Random Graphs and Complex Networks

Kaleidoscoopdag, Leiden

Remco van der Hofstad

May 16, 2007

Joint work with
- Henri van den Esker (TU Delft)
- Mia Deijfen (Stockholm)
- Gerard Hooghiemstra (TU Delft)
- Piet Van Mieghem (TU Delft)
- Dmitri Znamenski (Philips Research)

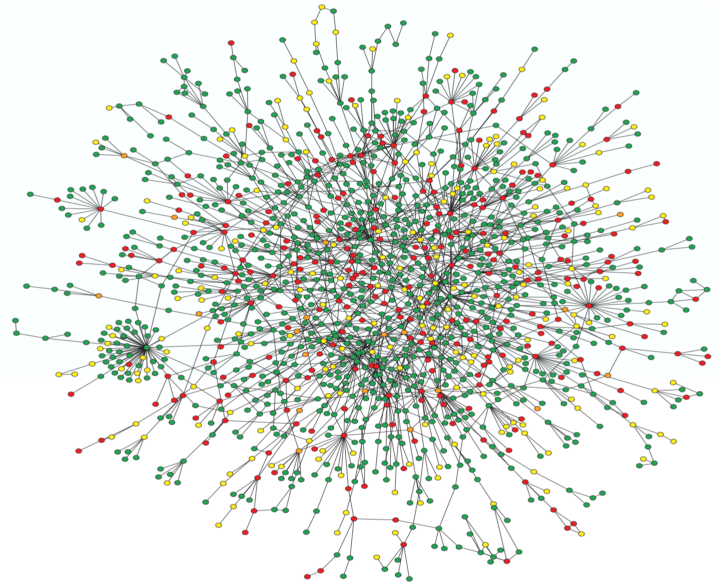
Six Degrees of Separation

"Everybody on this planet is separated only by six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice... It's not just the big names. It's anyone. A native in the rain forest. (...) An Eskimo. I am bound to everyone on this planet by a trail of six people. It is a profound thought."

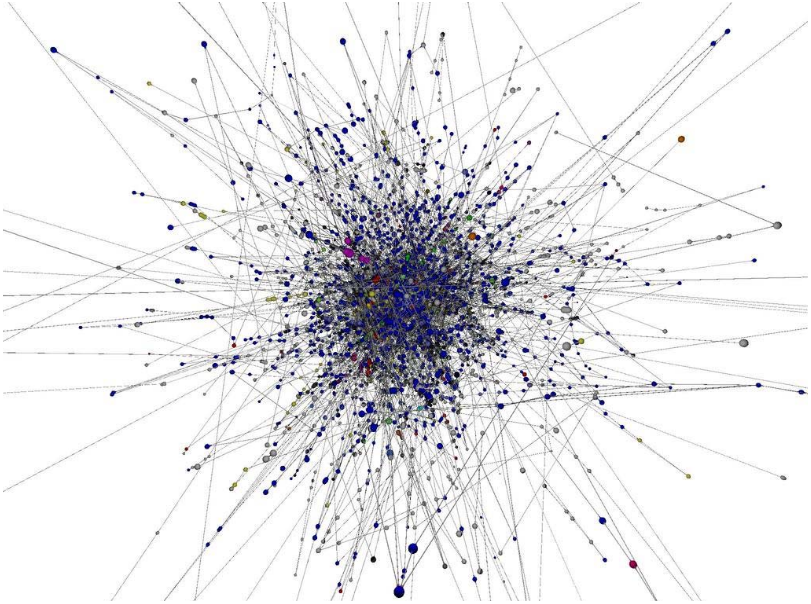
Small Worlds

Small Worlds

Scale-Free



Yeast protein interaction network
(Jeong, Mason, Barabási, and Oltvai (2001))



Internet topology in 2001
(http://www.fractalus.com/steve/stuff/ipmap/)

Real Networks

A notorious problem with many real networks, in particular social networks, is that they are

incredibly hard to measure!

Real Networks

A notorious problem with many real networks, in particular social networks, is that they are

incredibly hard to measure!

Restrict to graphs of which extensive measurements are available: movie actor network, collaboration networks, Internet, WWW, etc...

Real Networks

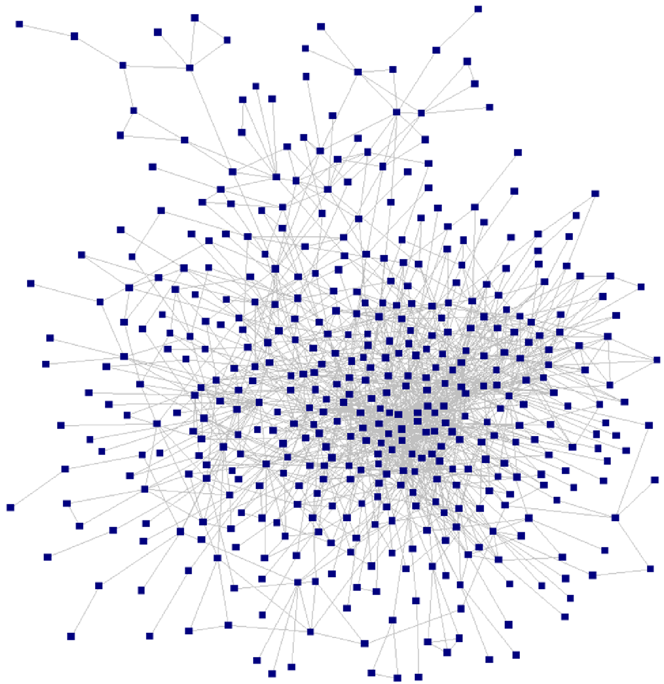
A notorious problem with many real networks, in particular social networks, is that they are

incredibly hard to measure!

Restrict to graphs of which extensive measurements are available: movie actor network, collaboration networks, Internet, WWW, etc...

Focus as a primary example on Collaboration graph in mathematics.

http://www.ams.org/mathscinet/otherTools.html



Collaboration graph in mathematics
(`http://www.orgnet.com/Erdos.html`)

Power-law degree sequences

Degree sequence $(f(1), f(2), f(3), \ldots)$ of graph:

$f(1)$ is number of elements with degree equal to 1,
$f(2)$ is number of elements with degree equal to 2, ...

Power-law degree sequences

Degree sequence $(f(1), f(2), f(3), \ldots)$ of graph:

$f(1)$ is number of elements with degree equal to 1,
$f(2)$ is number of elements with degree equal to 2, ...

$$f(k) = Ck^{-\tau}.$$

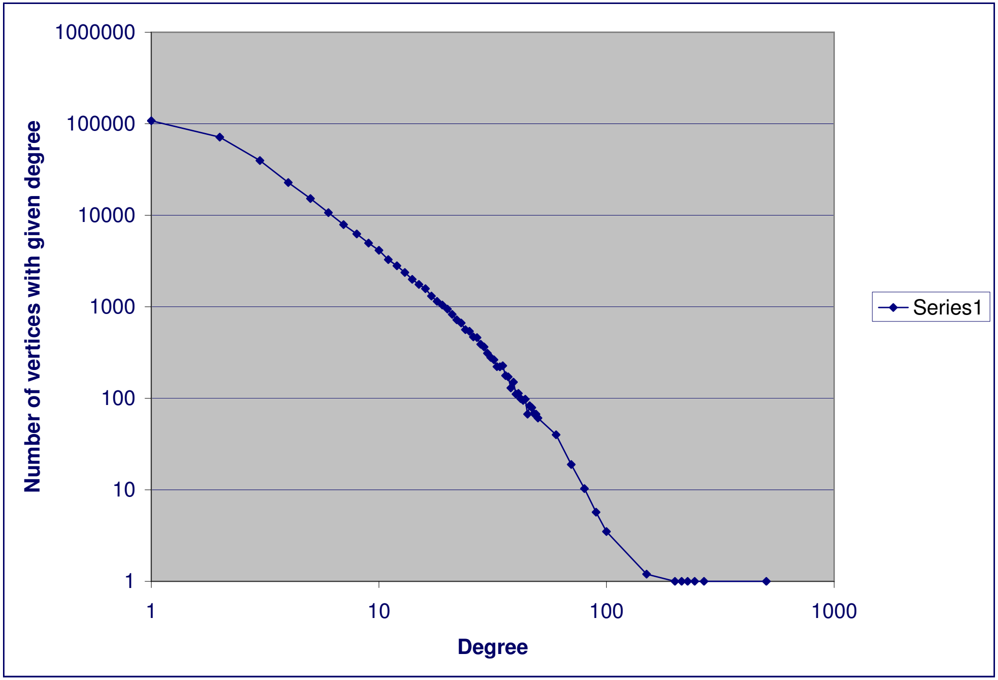
Power-law degree sequences

Degree sequence $(f(1), f(2), f(3), \ldots)$ of graph:

$f(1)$ is number of elements with degree equal to 1,
$f(2)$ is number of elements with degree equal to 2, ...

$$f(k) = Ck^{-\tau}.$$

Is dubbed scale-free phenomenon.



Degree sequence of collaboration graph among mathematicians
(`http://www.oakland.edu/enp`)

Power law random graph: Configuration Model

Let N be number of vertices. Consider an i.i.d. sequence of degrees $D_1, D_2, \ldots, D_N$, with

$$\mathbb{P}(D_1 = k) = c_\tau k^{-\tau}$$

where c_τ is normalizing constant, and $\tau > 1$.

Power law random graph: Configuration Model

Let N be number of vertices. Consider an i.i.d. sequence of degrees $D_1, D_2, \ldots, D_N$, with

$$\mathbb{P}(D_1 = k) = c_\tau k^{-\tau}$$

where c_τ is normalizing constant, and $\tau > 1$.

Power law degree sequence built in in model!

Graph construction

How to construct graph with above degree sequence?
Assign to vertex j degree D_j.

$$L_N = \sum_{i=1}^{N} D_i$$

is total degree. Assume L_N is even.

Graph construction

How to construct graph with above degree sequence?
Assign to vertex j degree D_j.

$$L_N = \sum_{i=1}^{N} D_i$$

is total degree. Assume L_N is even.

Now have N separate vertices and incident to vertex i have D_i 'stubs' or half edges. Connect stubs to create edges as follows:

Number stubs from 1 to L_N in any order.
First connect first stub at random with one of *other* $L_N - 1$ stubs.
Continue with second stub (when not connected to first) and so on, until all stubs are connected...

Scaling in power law random graph

Hopcount H_N is distance between uniform pair of nodes in graph.

Scaling in power law random graph

Hopcount H_N is distance between uniform pair of nodes in graph.

- When $\tau > 3$ and $\nu > 1$ (HHV03)

$$H_N \approx \log_\nu N, \qquad \text{with} \qquad \nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}.$$

Scaling in power law random graph

Hopcount H_N is distance between uniform pair of nodes in graph.

- When $\tau > 3$ and $\nu > 1$ (HHV03)

$$H_N \approx \log_\nu N, \qquad \text{with} \qquad \nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}.$$

- When $\tau \in (2,3)$, (Norros and Reittu, HHZ05)

$$H_N \approx 2\frac{\log\log N}{|\log(\tau-2)|}.$$

Scaling in power law random graph

Hopcount H_N is distance between uniform pair of nodes in graph.

- When $\tau > 3$ and $\nu > 1$ (HHV03)

$$H_N \approx \log_\nu N, \qquad \text{with} \qquad \nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}.$$

- When $\tau \in (2,3)$, (Norros and Reittu, HHZ05)

$$H_N \approx 2\frac{\log\log N}{|\log(\tau-2)|}.$$

- When $\tau \in (1,2)$, (HHZ04)

$$H_N \quad \text{uniformly bounded}.$$

Consequences

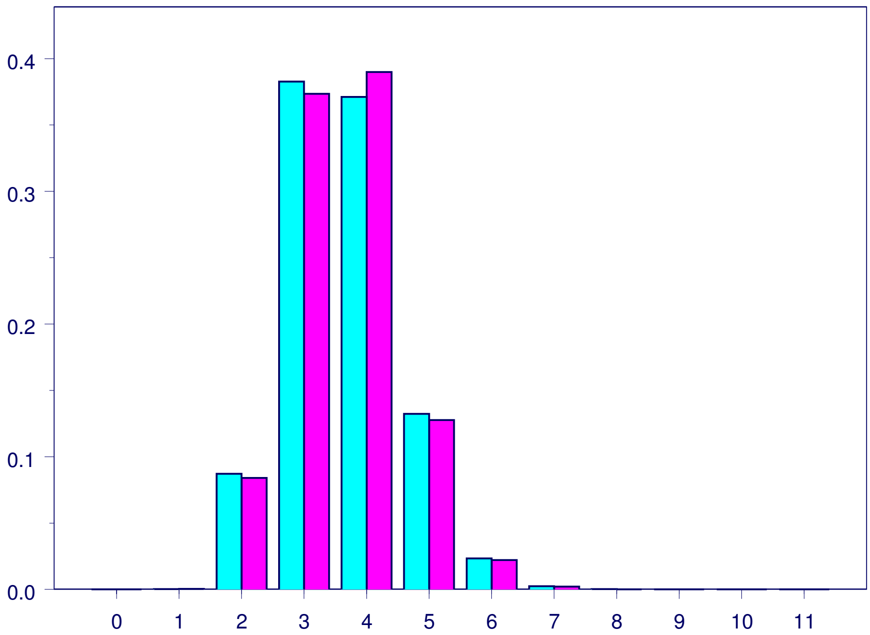
Above result explains small-world phenomenon, i.e., distances in graphs are (ultra-)small.

Consequences

Above result explains small-world phenomenon, i.e., distances in graphs are (ultra-)small.

Extension: Fluctuations around leading order are uniformly bounded, and we compute its distribution.

Allows us to compare distances to measurements.



Number of AS traversed in hopcount data (blue) compared to the model (purple) with $\tau = 2.25, N = 10,940.$

Preferential Attachment

Configuration model yields satisfactory model for complex network, but does not explain the reason for scale-free nature networks.

Preferential Attachment

Configuration model yields satisfactory model for complex network, but does not explain the reason for scale-free nature networks.

Preferential attachment offers convincing explanation.

In preferential attachment models, network is growing in time, in such a way that new vertices are more likely to be connected to vertices that already have high degree.

Preferential Attachment

Configuration model yields satisfactory model for complex network, but does not explain the reason for scale-free nature networks.

Preferential attachment offers convincing explanation.

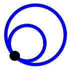
In preferential attachment models, network is growing in time, in such a way that new vertices are more likely to be connected to vertices that already have high degree.

Each of m edges of t^{th} vertex connects independently. Probability it connects to the i^{th} vertex equals

$$\frac{d_i(t-1)+\delta}{(2m+\delta)(t-1)},$$

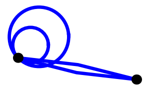
where $\delta \geq m$ is parameter model.

Preferential Attachment



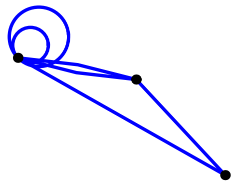
$(m = 2, \delta = 0)$

Preferential Attachment



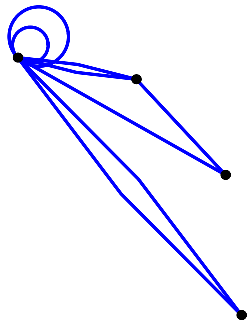
$(m = 2, \delta = 0)$

Preferential Attachment



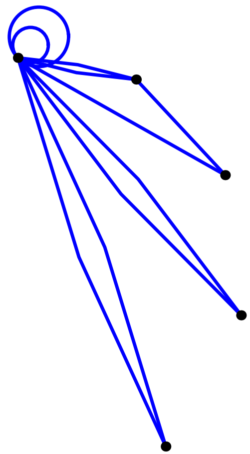
$(m = 2, \delta = 0)$

Preferential Attachment



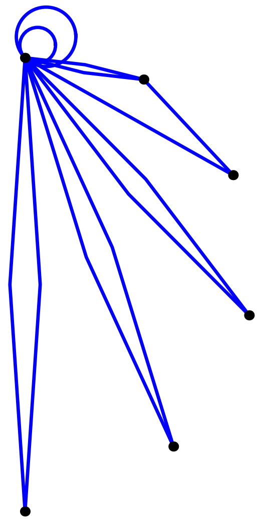
$(m = 2, \delta = 0)$

Preferential Attachment



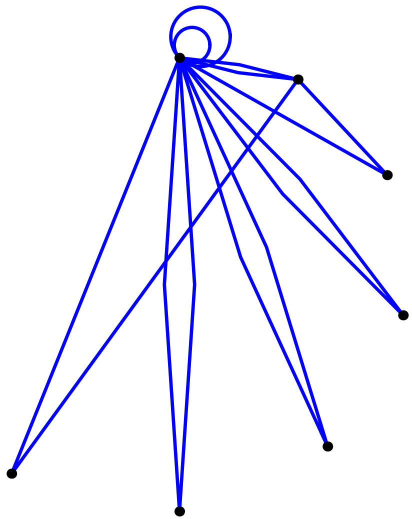
$(m = 2, \delta = 0)$

Preferential Attachment



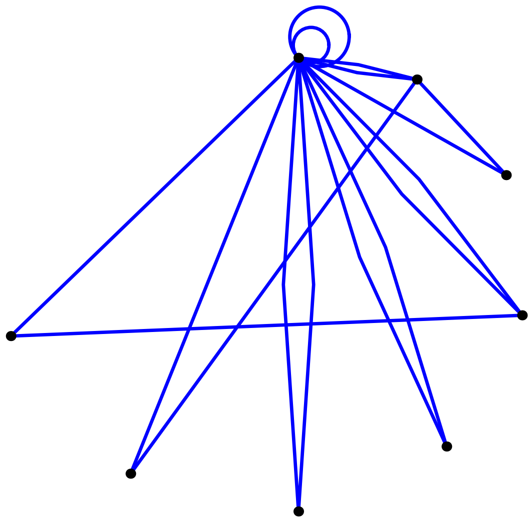
$(m = 2, \delta = 0)$

Preferential Attachment



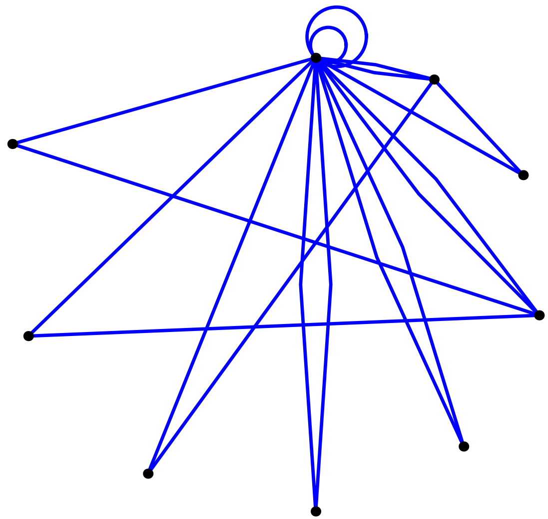
$(m = 2, \delta = 0)$

Preferential Attachment



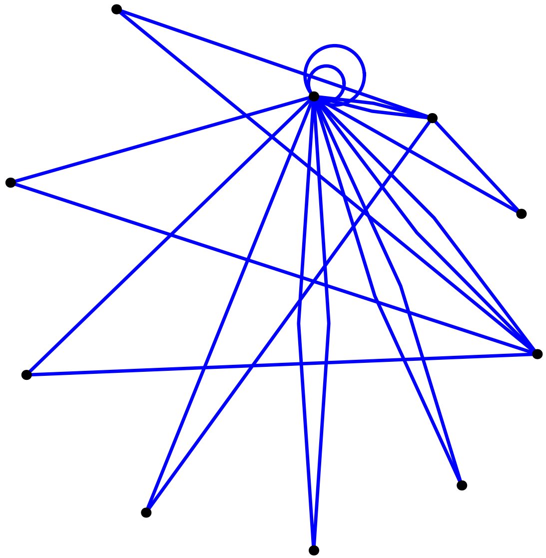
$(m = 2, \delta = 0)$

Preferential Attachment



$(m = 2, \delta = 0)$

Preferential Attachment

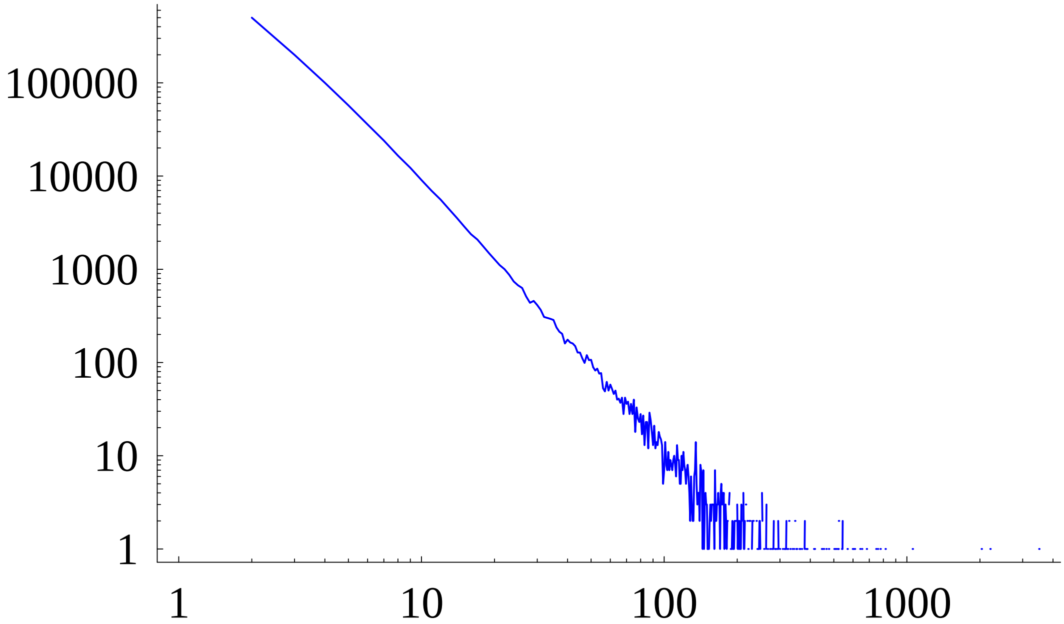


$(m = 2, \delta = 0)$

Preferential Attachment

$(m = 2, \delta = 0)$

Scale-Free Nature



$$(m = 2, \delta = 0, \tau = 3 + \frac{\delta}{m} = 3)$$

"...the scale-free topology is evidence of organizing principles acting at each stage of the network formation. (...) No matter how large and complex a network becomes, as long as preferential attachment and growth are present it will maintain its hub-dominated scale-free topology."

Challenges PA models

Non-rigorous literature predicts that scaling distances in preferential attachment models similar to the one in configuration model with equal power-law exponent degrees.

No rigorous foundation for this prediction!
Strong form of universality.