Chapter 5 of: Social and Economic Network / Matthew O. Jackson.

Growing Random Networks

Growing number of nodes over time.

- Web.
- Scientific networks.
- Citation Networks.
- Human societies.

Why growing random networks.

- Dynamisms.
- Heterogeneity.
- Varying degree distribution.

Random Network Model: two extremes.

Every node i is born at time i. A newborn node forms link(s) to elder node(s).

Newborn nodes uniformly randomly select other nodes to link to.	Newborn nodes select other nodes to link to based on their degree at time.

Random Network Model: two extremes.

Every node *i* is born at time *i*. A newborn node forms link(s) to elder node(s).

Newborn nodes uniformly randomly select other nodes to link to. **Growing version of Erdös-Renyi model.**	Newborn nodes select other nodes to link to based on their degree at time. **Preferential Attachment Model.**

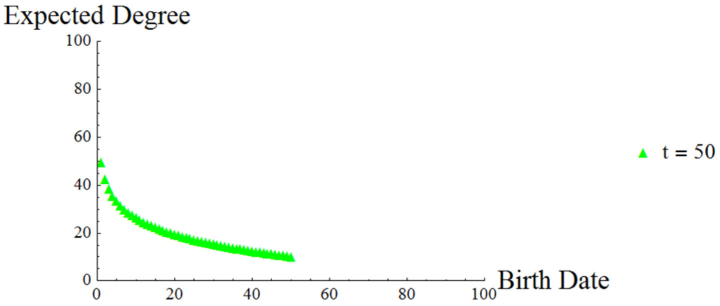
Growing Erdos-Renyi Model

- Start with m nodes fully connected.
- New node forms m links to existing nodes.
- $P\,[\text{an existing node getting a new link}] = \frac{m}{t}$.
- For node $m < i < t$:

$$E[d_i] = \underbrace{m + \frac{m}{i+1} + \frac{m}{i+2} + \cdots + \frac{m}{t}}_{\sim m\left(1+\log\left(\frac{t}{i}\right)\right)}$$

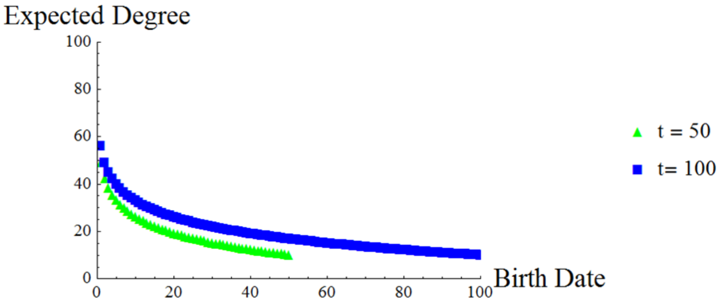
Expected Degree Distribution

Let $m = 10$.



Expected Degree Distribution

Let $m = 10$.

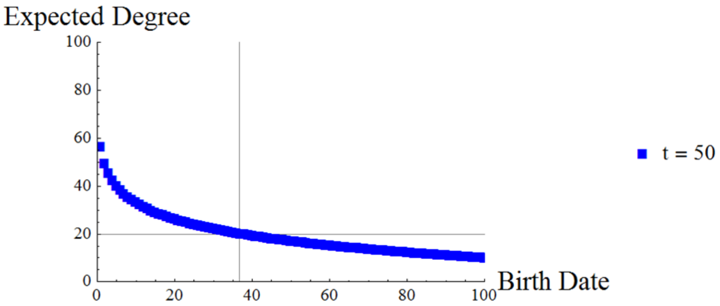


When $t = 100$, what is the fraction of nodes with degree ≤ 20 ($F(20)$).

$$10\left(1+\log\left(\frac{100}{i}\right)\right) < 20 \quad \Longrightarrow \; i > 100e^{-\frac{(20-10)}{10}}$$
$$\Longrightarrow \; i \sim 36.7$$

Thus the fraction of nodes with degree less than 20 is:

$$\frac{100-36.7}{100} = 0.633.$$


Expected Degree
100
80
60
40
20
0
0
20
40
60
80
100
Birth Date
t = 50

Nodes that have expected degree less than d at t:

$$i > te^{-\frac{d-m}{m}}$$

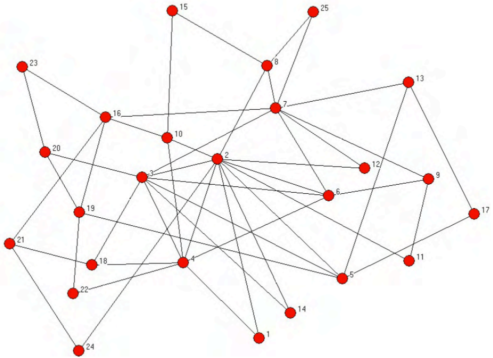
Expected degree distribution at time t:

$$F_t(d) = \left(\frac{t - te^{\frac{d-m}{m}}}{t} \right) = 1 - e^{\frac{m-d}{m}}$$

Actual degree distribution: $F_t(d)$ turns out to be a good approximation of actual degree distribution (bearing some conditions).

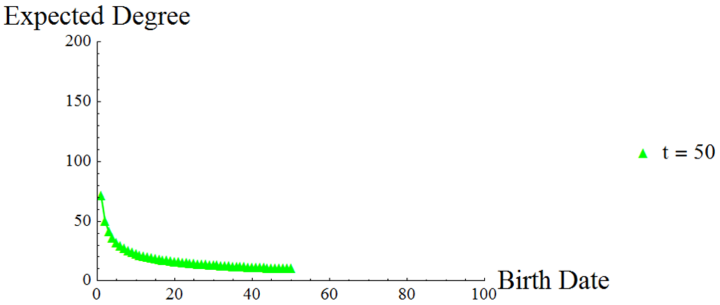
Preferential Attachment

- It enables us to get power law distributions (fat-tail).
- Realistic scenarios for power-law explanation:
  - Rich get Richer.
  - New objects enter over time.



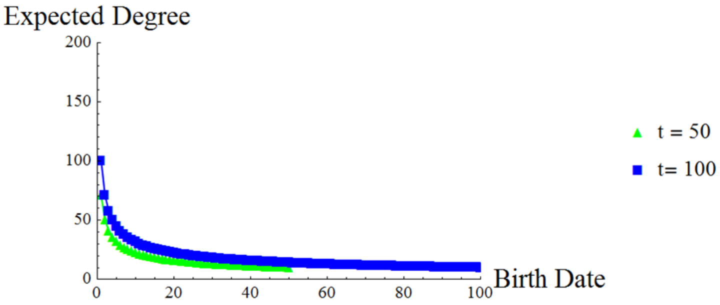
- There exist tm links at time t.
- P [connecting to node i] $= \frac{d_i(t)}{2tm}$.
- Continuous approximation!
- $d_i(t) = m\left(\frac{t}{i}\right)^{\frac{1}{2}}$

Let $m = 10$.

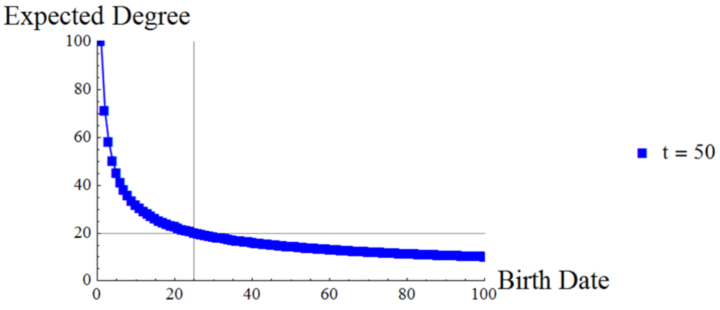


- There exist tm links at time t.
- $P\left[\text{connecting to node } i\right] = \frac{d_i(t)}{2tm}$.
- Continuous approximation!
- $d_i(t) = m\left(\frac{t}{i}\right)^{\frac{1}{2}}$

Let $m = 10$.

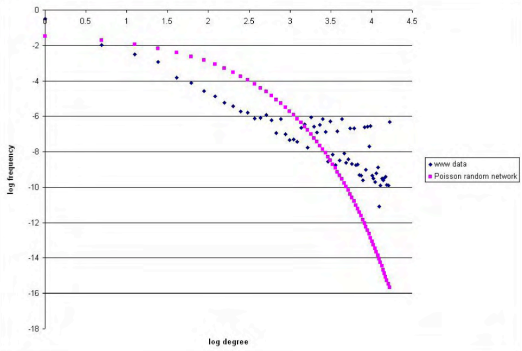


Similar to the first model (newborns form links uniformly at random), when $t = 100$, what is the fraction of nodes with degree ≤ 20 ($F(20)$)?

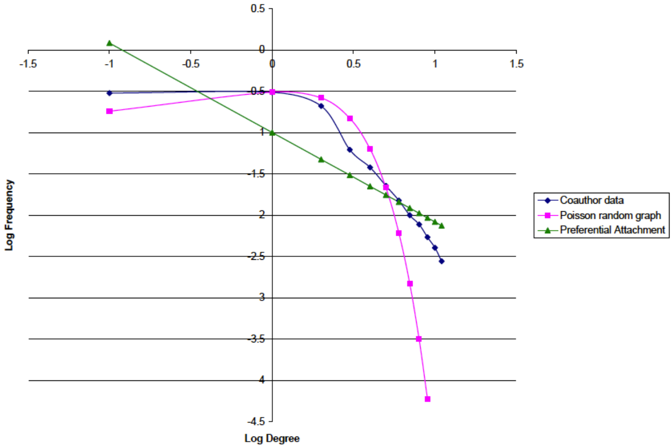


$$F_t(d) = \frac{\left(t - tm^2/d^2\right)}{t} = 1 - m^2/d^2 \implies f_t(d) = 2m^2/d^3.$$

Why is $F_t(d) = 1 - m^2/d^2$ more interesting than $F_t(d) = 1 - e^{\frac{m-d}{m}}$?



But is this always true?



Hybrid Model

- For a fraction of α, a new born links to α uniformly at random, and via searching neighbourhood for $1-\alpha$.
- Expected degree distribution:

$$1-\left(\frac{m+\frac{2\alpha m}{1-\alpha}}{d+\frac{2\alpha m}{1-\alpha}}\right)^{\frac{2}{1-\alpha}}$$

- When $\alpha \to 0: F_t(d) = 1-\left(\frac{m}{d}\right)^2$
- When $\alpha \to 1: F_t(d) \to 1-e^{\frac{med}{m}}$

Finding m and α: fitting to data.

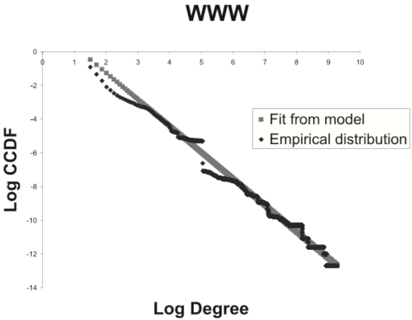
$$\log(1-F(d)) = c - \frac{2}{1-a}\log\left(d + am\frac{2}{1-a}\right)$$

One can estimate $\frac{2}{1-\alpha}$ by regression, or simply consider a grid in $\alpha \in \{0+\varepsilon, 2\varepsilon, \cdots, 1-\varepsilon\}$.

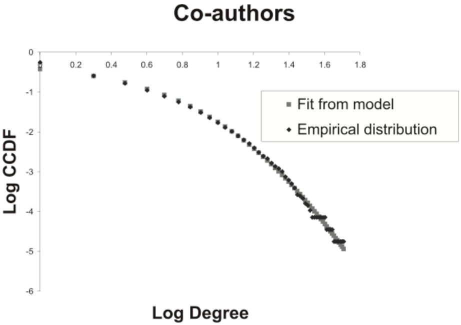
Or maybe solve:

$$\begin{aligned} \min \quad & \int_{-\infty}^{\infty} \left(F(d) - F_{\alpha(\beta)}(d)\right)^2 \mathbf{d}\beta \\ \text{s.t} \quad & \alpha(\beta) \in [0,1] \end{aligned} \tag{1}$$

Examples:

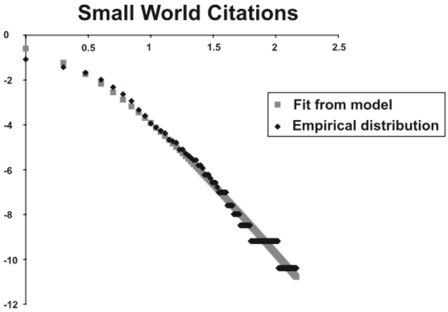

WWW
0
1
2
3
4
5
6
7
8
9
10
0
-2
-4
-6
-8
-10
-12
-14
Fit from model
Empirical distribution
Log CCDF
Log Degree

Examples:

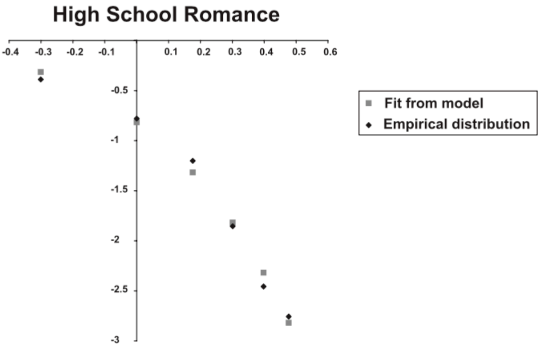


Co-authors
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
-1
-2
-3
-4
-5
-6
Log CCDF
Log Degree
Fit from model
Empirical distribution

Examples:


Small World Citations
0
-2
-4
-6
-8
-10
-12
0.5
1
1.5
2
2.5
Fit from model
Empirical distribution

Examples:



High School Romance
-0.4
-0.3
-0.2
-0.1
0.1
0.2
0.3
0.4
0.5
0.6
-0.5
-1
-1.5
-2
-2.5
-3
Fit from model
Empirical distribution