A Scale-Free Network Model with Arbitrary Clustering

Petter Holme

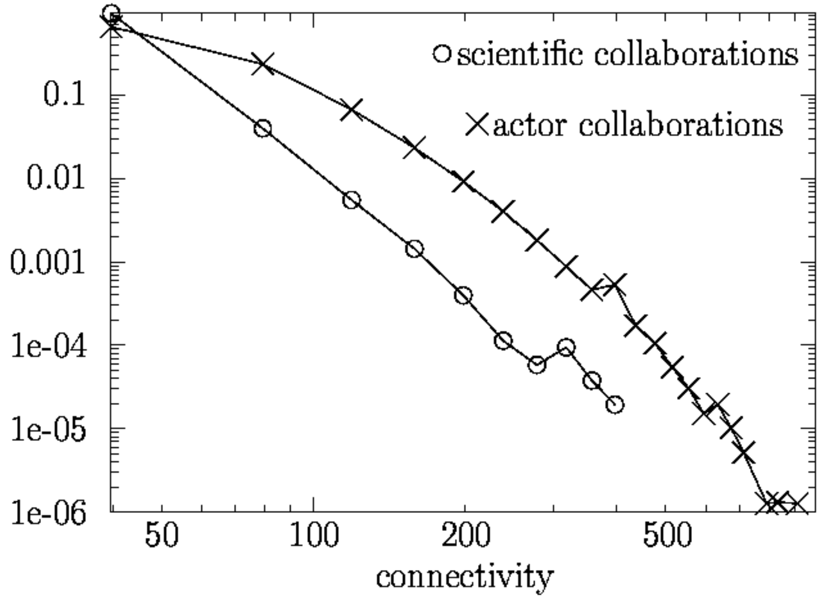
Department of Theoretical Physics, Umeå University, Sweden

2th August 2001

Contents

SOCIAL NETWORKS:
Power Law Connectivity Distribution

. . . more or less.



High Clustering

Social transitivity: If A knows B and A knows C then B is likely to know C.

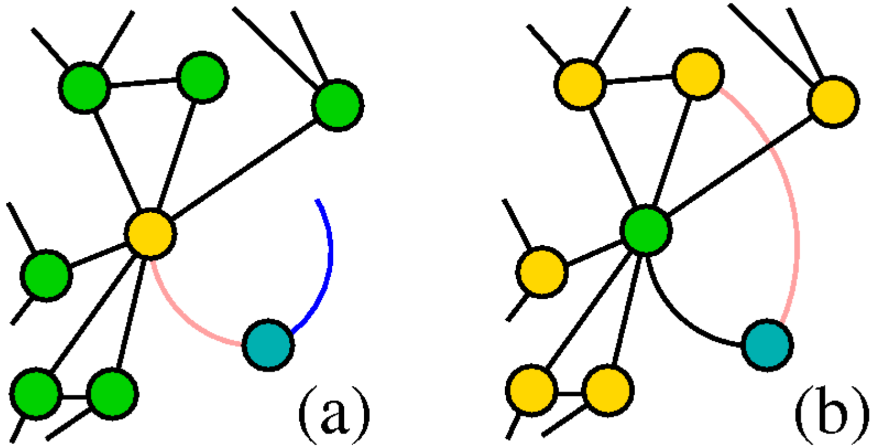
The ordinary scale-free network model is defined as follows:

Initial condition To start with the network consists of m_0 nodes and no links.

Growth One node v with m links is added every time step.

Preferential attachment A link is added to an old node in proportion to it's connectivity. Or, more precisely: the probability for a node w to be attached to is:

$$P_w = \frac{k_w}{\sum_{v \in V} k_v}$$



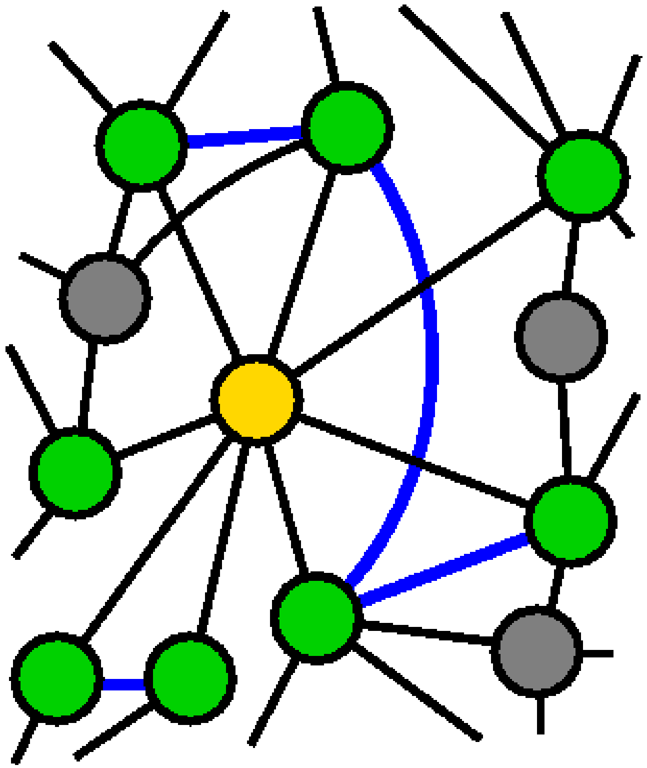
To get finite clustering for $N \to \infty$ we include:

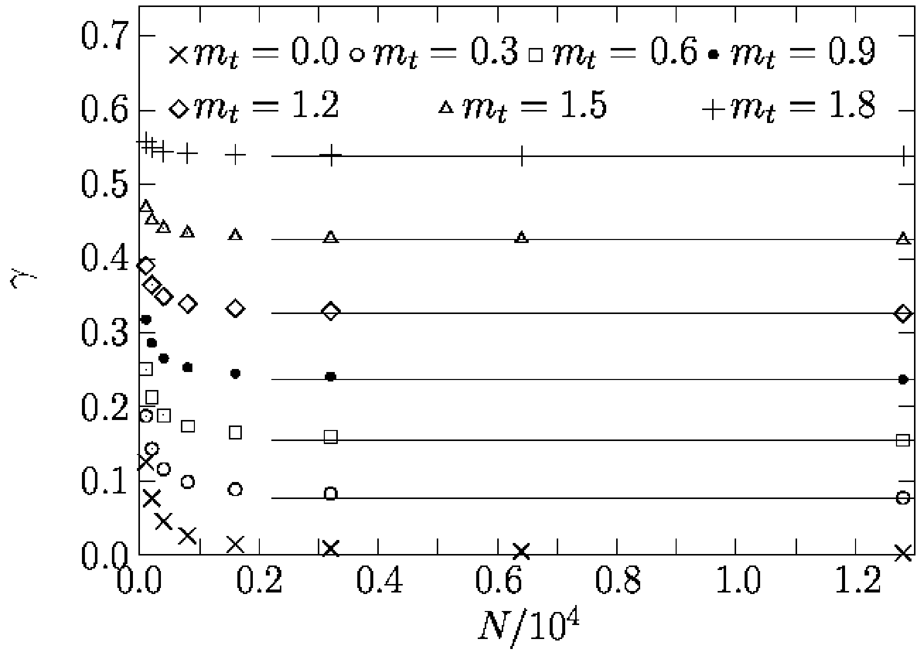
Triad Formation If a link between v and w was added in the previous step, then add a link from v to a randomly chosen neighbor of w (if not all neighbors of w are linked with v), otherwise do a PA step as above.

Clustering Coefficient

$$C_v = |E(\Gamma_v)| / \binom{k_v}{2}$$

where $|E(\cdot)|$ gives a sub-graph's total number of edges.





The clustering coefficient as a function of network size for different number of triad formation steps m_t. Straight lines show asymptotic γ-values.

Connectivity Distribution

TF steps don't change the connectivity distribution of standard scale-free network:

In a PA step the local connectivity is changed as:

$$\frac{\Delta k_v}{\Delta t} = 2\frac{k_v}{\Sigma_{w\in V} k_w}$$

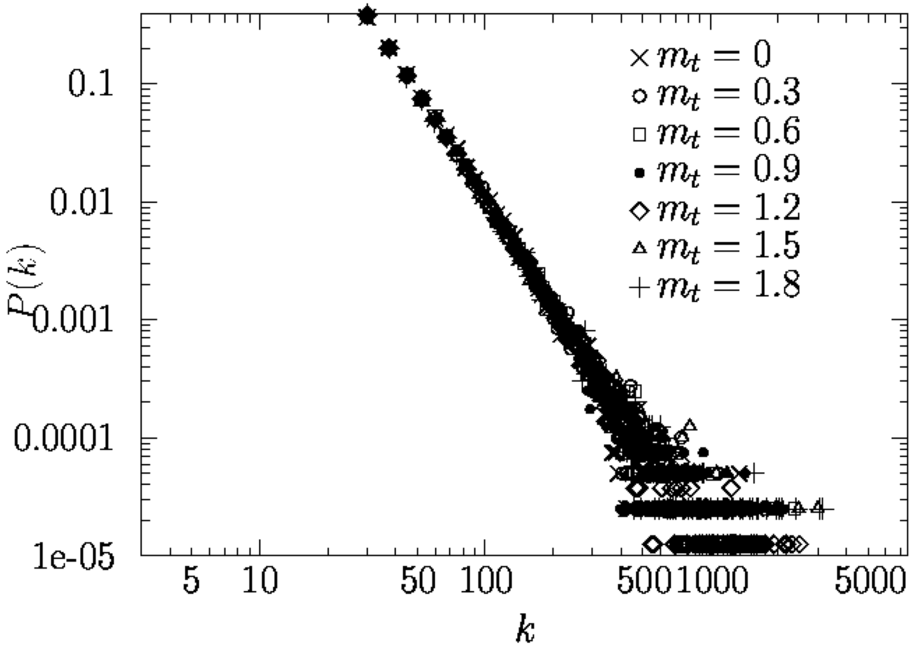
In a following TF step the connectivity changes as:

$$\frac{\Delta k_v}{\Delta t} = 2\frac{\Sigma_{w\in\Gamma(v)} k_w(1/k_w)}{\Sigma_{w\in V} k_w} = \frac{k_v}{\Sigma_{w\in V} k_w}$$

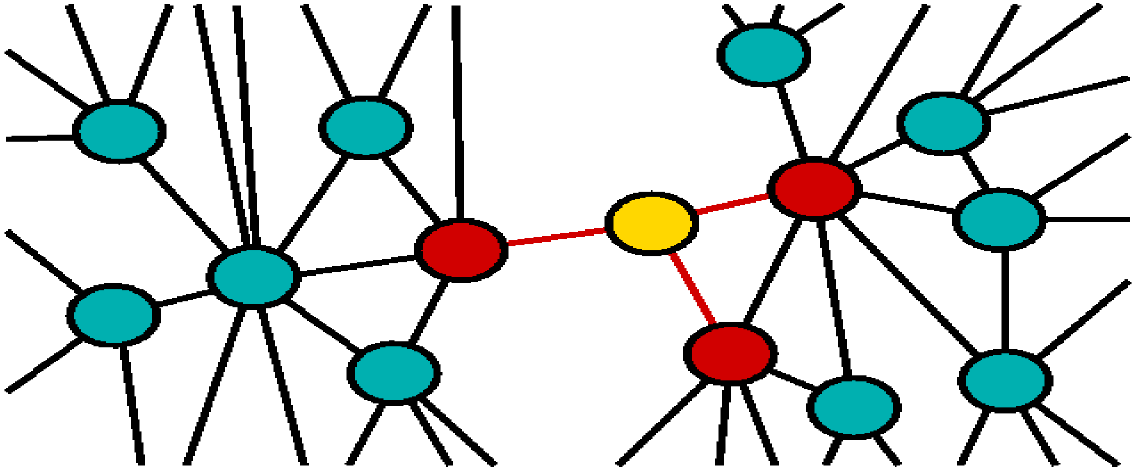
For a particular node the connectivity increases as a square root of the number of time steps:

$$\frac{\Delta k_v}{\Delta t} = \frac{2mk_v}{\Sigma_{w\in V} k_w} = \frac{2mk_v}{2mt} \Rightarrow k_v \propto t^{1/2}$$

... which is equal to the standard scale-free network model, and leads to a power-law connectivity distribution with slope –3.



Connectivity distribution for parameters $m = m_0 = 3$, $N = 10^6$ and different m_t.



- Nodes

$$C_B(v) = \sum_{w \neq w' \in V} \frac{\sigma_{ww'}(v)}{\sigma_{ww'}}$$

where $\sigma_{ww'}$ is the number of geodesics (shortest paths) between w and w', and $\sigma_{ww'}(v)$ is the number of geodesics between w and w' that passes v.

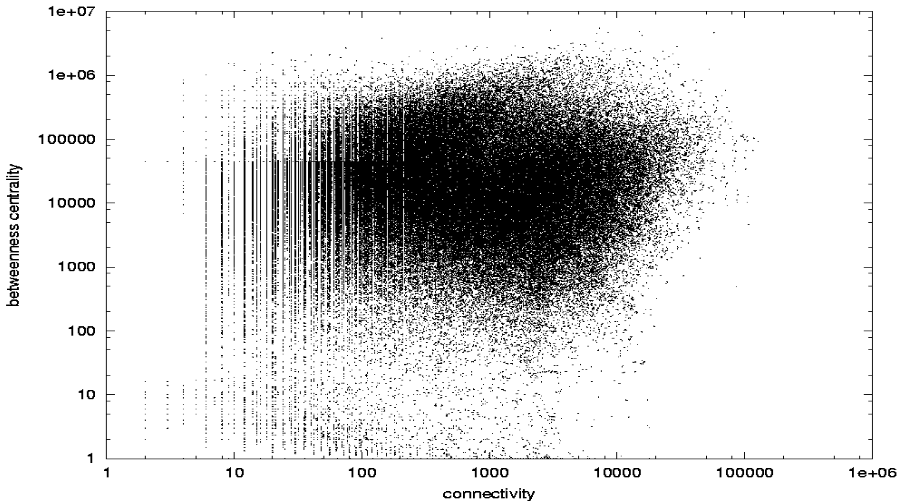
- Links

$$C_B(e) = \sum_{w \neq w' \in V} \frac{\sigma_{ww'}(e)}{\sigma_{ww'}}$$

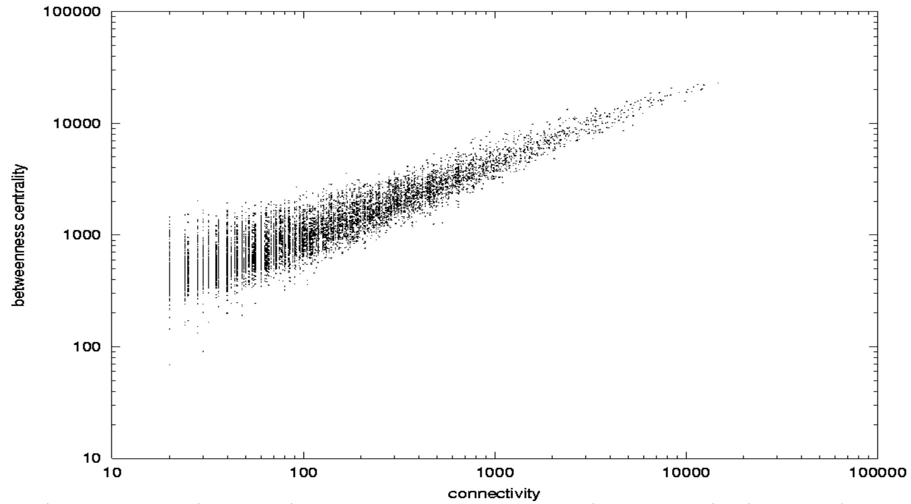
where $\sigma_{ww'}(e)$ is the number of geodesics between w and w' that includes e. (Note that $\sigma_{ww'}(v), \sigma_{ww'}(e) \in \{0, 1\}$.)

Motivation 2 (?): Betweenness Centrality *(continued)*

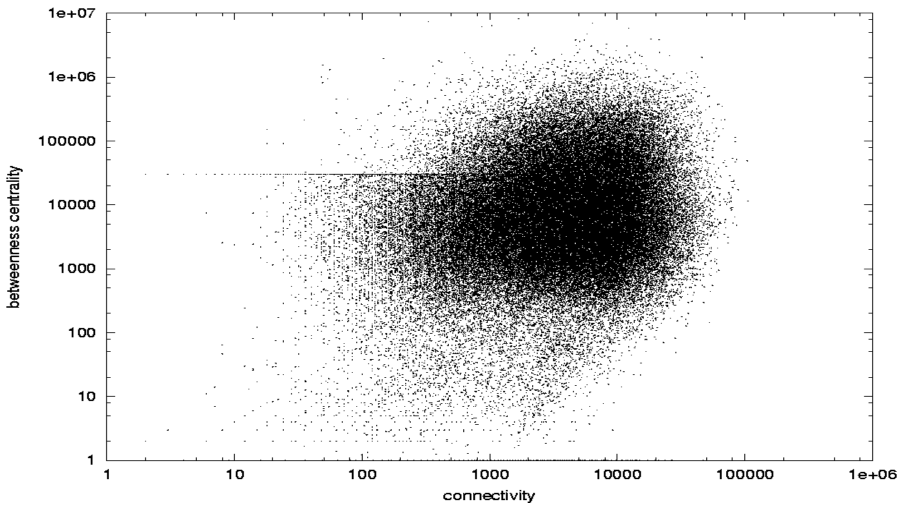
Scientific Collaborations Links



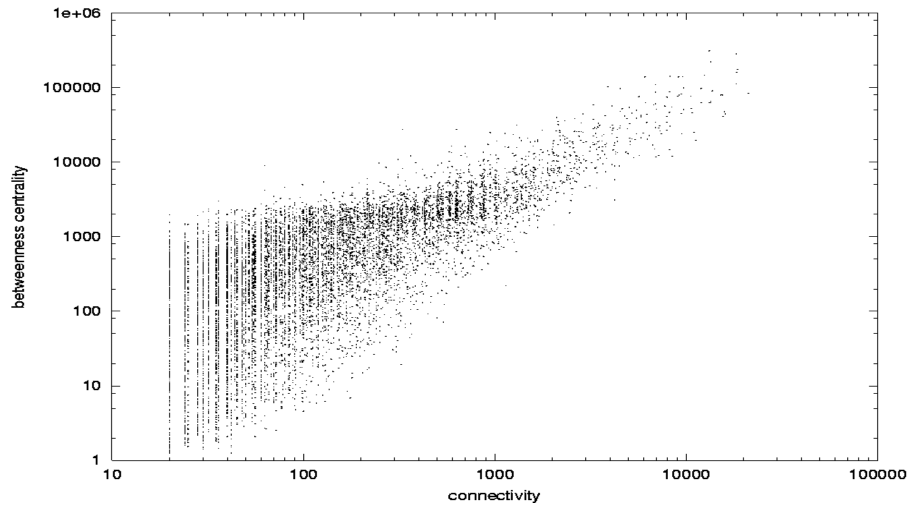
Scale Free Network Model Links



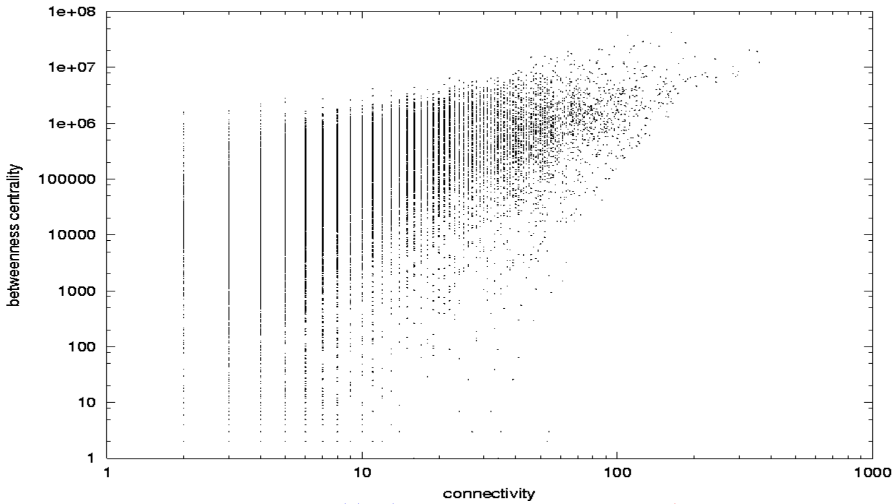
Actor Collaborations Links



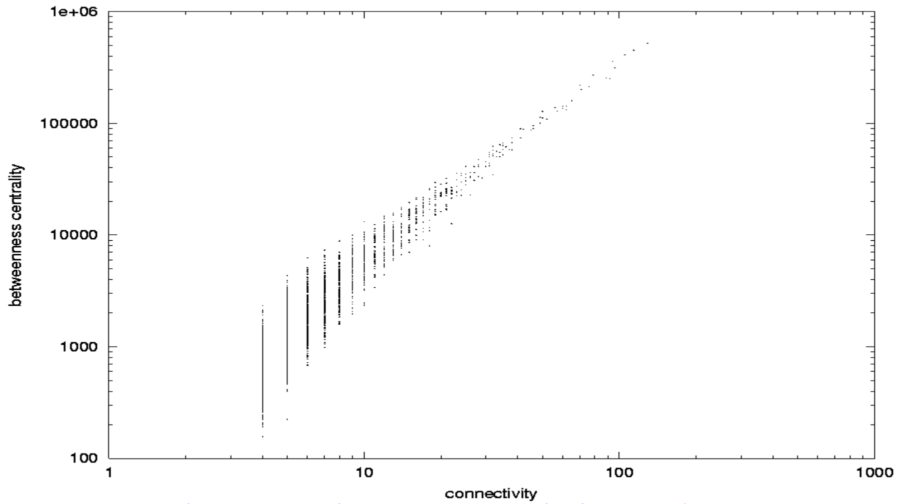
Clustered Scale Free Network Model Links



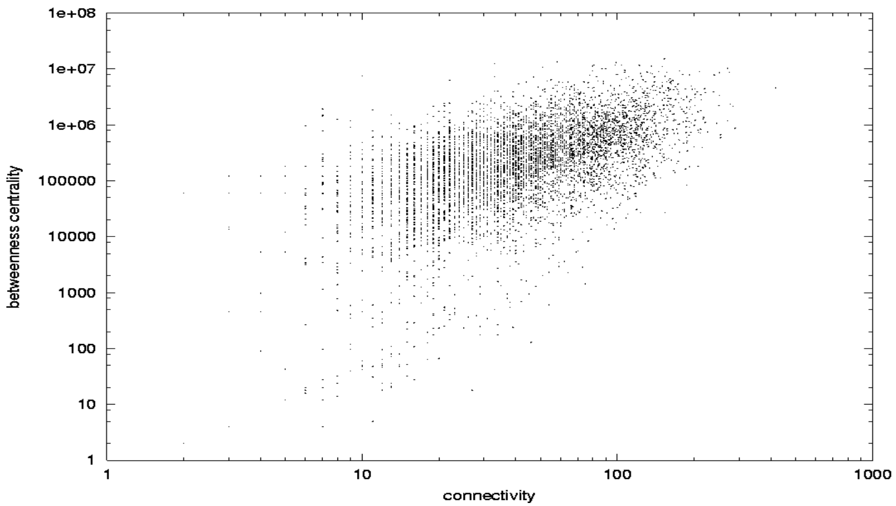
Scientific Collaborations Nodes



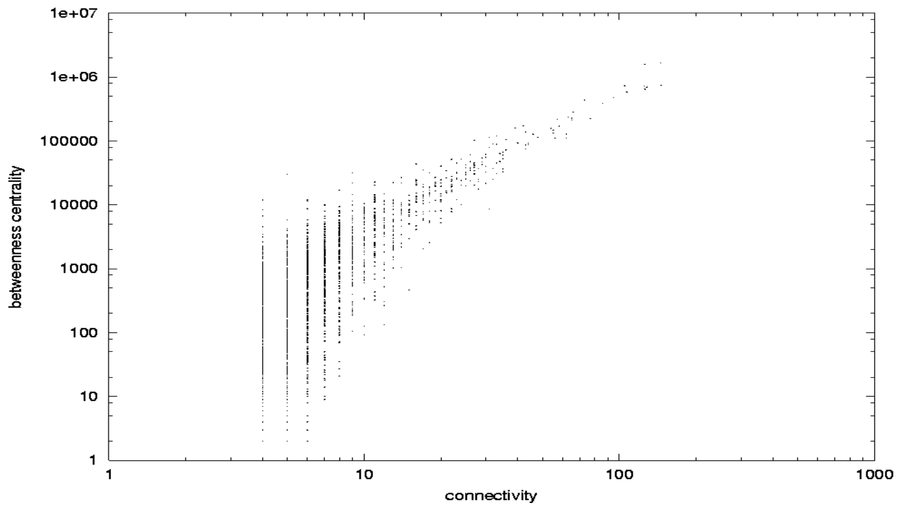
Ordinary SFN Model Nodes



Actor Collaborations Nodes



Clustered SFN Model Nodes



ATTACK VULNERABILITY

Measured quantity: The increase of the characteristic length if nodes (links) are removed in order of connectivity.

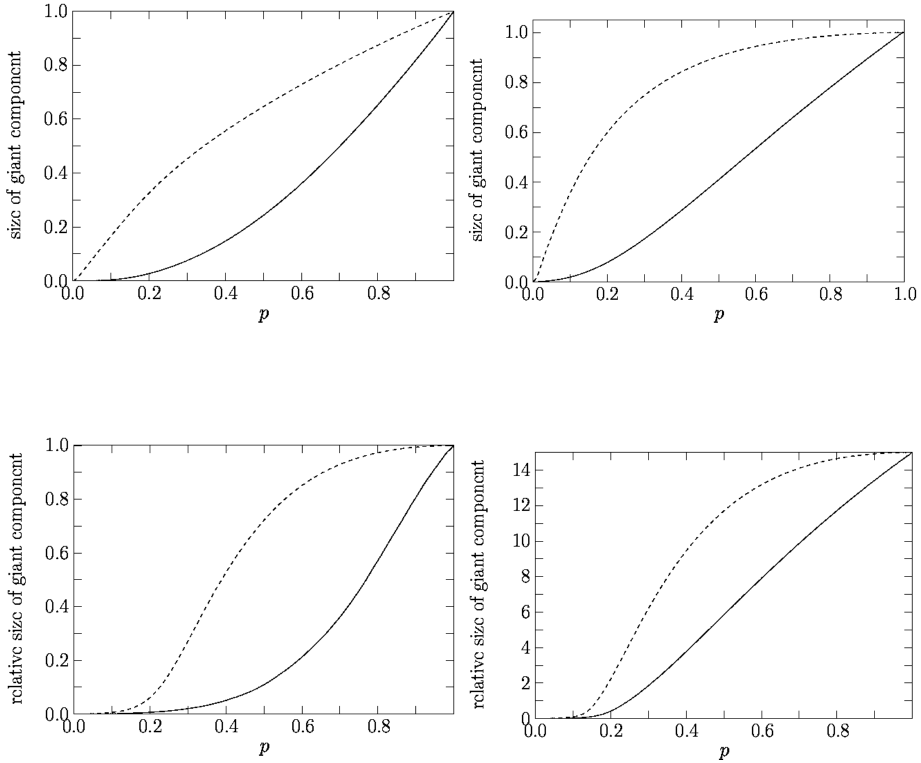
- Clustered SFN more sensitive to attack than standard SFN.
- Standard SFN more sensitive to attack than clustered SFN.

PERCOLATION

Measured quantity: The size of the largest connected active cluster (if a fraction p of the network is active).

- Clustered SFN performs better than standard SFN in link percolation.
- Standard SFN performs better than clustered SFN in link percolation.

- The clustered and standard SFN are both similar to real systems: $p_{c,\mathrm{node}} \approx 0$.



Site (solid) and bond (dashed) percolation for WWW, yeast metabolic system, clustered, and standard SFN respectively.