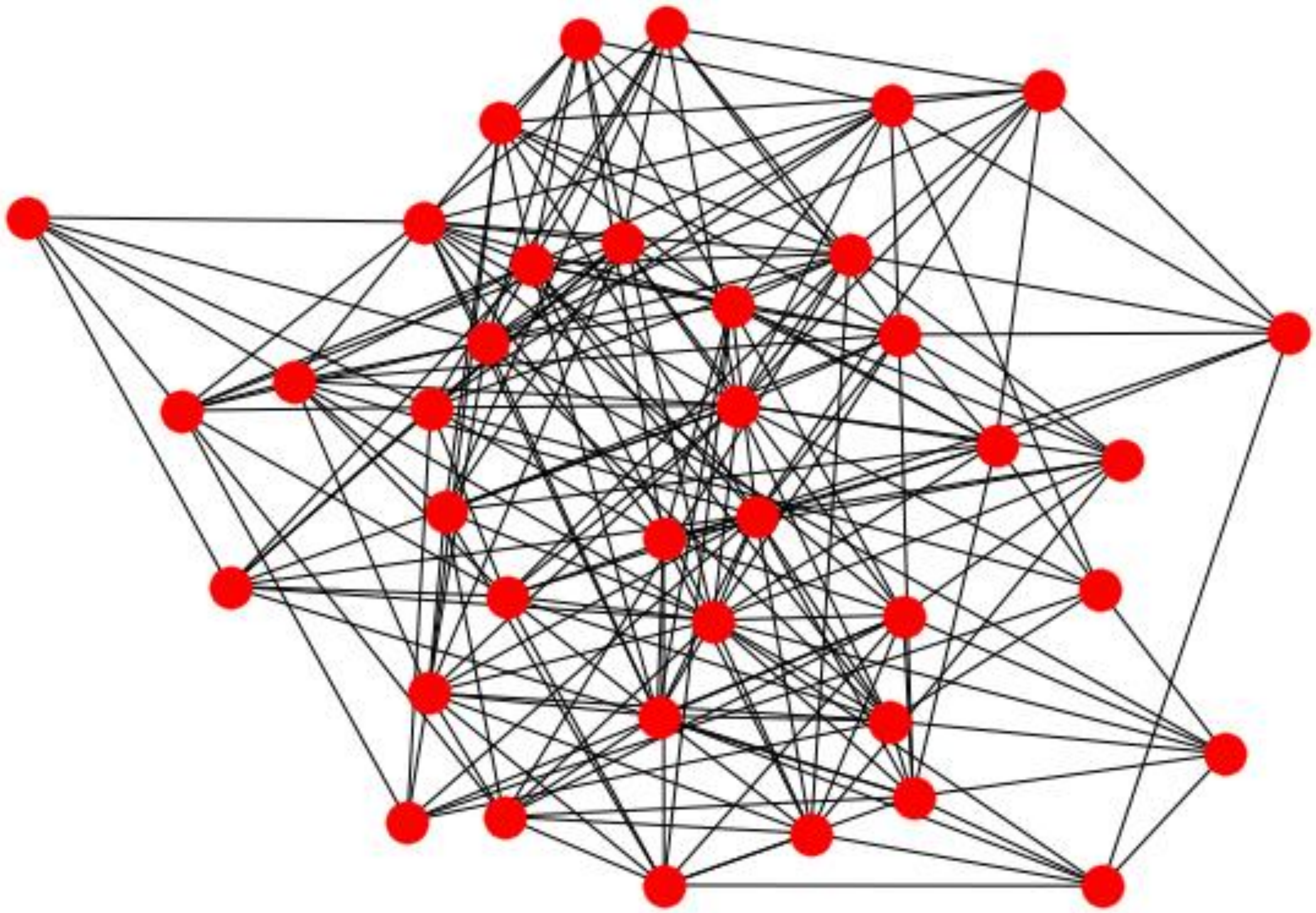
Betweenness centrality of a node i gives you the number of shortest paths that goes through the node i.

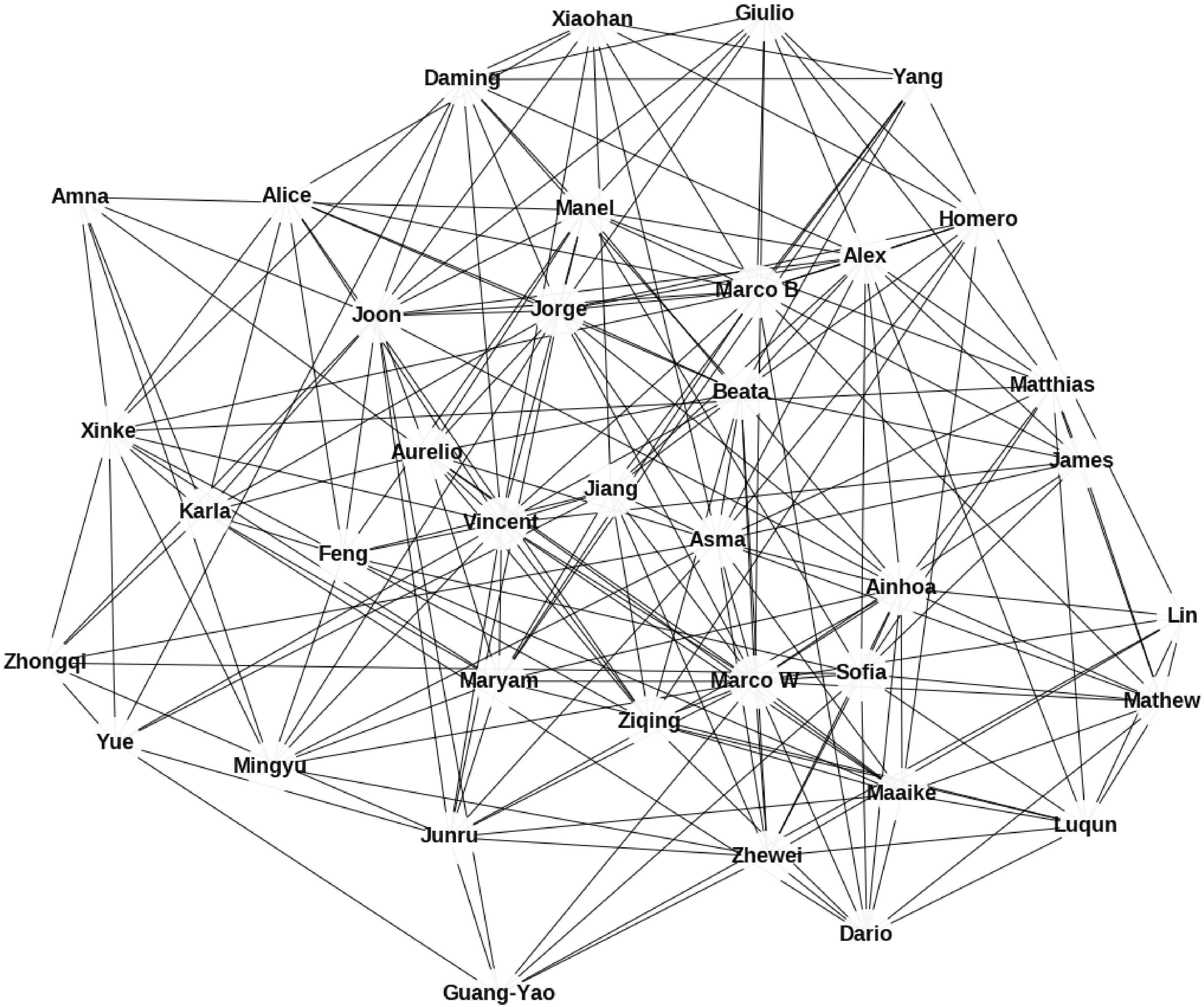
$$b_i = \sum_{j,k \in N, j \neq k} \frac{n_{jk}(i)}{n_{jk}}$$

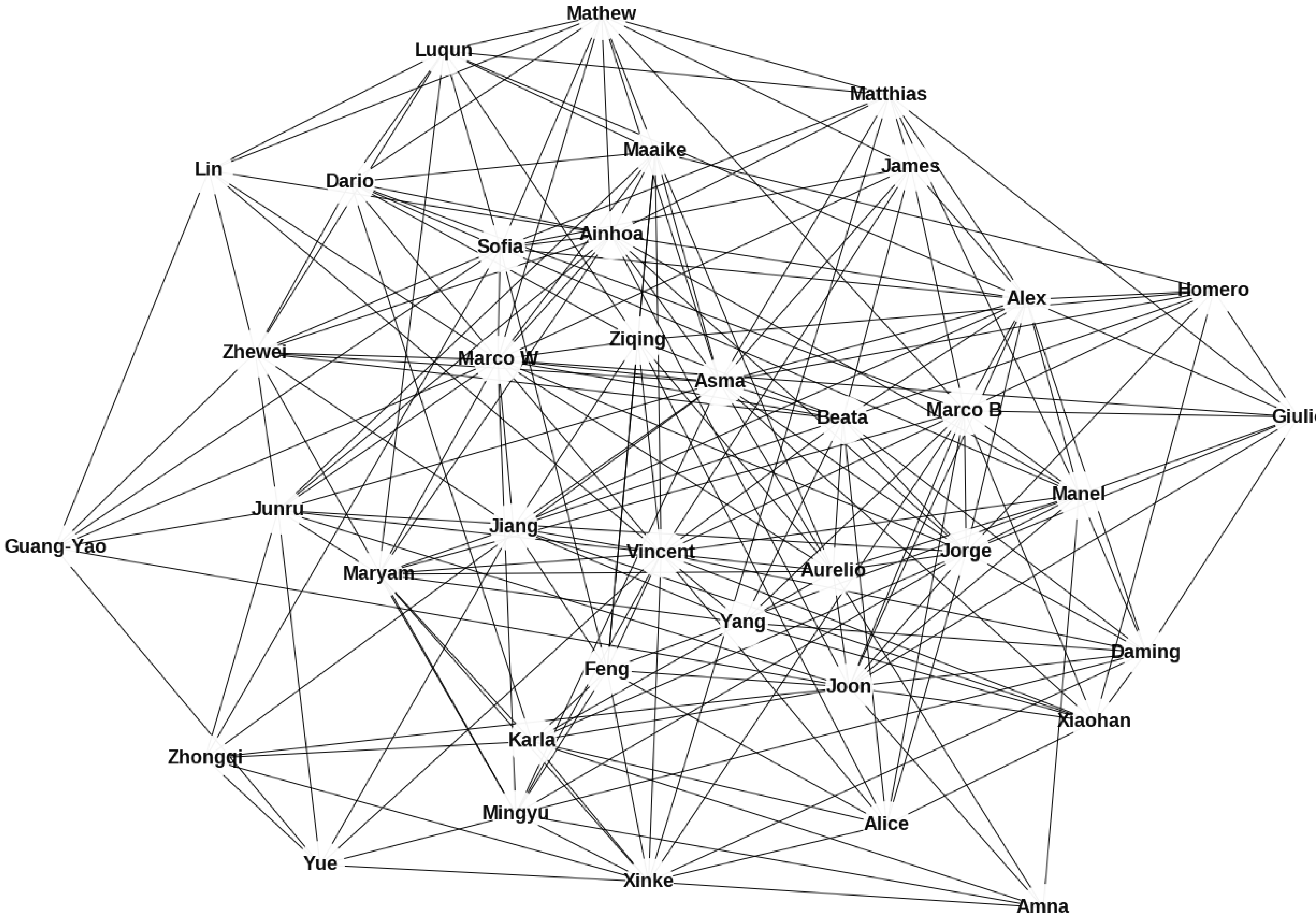
where n_{jk} is the number of the shortest paths between nodes $_j$ and $_k$, and $n_{jk}(i)$ is the number of these paths that go through node $_i$.

A node with high betweenness centrality would have more control over the network, because more information will pass through that node.

How do we analyze complex networks?




Xiaohan
Giulio
Daming
Yang
Amna
Alice
Manel
Homero
Alex
Marco B
Joon
Jorge
Beata
Matthias
Xinke
Aurelio
James
Jiang
Karla
Vincent
Asma
Feng
Ainhoa
Lin
Zhongqi
Maryam
Marco W
Sofia
Mathew
Ziqing
Yue
Mingyu
Maaike
Luqun
Junru
Zhewei
Dario
Guang-Yao



Mathew
Luqun
Matthias
Maaike
James
Lin
Dario
Ainhoa
Sofia
Alex
Homero
Ziqing
Zhewei
Marco W
Asma
Marco B
Beata
Giulio
Manel
Junru
Jiang
Guang-Yao
Vincent
Jorge
Aurelio
Maryam
Yang
Daming
Feng
Joon
Xiaohan
Karla
Zhongqi
Mingyu
Alice
Yue
Xinke
Amna

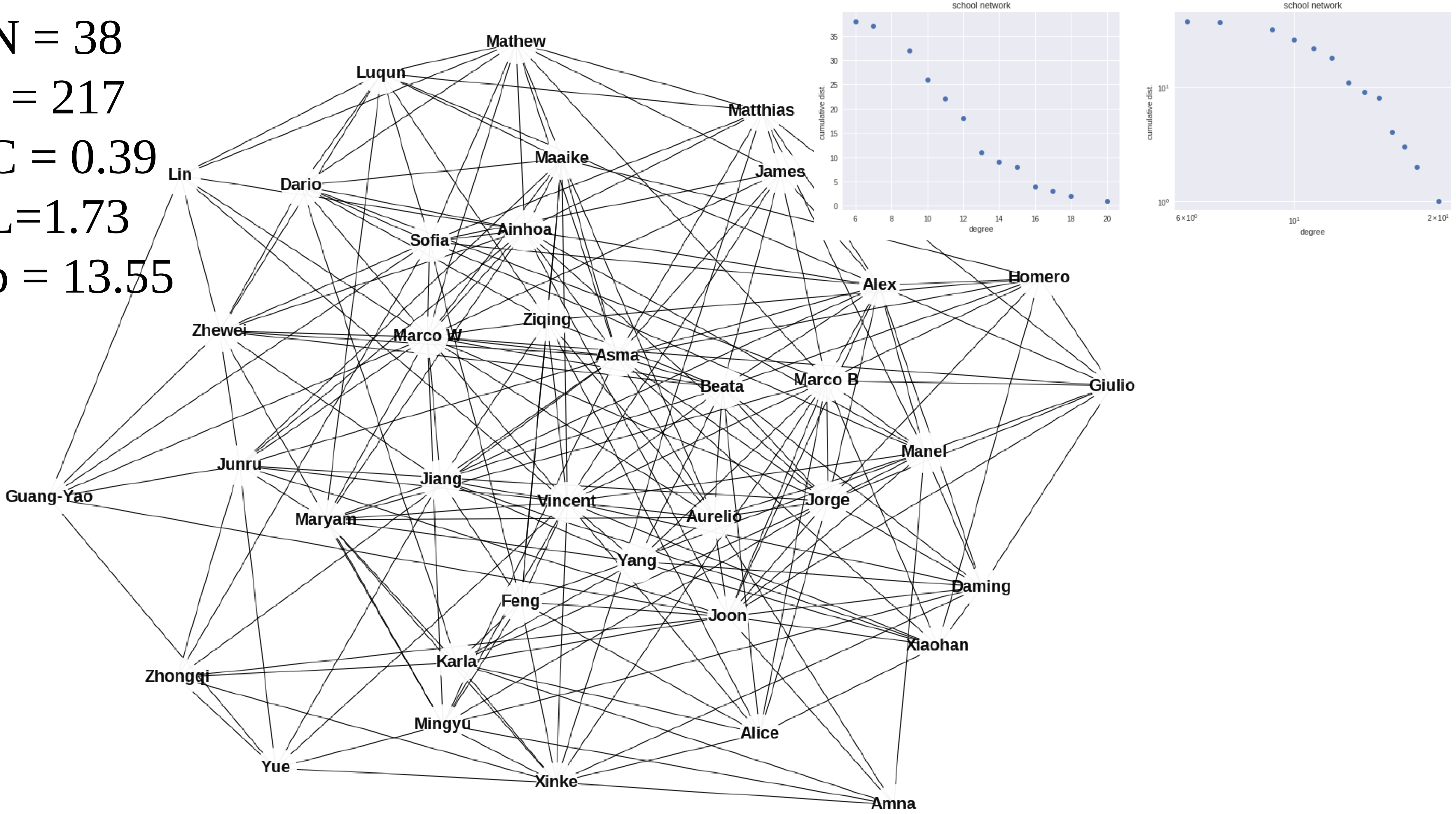
$N = 38$

$l = 217$

$C = 0.39$

$L = 1.73$

$b = 13.55$



Question:

What is your group’s average clustering coefficient, and average betweenness centrality?