Advanced Algorithms and Computational Models (module A)

Degree correlation

Giacomo Fiumara
giacomo.fiumara@unime.it

2020-2021

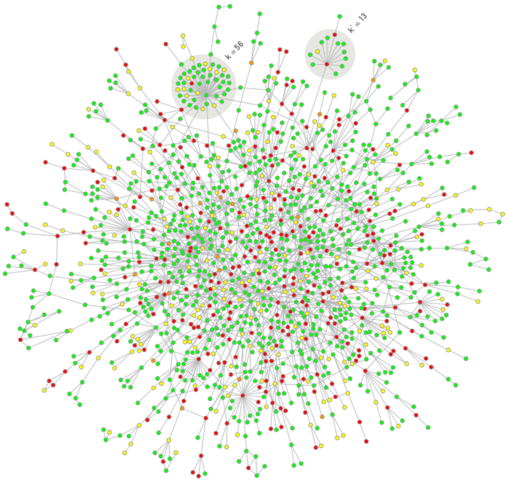
Introduction

- Angelina Jolie and Brad Pitt, Ben Affleck and Jennifer Garner, Harri- son Ford and Calista Flockhart, Michael Douglas and Catherine Zeta-Jones, Tom Cruise and Katie Holmes, Richard Gere and Cindy Crawford
- They are Hollywood stars that are or were married. Thanks to them we take for granted that celebrities marry each other
- Is this normal? What is the true chance that a celebrity marries another celebrity?

Introduction

- Assuming that a celebrity could date anyone from a pool of about a hundred million (10^8) eligible individuals worldwide, the chances that their mate would be another celebrity from a generous list of 1,000 other celebrities is only 10^{-5}
- Therefore, if dating were driven by random encounters, celebrities would never marry each other.
- This phenomenon highlights the structure of the social network
- Celebrities are hubs, and celebrity dating and joint board membership show an interesting property of social networks: **hubs tend to have ties to other hubs**

Introduction



Introduction

- This property is not present in all networks
- For example, the protein interaction network of yeast shows its scale-free structure
- There is a large number of degree-1 and degree-2 proteins coexisting with a few highly connected hubs
- These hubs tend to avoid linking to each other
- This behaviour is unusual. If eache node chooses randomly the nodes it connects to, then the probability that nodes with degree k and k' link to each other is

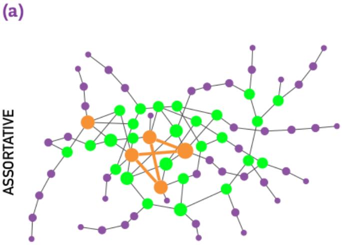
$$p_{k,k'} = \frac{kk'}{2L}$$

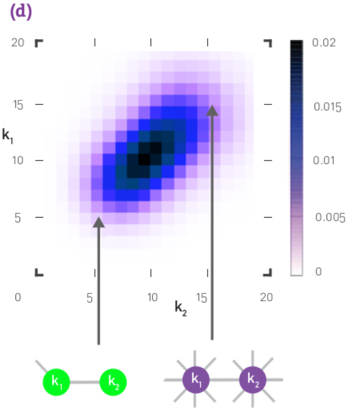
Introduction

- Therefore, hubs are more likely to connect to each other
- Yet, in the protein interaction network there are no direct links between the hubs, but a large number of direct links between small degree nodes can be observed
- In summary, in social networks hubs tend to connect each other, while in the protein interaction network the opposite is true
- These patterns are a manifestation of a general property of real networks and are call **degree correlations**

Degree correlation matrix

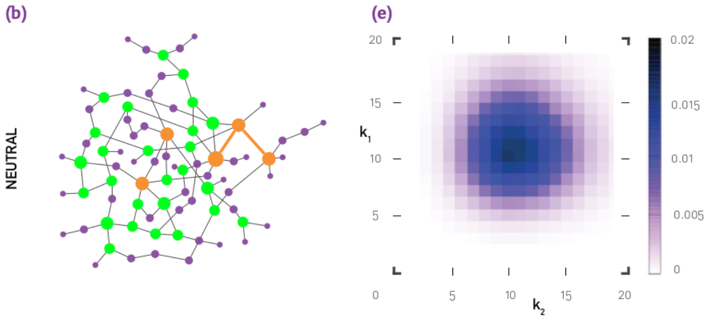
Assortative networks



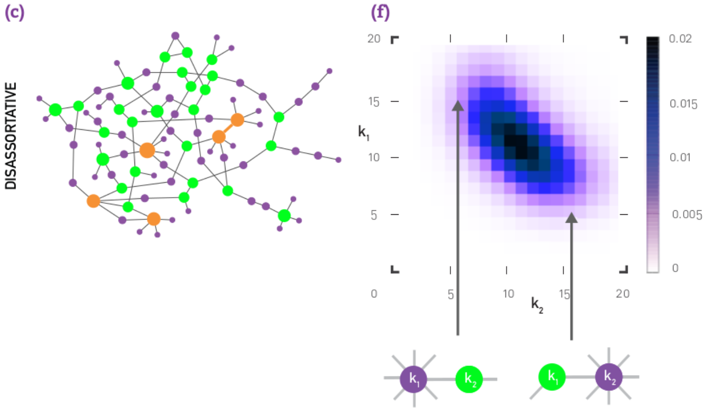


Degree correlation matrix

Neutral networks



Degree correlation matrix
Disassortative networks



Degree correlation matrix

- The information about degree correlations is captured by the *degree correlation matrix* e_{ij}, which is the probability of finding a node with degrees i and j at the two ends of a randomly selected edge.
- e_{ij} is a probability, therefore it is normalized

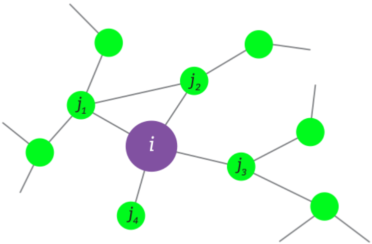
$$\sum_{i,j} e_{ij} = 1$$

- Information about degree correlations is carried by the degree correlation matrix
- As a matter of fact, it contains a huge amount of elements, thus making quite impossible a visual inspection

Measuring degree correlations

- Degree correlations capture the relationship between the degrees of nodes that link to each other
- One way to quantify their magnitude is to measure for each node i the average degree of its neighbours:

$$k_{nn}(k_i) = \frac{1}{k_i} \sum_{j=1}^{N} A_{ij} k_j$$



Measuring degree correlations

- The *degree correlation function* calculates for all nodes with degree k

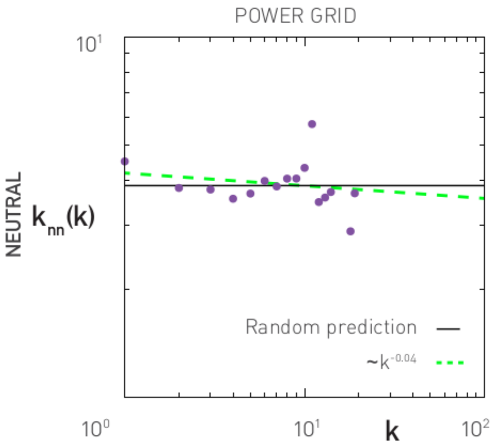
$$k_{nn}(k) = \sum_{k'} k' P(k'|k)$$

- where $P(k'|k)$ is the conditional probability that following a link of a k-degree node a degree-k' node is reached
- Therefore, $k_{nn}(k)$ is the average degree of the neighbours of all degree-k nodes

$$k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

- In a neutral network the average degree of a node's neighbours is independent of the node's degree k and depends only on the global network characteristics $\langle k \rangle$ and $\langle k^2 \rangle$
- Therefore, plotting $k_{nn}(k)$ in function of k results in a horizontal line at $\langle k^2 \rangle \langle k \rangle$

Measuring degree correlations
Neutral networks



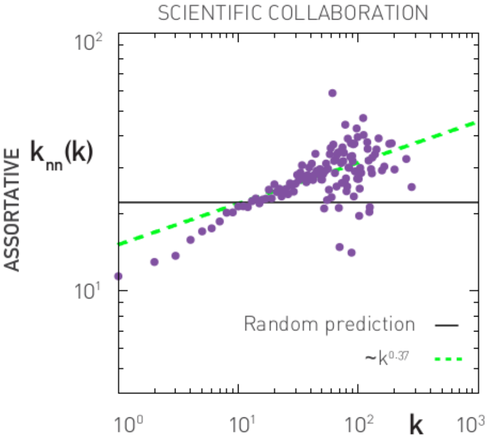
Measuring degree correlations

Assortative networks

- In assortative networks hubs tend to connect to other hubs, hence the higher is the degree k of a node, the higher is the average degree of its nearest neighbours
- Therefore, $k_{nn}(k)$ increases with k, as observed for scientific collaboration networks

Measuring degree correlations

Assortative networks



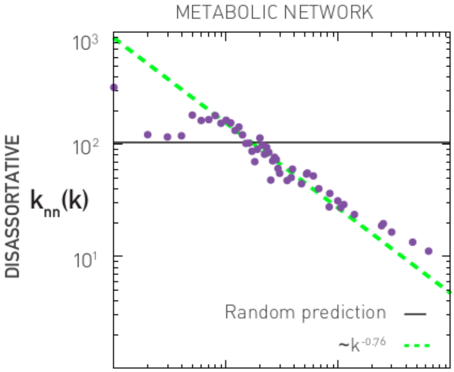
Measuring degree correlations

Disassortative networks

- In disassortative networks hubs tend to connect low-degree nodes
- Therefore, $k_{nn}(k)$ decreases with k, as observed for scientific collaboration networks

Measuring degree correlations

Disassortative networks



Measuring degree correlations

The degree correlation function can be approximated as:

$$k_{nn}(k) = ak^{\mu}$$

The nature of degree correlations is determined by the sign of the correlation exponent μ

- **Assortative networks:** $\mu > 0$
  For the science collaboration network $\mu = 0.37 \pm 0.11$
- **Neutral networks:** $\mu = 0$
  For the power grid $\mu = 0.04 \pm 0.05$
- **Disassortative networks:** $\mu < 0$
  For the metabolic network $\mu = -0.76 \pm 0.04$