BA and BB models

Vander Luis de Souza Freitas
vander.freitas@ufop.edu.br

UFOP

Agenda

- Last class we discussed:
  - Watts-Strogatz model
  - The Scale-Free Property
  - Generative algorithms

Agenda

- Today:
  - Discussion about the homework
  - Growth and Preferential Attachment
  - The Barabási-Albert Model
  - The Bianconi-Barabási Model
  - Directions and weights

Discussion

Ideas for the final project?

Growth and Preferential Attachment

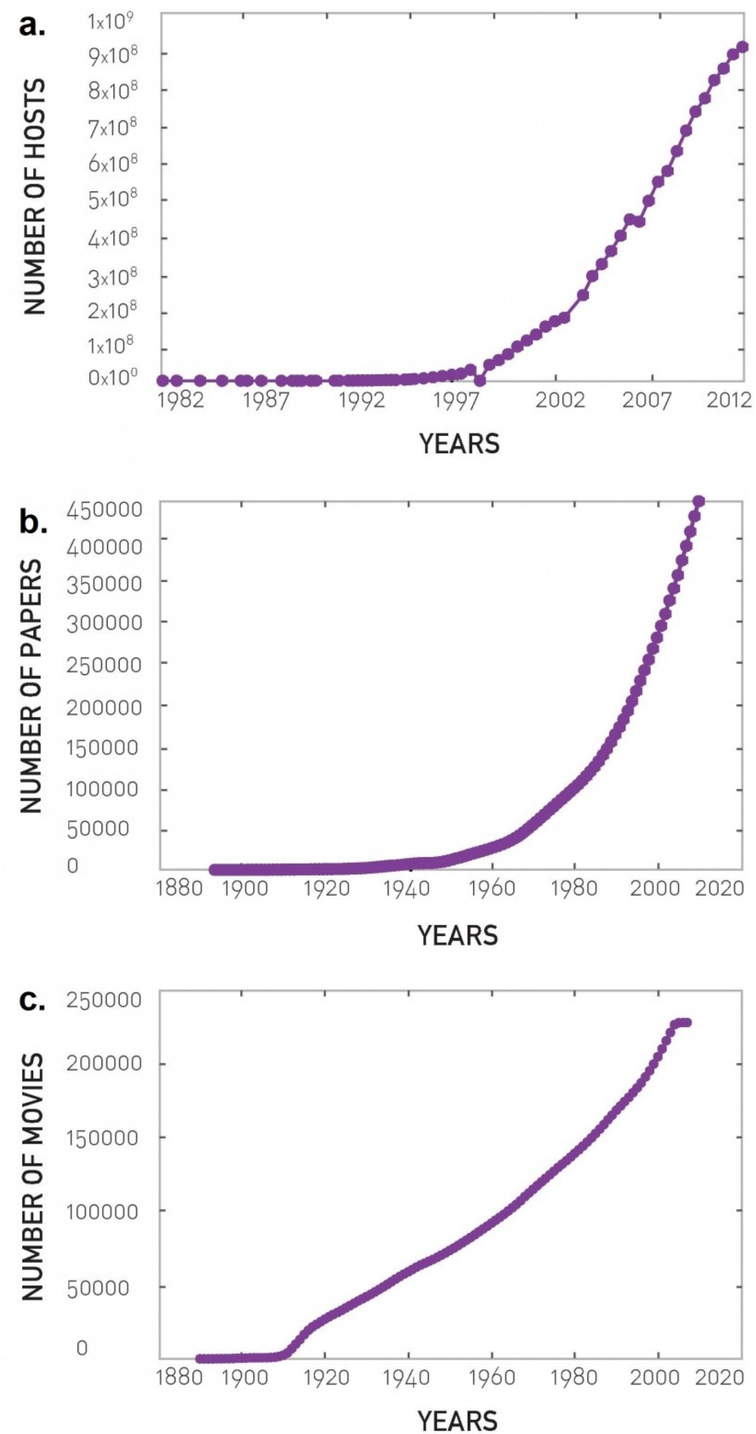
Growth: networks are not static, but grow via the addition of new nodes.

a. number of WWW hosts.

b. number of papers published in Physical Review.

c. Number of movies listed in IMDB.com.

Networks expand through the addition of new nodes



http://networksciencebook.com/chapter/5

Growth and Preferential Attachment

- **Preferential attachment:** nodes prefer to link to the more connected nodes.
    - Examples:
        - We all know Google and Facebook. New webpages usually have links to them.
        - A well-cited paper is more likely to be cited by new papers.
        - An actor that played in many movies is well known and has more chances to be in a new movie.

The random network model differs from real networks in two important characteristics: **Growth** and **Preferential Attachment**

http://networksciencebook.com/chapter/5

The Barabási-Albert Model

Also known as the BA model or the scale-free model.

We start with m_0 nodes and randomly chosen links. All nodes must have at least one link. Another possibility is to start with a fully connected network.

Growth:

- At each timestep we add a new node with $m(\leq m_0)$ links connecting to m nodes already in the network.

Preferential Attachment:

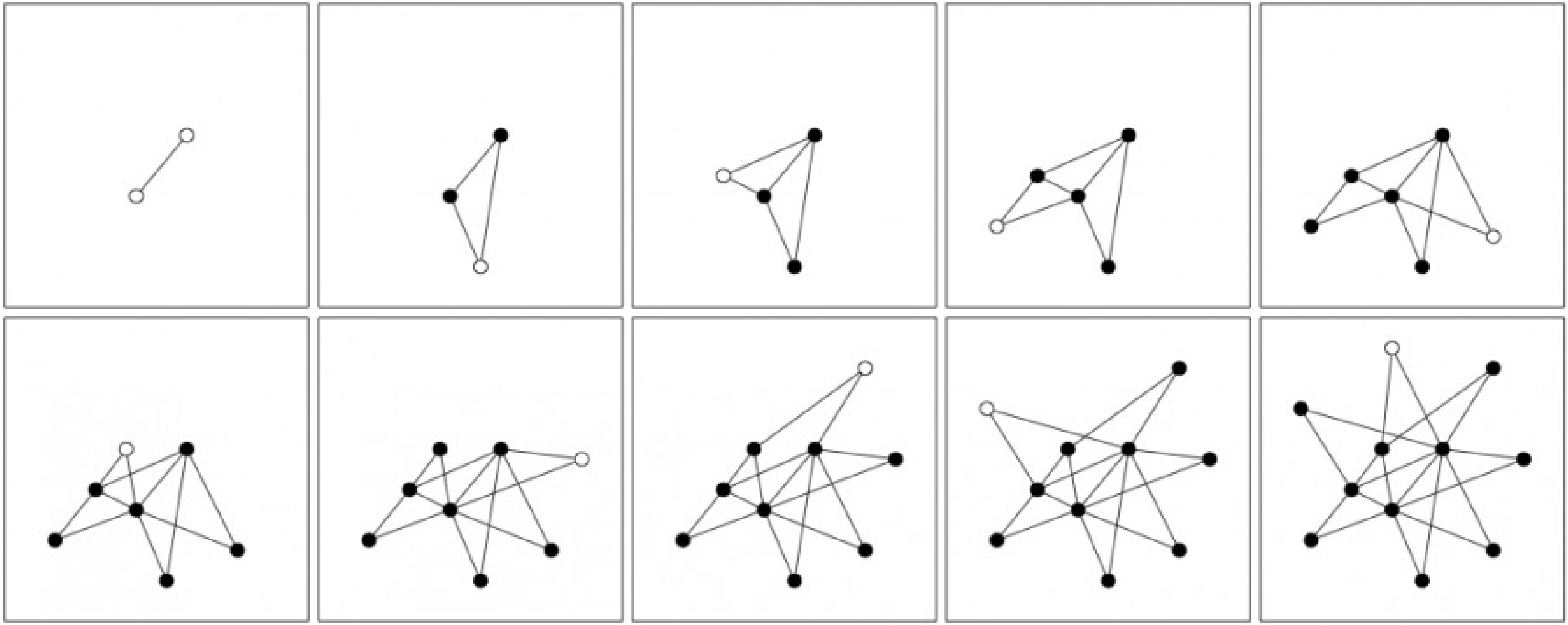
- The probability $\Pi(k)$ that a link of the new node connects to node *i* depends on the degree k_i as

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (1)$$

http://networksciencebook.com/chapter/5

The Barabási-Albert Model

Evolution of the Barabási-Albert Model



Initially: *N=2* nodes. Parameter: $m = 2$

Hubs are the result of the rich-gets-richer phenomenon

http://networksciencebook.com/images/ch-05/video-5-2.webm

http://networksciencebook.com/chapter/5

Degree dynamics

The rate at which an existing node *i* acquires links as a result of new nodes connecting to it is

$$\frac{dk_i}{dt} = m\Pi(k_i) = m\frac{k_i}{\sum_{j=1}^{N-1} k_j} \tag{2}$$

which leads to

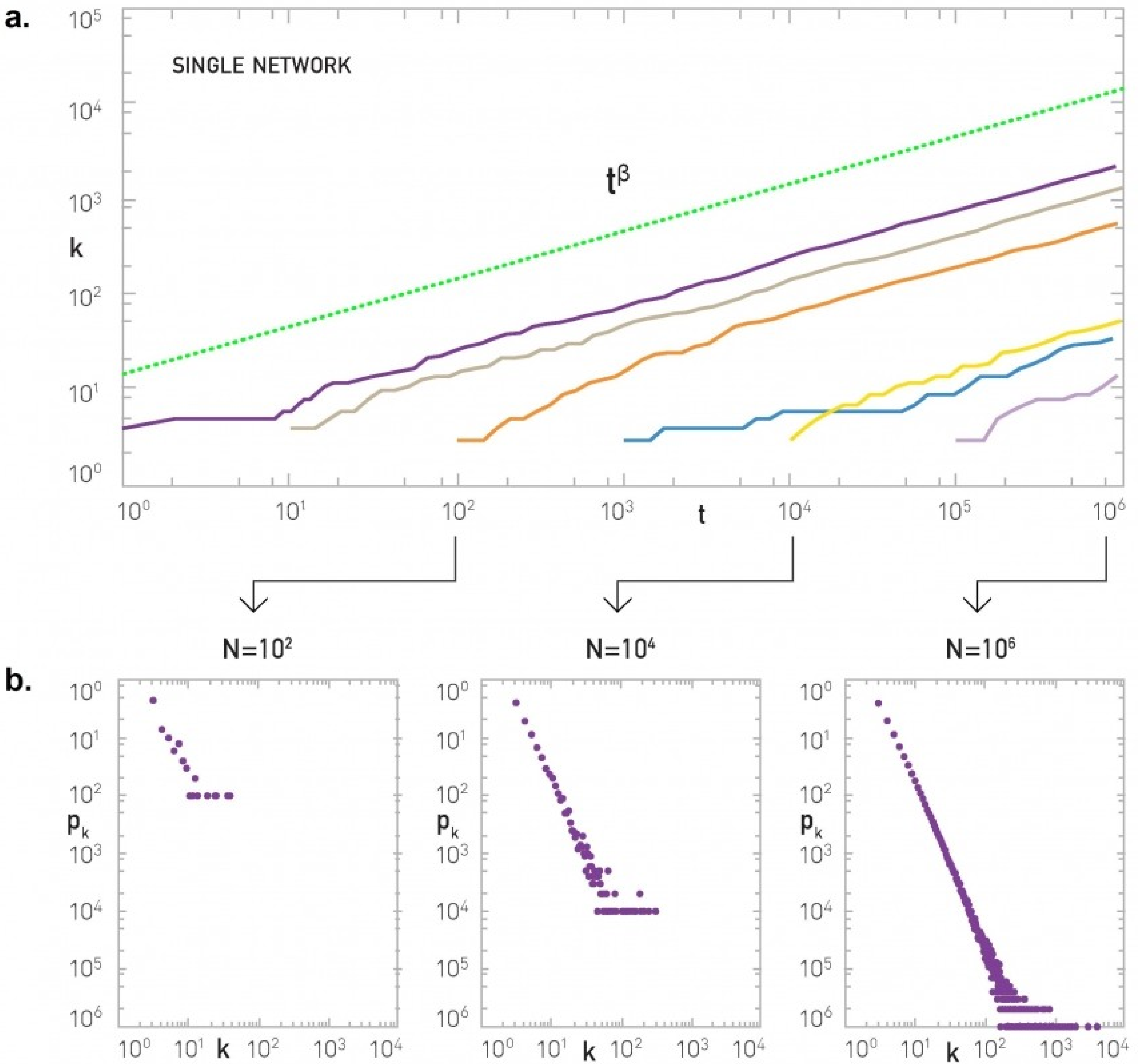
$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta} \tag{3}$$

in which t_i is the time when node *i* joins the network, with the *dynamical exponent*

$$\beta = \frac{1}{2}$$

The Barabási-Albert Model

Degree dynamics



http://networksciencebook.com/chapter/5

The Barabási-Albert Model

Degree distribution

Using the continuum theory

$$p(k) \approx 2m^{1/\beta} k^{-\gamma}$$

with

$$\gamma = \frac{1}{\beta} + 1 = 3$$

Thus, it relates network's topology and dynamics.

http://networksciencebook.com/chapter/5

The Barabási-Albert Model

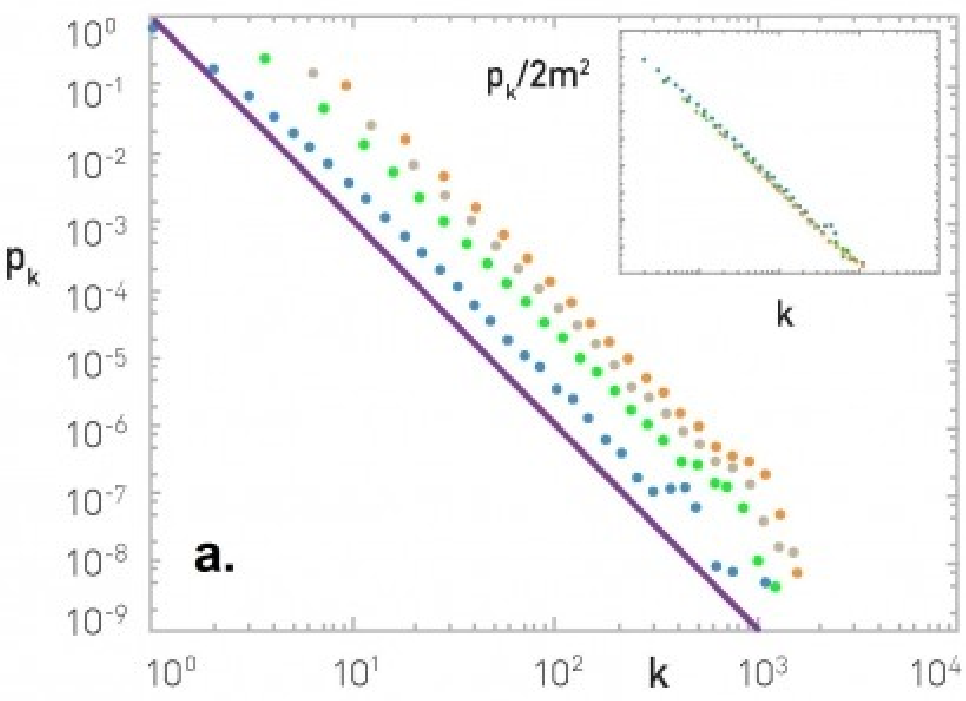
Degree distribution

a. networks with *N=100,000* and *m=1* (blue), *m=3* (green), *m=5* (grey), and *m=7* (orange). Purple line has slope *-3*. As the lines are parallel, γ is independent of m.

The exact degree distribution of the BA model:

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)}$$

The degree distribution is stationary, explaining why networks with different history, size and age develop a similar degree distribution.

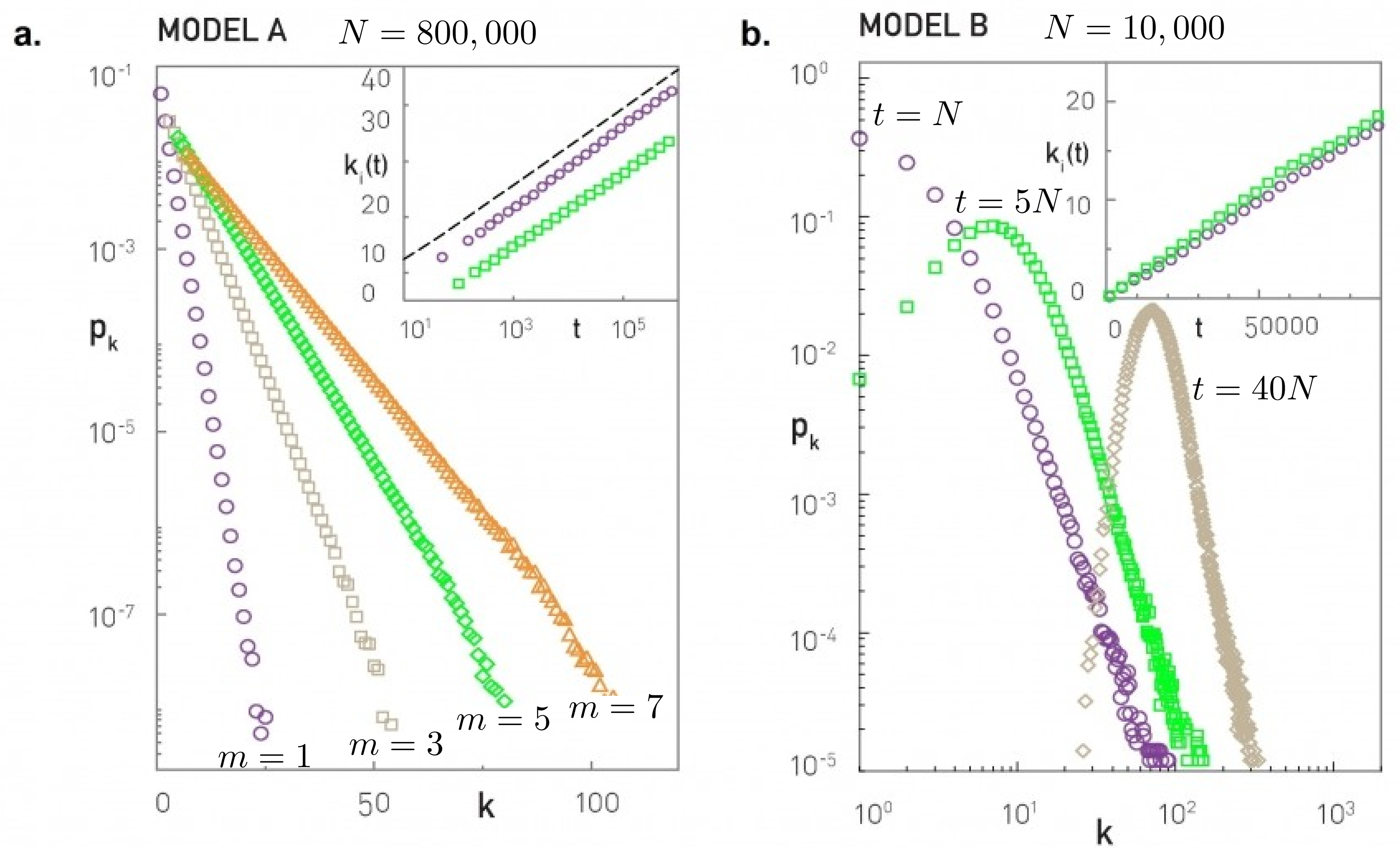


http://networksciencebook.com/chapter/5

Preferential Attachment and Growth are both important for the BA model

Without PA

Without Growth



http://networksciencebook.com/chapter/5

The Barabási-Albert Model

The BA model is unable to describe many characteristics of real systems:

- The model predicts $\gamma = 3$, while many real networks have $\gamma \in [2, 5]$.
- Directed networks, like the WWW.
- The disappearance of links and nodes.
- Intrinsic characteristics of the system, like the novelty of a paper or the utility of a webpage.

The AB model is minimal, a proof of principle.

http://networksciencebook.com/chapter/5

Evolving Networks

Late 1990s: Alta Vista and Inktomi have dominated the search market.

2000: Google became the biggest hub of the Web.

2011: Facebook took over as the Web's biggest node.

Evidences show that the growth rate of a node does not depend on its age alone (as in the BA model).

The Bianconi-Barabási Model

Some nodes have an intrinsic property (*fitness*) that propels them ahead of the pack.

The Bianconi-Barabási Model

Growth:
In each timestep a new node *j* with *m* links and fitness η_j is added to the network, where η_j is a random number chosen from a fitness distr. $\rho(\eta)$. Once assigned, a node's fitness does not change.

Preferential Attachment:
The probability that a link of a new node connects to node *i* is proportional to the product of node *i*'s degree k_i and its fitness η_i:

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}$$

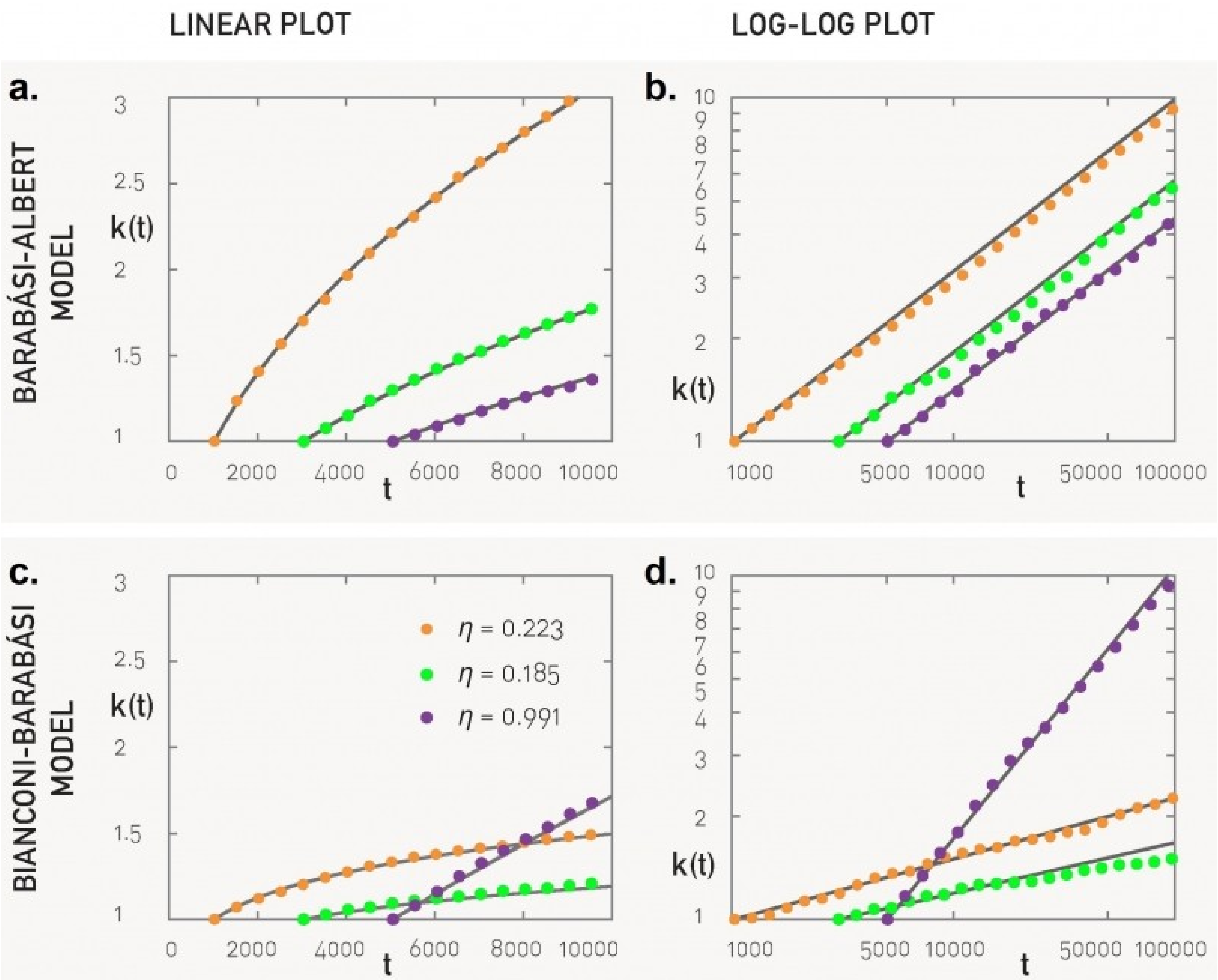
Video: http://networksciencebook.com/images/ch-06/video-6-1.webm

The Bianconi-Barabási Model

Degree dynamics

Unlike the BA model, here each node has its own dynamic exponent $\beta(\eta_i)$, that depends on η_i, such that

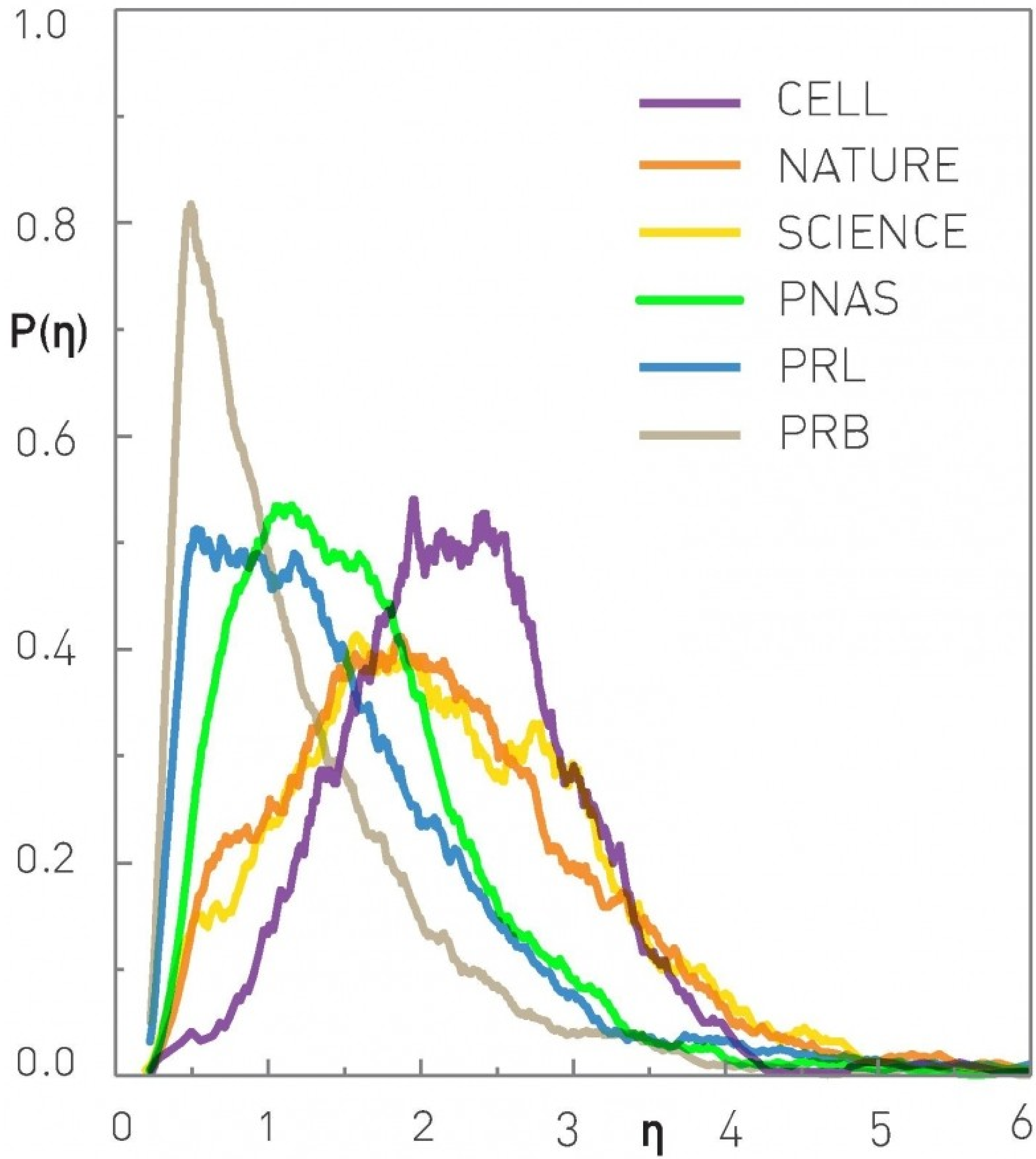
$$k(t, t_i, \eta_i) = m\left(\frac{t}{t_i}\right)^{\beta(\eta_i)}$$



http://networksciencebook.com/chapter/6

Measuring fitness

Example:
Fitness distribution of research papers.



http://networksciencebook.com/chapter/6

Attractiveness Model

If the initial condition for the BA model has isolated nodes, they will have zero probability of getting links. The attractiveness model overcomes this problem:

$$\Pi_i = \frac{A + k_i}{\sum_j (A + k_j)}$$

The case of $A=0$ yields the BA model.

For any value of A, it builds networks with heavy-tailed degree distributions, with the corresponding slope depending on A.

MENCZER, F.; FORTUNATO, S.; DAVIS, C. A First Course in Network Science. Cambridge: Cambridge University Press, 2020. 300 p.

Directions and weights

Directed networks (digraphs):

- WWW
- Twitter
- Instagram
- Wikipedia
- Citation network

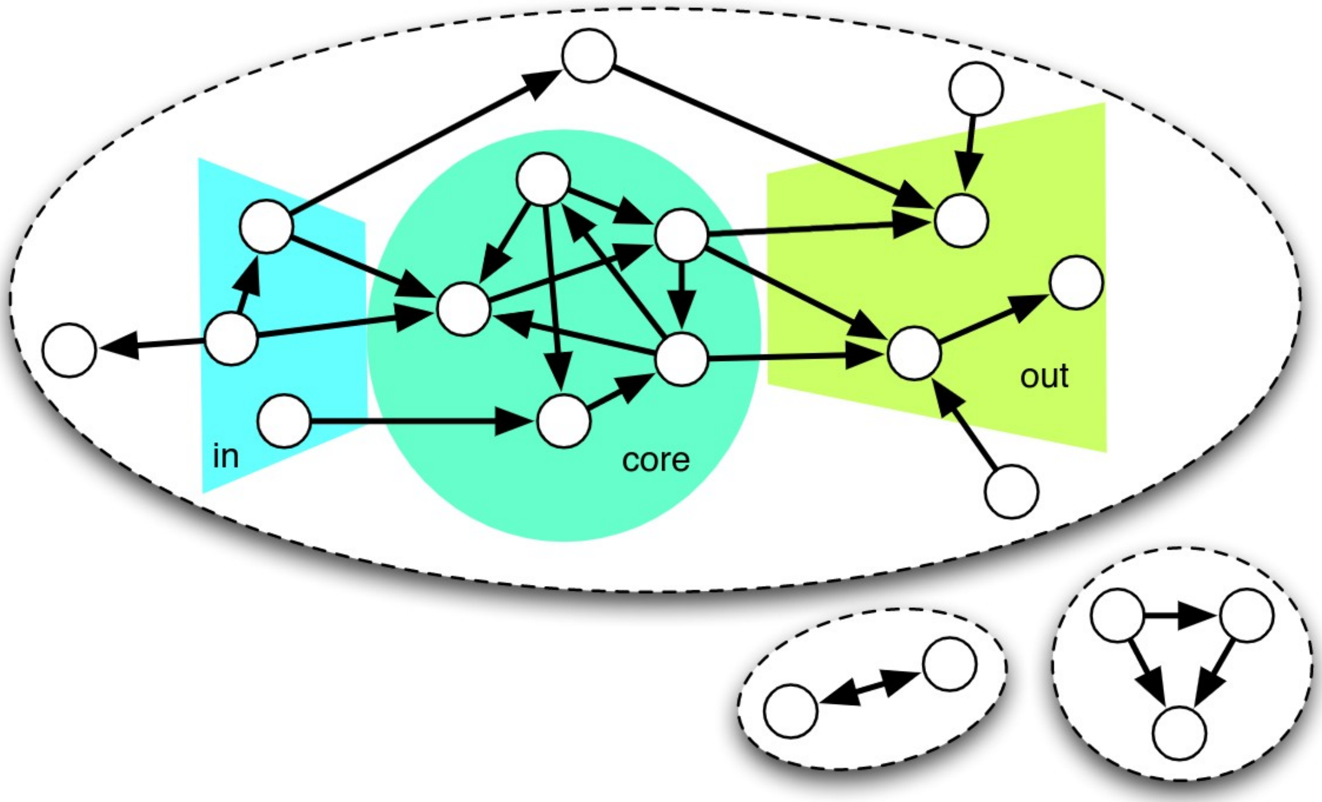


Fig. 4.4 The bow-tie structure of the Web graph. Components are highlighted by dashed ovals. The giant component has a giant strongly connected component (sometimes referred to as "core"), an in-component ("in"), and an out-component ("out").

https://github.com/CambridgeUniversityPress/FirstCourseNetworkScience/blob/master/sample/chapters/chapter4.pdf

Directions and weights

Co-occurrence networks

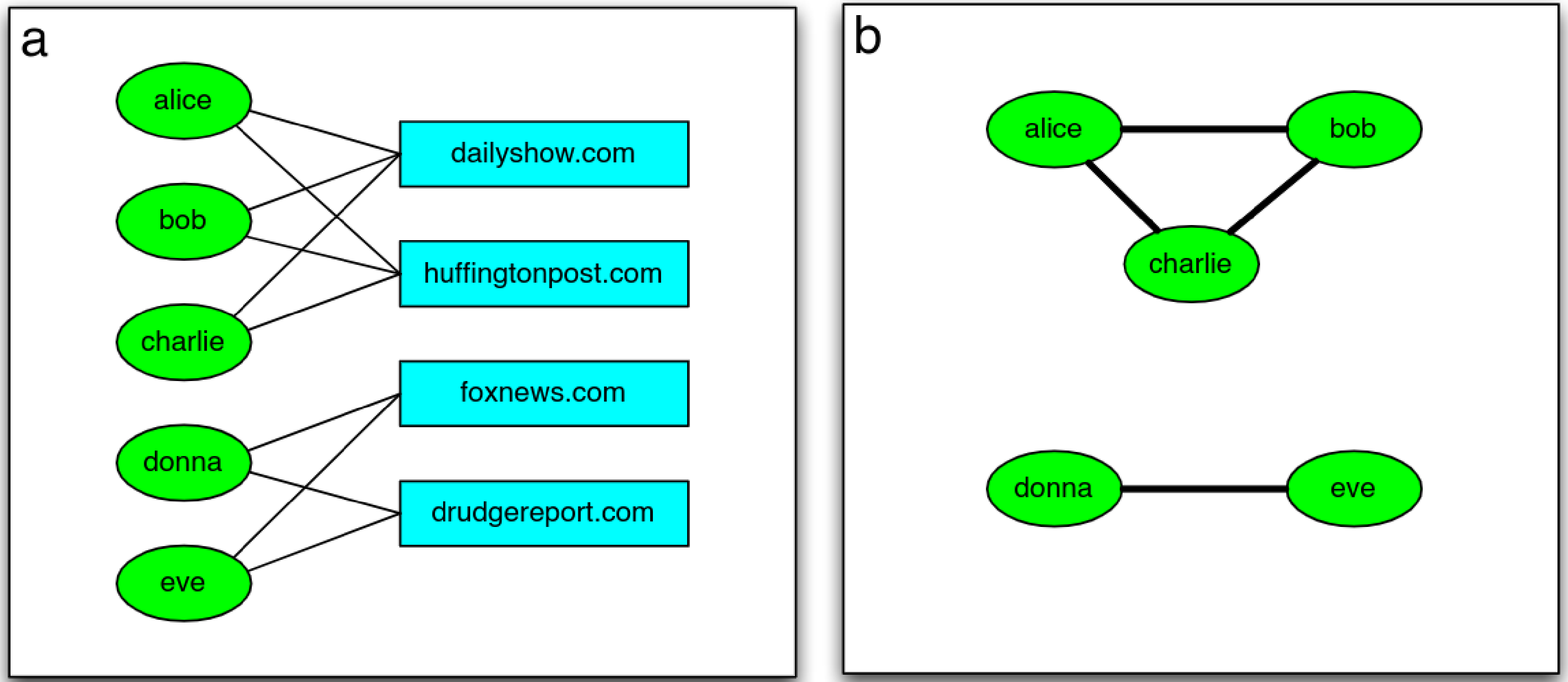


Fig. 4.14 (a) A bipartite network induced by "like" relationships. (b) A user co-occurrence network derived from projecting the "like" network onto user nodes.

https://github.com/CambridgeUniversityPress/FirstCourseNetworkScience/blob/master/sample/chapters/chapter4.pdf

Directions and weights

Co-occurrence networks

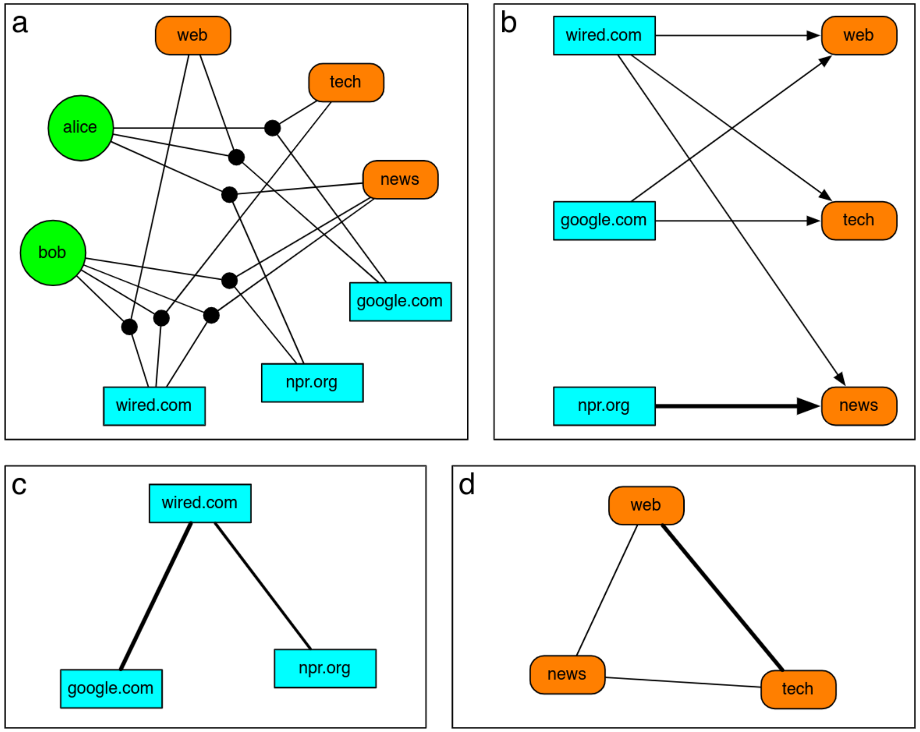


Fig. 4.15 Example of a folksonomy and derived bipartite and co-occurrence networks. (a) Two users (`alice` and `bob`) annotate three resources (`npr.org`, `wired.com`, `google.com`) using three tags (`news`, `web`, `tech`), resulting in seven triples. (b) Projecting the folksonomy onto resources and tags, we obtain a bipartite network. Link weights correspond to numbers of triples, or numbers of users. The link from `npr.org` to `news` has a larger weight because both users agree on that annotation. (c) A resource co-occurrence network. The resources `wired.com` and `google.com` are more similar because they co-occur with two tags, `web` and `tech`. (d) A tag co-occurrence network. The link between `web` and `tech` has a larger weight because of the similarity in their resources: the two tags co-occur with two resources, `wired.com` and `google.com`.

https://github.com/CambridgeUniversityPress/FirstCourseNetworkScience/blob/master/sample/chapters/chapter4.pdf

Directions and weights

Link filtering

Box 4.3 **Network backbone**

In networks with broad distributions of link weights, using a global threshold to prune links is inappropriate. Instead, we can use the weight fluctuations for each node to identify the links to be preserved — those that carry most of the weight. Given a node i with degree k_i and strength s_i, let us evaluate a link against a null model in which the weights are distributed randomly on the k_i links adjacent to i, with the constraint that their sum equals s_i. The probability that a link has weight w_{ij} or larger under this hypothesis is:

$$p_{ij} = \left(1 - \frac{w_{ij}}{s_i}\right)^{k_i - 1}. \tag{4.2}$$

So, if link ij has weight w_{ij} from Eq. 4.2 we compute the probability p_{ij} that such value is compatible with the null model: if $p_{ij} < \alpha$, where α is a parameter that represents the desired significance level, the link is preserved, otherwise it is removed. Lower values of α lead to sparser networks, as fewer links are preserved. Since a link is connected to two nodes, we can obtain two values for p_{ij} by plugging the strength and degree of either node into Eq. 4.2. We can then use the larger or smaller of these values, depending on whether we wish to prune more or less aggressively. This link filtering procedure extracts a *network backbone*, which is supposed to preserve the essential structure and global properties of the network.

https://github.com/CambridgeUniversityPress/FirstCourseNetworkScience/blob/master/sample/chapters/chapter4.pdf

Directions and weights

Network backbone

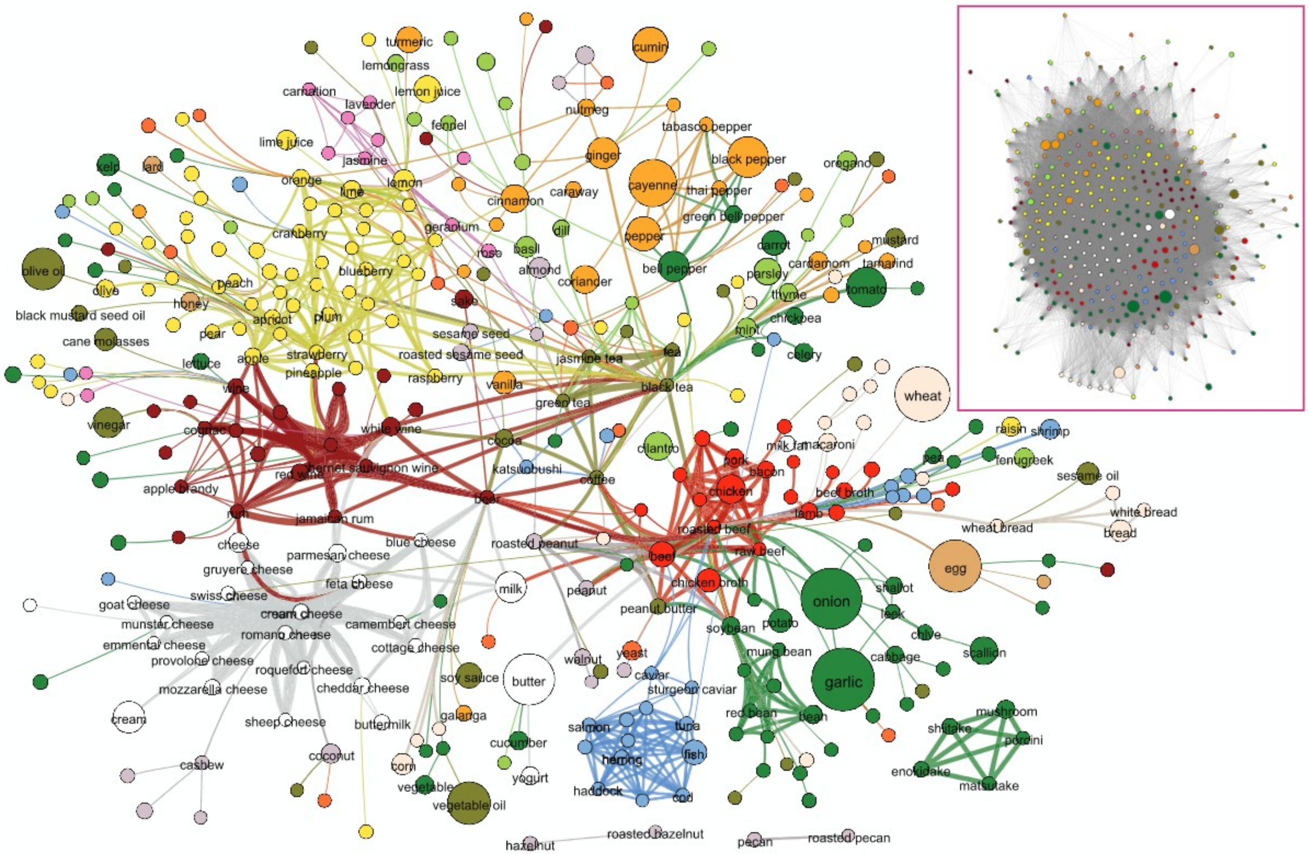


Fig. 4.18 A flavor network: each node denotes an ingredient, its color indicates a food category, and its size reflects the ingredient prevalence in recipes. Two ingredients are connected if they share flavor compounds, link width representing the number of shared compounds. The full network is shown in the inset, while the main image visualizes the backbone network with the significant links identified by the method in Box 4.3, using $\alpha = 0.04$. Images adapted from Ahn et al. (2011) under CC BY 4.0 license.

https://github.com/CambridgeUniversityPress/FirstCourseNetworkScience/blob/master/sample/chapters/chapter4.pdf

Global metric

Heterogeneity parameter:

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle^2}$$

For a normal or narrow distribution with a sharp peak at some value:

$$\kappa \approx 1$$

For a heavy-tailed distribution: $\kappa \gg 1$