SCALE-FREE NETWORKS: Background, evolutionary models & simulation

John Pui-Fai SUM

PART I: BACKGROUND

Stanley Milgram (60s) [A.L. Barabasi, *Linked*, Plume, 2003]

Stanley Milgram, a Harvard researcher who in 1967 conducted a series of mailing experiments

- Initial senders are selected randomly from Kansas or Nebraska

- Forwarding mail to one of the two persons living/working in Boston

- Only the name of the persons and their careers are specified

- Each mail receiver will have to forward the mail to a friend

- Receiver knows the target person $\rightarrow$ send the mail directly to the target person

- Receiver does not know the target person $\rightarrow$ sends the mail to whoever appropriate

- Lot of mails have been lost (about 75 percent)

- Average forward steps is about six!

Duncan Watts (Mid 90s)

Associate Professor, Department of Sociology, Columbia University

1997 PhD Cornell University, Department of Theoretical and Applied Mechanics

Thesis title: The structure and dynamics of small-world systems

Contribution

> Discover behvaior of various real networks
> (File actors, Power grid, neural net)
> Mechanisms for the formation of such networks
> Dynamic behavior of such small world networks

1. Discoveries – Many real networks are not random

- <u>Social network</u>: Movie actor collaboration network

- <u>Technology network</u>: Power grid of western United State (4941 generators, transformers and substations)

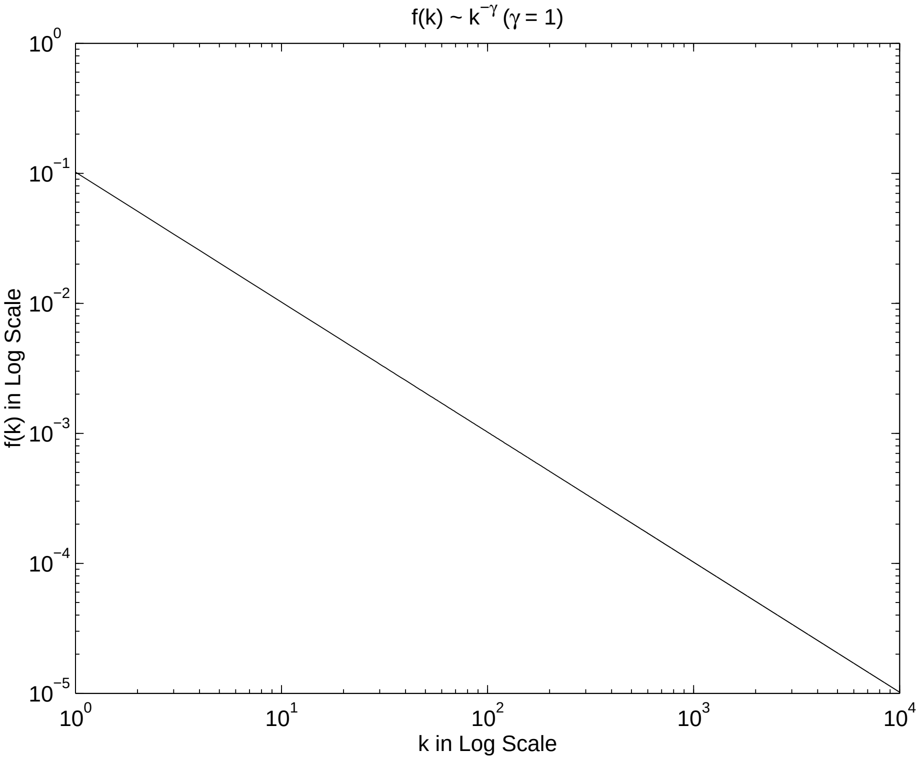
- <u>Biological network</u>: Neural network of a nematode worm *C. elegans* (282 neurons)

- Node degree distribution is not Poisson, i.e. not random network.

- Node degree distributions follow power law.

Power law network

$$f(k) = c_0 k^{-\gamma} \tag{1}$$
$$\log f(k) = \log c_0 - \gamma \log k$$



Average distance ℓ

$$\ell = \frac{2}{n(n-1)} \sum_{i>j} d_{ij} \tag{2}$$

where d_{ij} is the shortest path distance between i and j in a n nodes network

Clustering coefficient C

$$C = \frac{1}{n} \sum_{i} C_i \tag{3}$$

$$C_i = \frac{\text{\# of triangles connected to i}}{\text{\# of triples centered on i}}$$

	ℓ	ℓ_{rand}	C	C_{rand}
Movie actors	3.65	2.99	0.79	0.00027
Power grid	18.7	12.4	0.08	0.005
C. Elegans	2.65	2.25	0.28	0.05

Conclusion: *We are likely living in a small world. Real network is in between regular and random network.*

2. Network Formation – Watt-Strogatz Networks (1998)

- Starting with a regular network, a ring lattice.

- Each node is connected to four neighbor nodes, two at each side.

- Rewiring p percentage partial edges

Observation from simulation

- ℓ decreases as p increases

- C decreases as p increases

- Sharp drop of ℓ is earlier than sharp drop of $C \rightarrow$ there is a range of p that the network can have small ℓ but large C.

- C.F. random network, both ℓ and C are small.

A.L. Barabasi & R. Albert, Marc Newman, Dorogovtsev & Mendes etc.(Late 90s)

Contribution

> Discover Power-law like networks
> Network evolution mechanisms
> Properties analysis

... *How do various microscopic processes influence the network topology ? ... Are there quantities beyond degree distribution that could help in classifying networks? ... These results signal the emergence of a self-consistent theory of evolging networks, offering unprecedented insights into network evolution and topology.* (P.76 of Albert & Barabasi, Statistical mechanics of complex networks, *Reviews of Modern Physics* Vol. 74, 47-97, 2002.)

1. Current discoveries

- World Wide Web (Adamic 1999)
- Internet router network (Faloutsos *et al.* 1999)
- Telephone call network (Aiello *et al.* 2000)
- Email message (Ebel *et al.* 2002)
- Sexual contacts (Liljeros *et al.* 2001)
- Research papers co-authorship (Newman 2001; Barabasi 2001)
- Words co-occurrence (Ferreri Cancho & Sole 2001)

- P2P network (Jovanovic 2001; Ripeanu *et al.* 2002; Saroiu *et al.* 2002)

- Software classes (Valverde *et al.* 2002)

Question: How those networks are being formed? Any general principle behind?

2. Formation of Power law networks

- *Preferential attachment* models – Scale-free networks (Albert & Barabasi 1999)
  - Initially m nodes ($s = 0, \ldots, m-1$) are fully connected
  - Node is added one at a time
  - m *new edges* are connected m different existing nodes selected randomly
  - For $t \gg 1$ and $k_s > m$

$$h(k+1, s, t) = m \frac{k_s}{\sum_j k_j} = \frac{k_s}{2t} \qquad (4)$$

since

$$\sum_j k_j = 2 \left\{ m^2 + \sum_{j=1}^{t-m+1} m \right\} = 2mt$$

- Exponent γ equals 3, i.e. $P(k) \propto k^{-3}$, for $k, t \gg m$

- Attachment with node decay (Dorogovtsev & Mendes 2000)
  - Initially m nodes ($s = 0, \ldots, m-1$) are fully connected
  - Node is added one at a time
  - m *new edges* are connected to randomly m different existing nodes
  - For large t and $k, s > m$

$$h(k, s, t) = \frac{k_s(t-s)^{-\lambda}}{\sum_j k_j(t-j)^{-\lambda}} \qquad (5)$$

    $0 < \lambda < 1$ is node decay rate

- Random edge attachment (Dorogovtsev & Mendes 2003)

- Initially m nodes ($s = 0, \ldots, m-1$) are fully connected

- Node is added one at a time

- *2 new edges* are connected to the two ends of a randomly selected edge

- Exponent γ equals 3, i.e. $P(k) \propto k^{-3}$

- Others

  - Degree correlation preference (Chung & Lu 2002)

  - Edge decay

  - Nonlinear preferential attachment

$$h(k, s, t) = \frac{k_s^{\beta}}{\sum_{j=0}^{t} k_j^{\beta}}$$

– Probabilistic edges rewiring

3. Properties in power law networks

- Node degree correlation, $f_{ij} = k_i k_j$, follows power law

- Contention L_i follows power law

$$L_i = \# \text{ of shortest paths via node } i$$

- Local cluster coefficient C_i follows power law

- Network resiliency (Barabasi *et al* 200x; Callaway *et al.* 2000; Cohen et al. 2001)

  – Random node removal: Network is connected even $0.6n$ nodes have been removed $\rightarrow$ fault tolerance

  - High degree first removal: Network will be disconnected if $0.2n$ nodes have been removed $\rightarrow$ intentional attack in-tolerable

- Network vulnerability (Newman *et al.* 2000; and collaborators)

- Size of the giant component (the largest connected subgraph)

- Bose-Einstein condensation: A significant fraction of nodes will be connected to just a few strong nodes

A Few Remarks

1. Definition on small world networks has not yet been concluded.

- Watt-Strogatz model is a small world network, something between regular and random.

- Barabasi-Albert model is scale-free network, node degree follows power law and this propoerty does not change with the size of the network

- Growing network refers a network that the size can grow(*). Idea is similar to the one in neural network but specifically for the network with node degrees follow
  - $k^{-\gamma}$

– $k^{-\gamma}\,\text{exp}(-\alpha k)$

– $\text{exp}(-\alpha k)$

- Evolution networks refers to a network that can grow(*) or decay(*) (i.e. evolve). The grow/decay can act on nodes or edges as well.

- Complex network is a general name for all these networks.

2. Power law is not the only distribution found in real networks other than Poisson

- $\frac{k^{-(1+\gamma)}}{\log k}$

- $\alpha k^{-\gamma} \log(\beta k)$

3. Random graph models

- Erdo-Renyi studied the properties of random graph in 1960. The existence of an edge between two nodes is depended on a fixed probability p between zero and one.

- It should be noted that there are two types of random graph usually denoted by $G_{n,M}$ and $G_{n,p}$.

- $G_{n,M}$ is the orginial Erdo-Renyi graph that consists of n nodes and M edges. M is pre-defined. Edges are assigned randomly to N locations, where

$$N = \begin{pmatrix} n \\ 2 \end{pmatrix}.$$

- $G_{n,p}$ refers to the graph in which a edge between two nodes is generated randomly with probability p.

- For n is large and $p = M/N$, their properties have been proved to be the same.

- The node degree distribution of both random graphs follow Poisson distribution.

PART II: EVOLUTIONARY MODELS

Notations & Defintions

Notation	Meaning
t	Time
m	Number of edges added to a new node
s	(Node index) The time node s is added
k	Node degree
$p(k,s,t)$	Probability that node s will have deg. k at $t \geq s$
$p(m,s,s) = 1$	Boundary condition
$P(k,t)$	Proportional of nodes that have deg. k at time t
$\bar{P}(k)$	Node degree distribution
$\bar{k}(s,t)$	Average node degree
$\bar{k}(t,t) = m$	Initial condition
$G(z)$	Moment generator function

$$
\begin{aligned}
s &= 0, 1, 2, 3, \ldots, t \\
p(m, s, m) &= 1 \ \forall s = 0, 1, 2, \ldots, m \qquad (6) \\
p(m, t, t) &= 1 \ \forall t \geq m \qquad (7) \\
P(k, t) &= (t + 1)^{-1} \sum_{s=0}^{t} p(k, s, t) \qquad (8)
\end{aligned}
$$

Total number of nodes are $t + 1$ at time t

$$
\begin{aligned}
\bar{P}(k) &= \lim_{t \to \infty} (t + 1)^{-1} \sum_{s=0}^{t} p(k, s, t) \qquad (9) \\
\bar{k}(s, t) &= \begin{cases} \sum_{k=m}^{t} k p(k, s, t) & s < t \\ m & s = t \end{cases} \qquad (10)
\end{aligned}
$$

Between the time t and s, total number of new nodes being added is $(t - s)$. The total number of new edges added on an existing node will not be larger than $(t - s)$. Therefore,

$$p(k, s, t) = 0 \quad \forall\, k \geq (t - s + m)$$

Precisely, Equation (8) and (10) should be written as follows :

$$P(k,t) = (t-k+m+1)^{-1} \sum_{s=0}^{t-k+m} p(k,s,t) \quad (11)$$

$$\bar{k}(s,t) = \begin{cases} \sum_{k=m}^{t-s+m} kp(k,s,t) & s \geq m \\ \sum_{k=m}^{t} kp(k,s,t) & 0 \leq s \leq m \end{cases} \quad (12)$$

Example $m = 2$, $t = 5$

(s,k)	0	1	2	3	4	5
0	0	0	+	+	+	+
1	0	0	+	+	+	+
2	0	0	+	+	+	+
3	0	0	+	+	+	0
4	0	0	+	+	0	0
5	0	0	1	0	0	0
6	1	0	0	0	0	0

$p(k,s,t)$ in the + locations have positive values.

General Model

$$\begin{aligned} p(k,s,t+1) &= h(k-1,s,t)p(k-1,s,t) \\ &+ (1-h(k,s,t))p(k,s,t) \quad (13) \end{aligned}$$

where $h(k,s,t)$ is the probability that an edge will be added to a node of degree k.

Analysis techniques

- Moment generator function

$$\begin{aligned} G(z) &= \sum_{k=1}^{\infty} p_k z^k \\ G(1) &= 1 \\ G'(1) &= \sum_{k=1}^{\infty} k p_k \\ G^{(k)}(0) &= k! p_k \end{aligned}$$

- Master equation, i.e. Equation (13)

- Continuous differential equation

$$\frac{\partial p(k,s,t)}{\partial t} \approx p(k,s,t+1) - p(k,s,t)$$
$$\frac{\partial p(k,s,t)}{\partial k} \approx p(k,s,t) - p(k-1,s,t)$$

Since $\bar{k}(s,t) = \int_m^{t-s+m} kp(k,s,t)dk$, for all $t \gg s$

$$\frac{\partial \bar{k}(s,t)}{\partial t} = (t-s+m)p(t-s+m,s,t) + \int_m^{t-s+m} k\frac{\partial p(k,s,t)}{\partial t}dk$$

<u>Useful equalities</u>

$$\frac{\partial p(k,s,t)}{\partial t} + \frac{\partial h(k,s,t)p(k,s,t)}{\partial k} = 0 \qquad (14)$$

$$\frac{\partial p}{\partial t} + h\frac{\partial p}{\partial k} + p\frac{\partial h}{\partial k} = 0$$

Assuming that

$$\bar{P}(k) \propto k^{-\gamma} \text{ and } \bar{k}(s) \propto s^{-\beta},$$

we have

$$s(\bar{k}, t) \propto \bar{k}^{-1/\beta}$$

$$\bar{P}(k) \propto \frac{\partial s}{\partial \bar{k}(s,t)} \propto \bar{k}^{-1-\frac{1}{\beta}}$$

Hence,

$$\gamma = 1 + \frac{1}{\beta} \tag{15}$$

$$\beta(\gamma - 1) = 1 \tag{16}$$

Barabasi-Albert Model

Master equation approach

$$\begin{aligned} p(k,s,t+1) &= \frac{k-1}{2t}p(k-1,s,t) \\ &\quad + (1-\frac{k}{2t})p(k,s,t) \\ tp(k,s,t+1) &= (k-1)p(k-1,s,t) \\ &\quad + (t-k)p(k,s,t) \end{aligned}$$

for $k \geq m$. Adding both side $p(k,t+1,t+1)$ and then sum up for s from 0 to t,

$$\begin{aligned} & (t+2)\bar{P}(k,t+1) \\ = & (k-1)\sum_{\tau=1}^{t}\frac{\tau+1}{m+1+2\tau}\bar{P}(k-1,t) \\ & - k\sum_{\tau=1}^{t}\frac{\tau+1}{m+1+2\tau}\bar{P}(k,t) \end{aligned} \qquad (17)$$

When $t \to \infty$

$$\sum_{\tau=1}^{t}\frac{\tau+1}{m+1+2\tau}\bar{P}(k,t) \approx \frac{t}{2}\bar{P}(k)$$

Equation (17) becomes

$$\bar{P}(k) \approx \frac{k-1}{k+2}\bar{P}(k-1) \tag{18}$$

For $k = m$,

$$\begin{aligned} p(m,s,t+1) &= (1-\frac{m}{t})p(m,s,t) \\ tp(m,s,t+1) &= (t-m)p(m,s,t) \\ (t+1)p(m,s,t+1) &= (t-m)p(m,s,t)+1 \end{aligned}$$

since $p(m,t,t) = 1$ for all t. Similarly,

$$\bar{P}(m) = \frac{2}{m+2}$$

Thus for $k \gg m$

$$\bar{P}(k) = \frac{2(m+1)m}{(k+2)(k+1)k} \approx 2m(m+1)k^{-3}$$

<u>C.D.E. approach</u>

$$\begin{aligned} \frac{\partial p(k,s,t)}{\partial t} &= -\frac{1}{2t}\frac{\partial p(k,s,t)}{\partial k} && (19) \\ \frac{\partial \bar{k}(s,t)}{\partial t} &= \frac{\bar{k}(s,t)}{2t} && (20) \end{aligned}$$

Solving Equation (20) with the boundary condition $\bar{k}(s,s) = m$,

$$\bar{k}(s,t) = m\sqrt{\frac{t}{s}} \text{ and } \beta = \frac{1}{2} \tag{21}$$

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

Node Decay Model

Master equation approach

$$\begin{aligned} p(k,s,t+1) &= \frac{(k-1)(t-s)^{-\lambda}}{E(t,\lambda)} p(k-1,s,t) \\ &+ (1 - \frac{k(t-s)^{-\lambda}}{E(t,\lambda)}) p(k,s,t) \end{aligned}$$

for $k \geq m$ and

$$\begin{aligned} E(t,\lambda) &= \sum_{s=0}^{t} k_s (t-s)^{-\lambda} \\ &\approx \int_0^t \bar{k}(u,t,\lambda)(t-u)^{-\lambda} du \end{aligned}$$

C.D.E. approach

$$\frac{\partial \bar{k}(s,t)}{\partial t} = \frac{\bar{k}(s,t)(t-s)^{-\lambda}}{E(t,\lambda)} \qquad (22)$$

with boundary condition $\bar{k}(t,t) = m$ and

$$E(t,\lambda) = \int_1^t \bar{k}(s,t)(t-s)^{-\lambda} ds$$

Lecture 3: Applications

Possible applications

Network modeling and analysis (Topology, formation mechanism, session length etc.)

Internet
World Wide Web
P2P network
Mobile phone network
Mobile P2P

Performance evaluation by simulation

Computational limitation
File distibution
Search queries
Routing algorithms
Search algorithms
Load balancing

Application example 1: IP search in Internet

Modified Ant Routing (Sum *et al.* 1999, Sum *et al.* 2001)

- Message from Node A to Node B
- Ants dispatched to the network from Node A
- Visit Node B
  - Not the destination → Select a random neighbor and go
  - B is the destination → Backtrack the path to Node A

Assumptions

- Unlimited resource for search

- Number of requests at each node are all the same

- Dying rate to mimic TTL

  – (A1) Node degree dependent: $(1+\Omega_i)^{-1}$

  – (A2) Constant rate: $0 < s < 1$

Model (Average case)

$$\vec{p}(t+1) = A\vec{p}(t) + km\vec{e}$$

$$\begin{aligned} \vec{p} &= (p_1, p_2, \ldots, p_n)^T \\ \vec{e} &= (1, 1, \ldots, 1)^T \end{aligned}$$

- $A \in R^{n \times n}$ is the transfer matrix

- Elements of A depend on dying models and network topology

<u>Results and extension</u>

- (A1): $p_i \leq km(1+\Omega_i)$

- (A2): $p_i \leq \frac{km}{s}$

(Sum 2003) *Searching in a Power law network with fixed topology (Internet), bounds of p_i follows Power-law.*

$$
\begin{aligned}
p_i \leq \widehat{p}_i &= km(1+\Omega_i) \\
P(\widehat{p}) &\propto \widehat{p}^{-a}
\end{aligned}
$$

Application example 2: P2P modeling

Parameters

- Search methods: Random search and/or hash table facilitated search

- File replication: With or without file replication mechanism

- Network structures: Unstructured P2P network and/or with structured overlay network

- Immunization: With or without immunization

- Node natures: Always on-line or randomly on-line

- Node classes : Ultrapeers, ultrapeer capable, shielded leaf or leaf
- Computational power reserved for peer operation: Limited or unlimited.
- Node failure models: Due to overload or attack

Network properties (Structural, Efficiency, Vulnerability)

Structural

1. Node degree distribution
2. Cluster coefficient: The proportion that the neighbor nodes of a node will also be neighbors of each other

Efficiency

1. Average shortest path: Measure the average distance between pairs of nodes

2. Load distribution: The bounds on the loading of a node

3. Latency : The time taken a search message to travel from one node to another.

4. Betweenness : It is calculated as the fraction of shortest paths between nodes pairs that pass through the node of interest.

Vulnerability

1. Betweenness: A measure for loading of a node due to search.

2. Connectivity: Measure the network fragmentation

3. Size of giant component: Measure the size a search can reach.

4. Transmissibility: The spreading rate of a virus over the network.

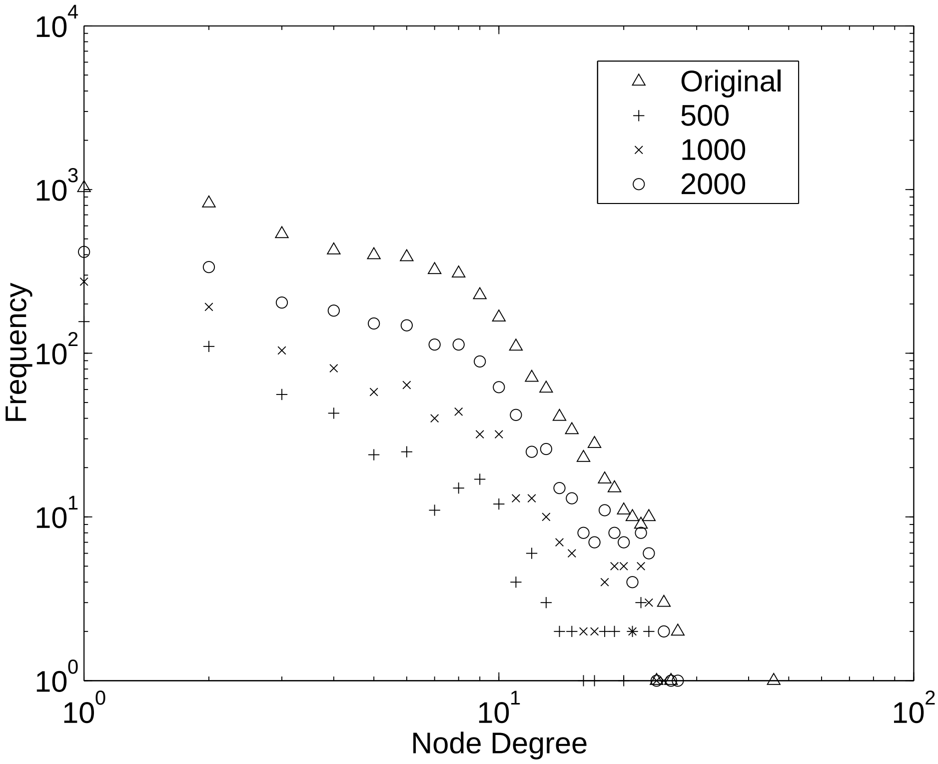
Case Studies: Gnutella

Measurements (Sum & Wong 2003)

- Gnutella v0.6
- Ultrapeer capable (long expected uptimes)
- Compared with Ripeanu *et al.* (2002)
- Distributions on node degree and session up time

Node Degree – Power Law Distribution

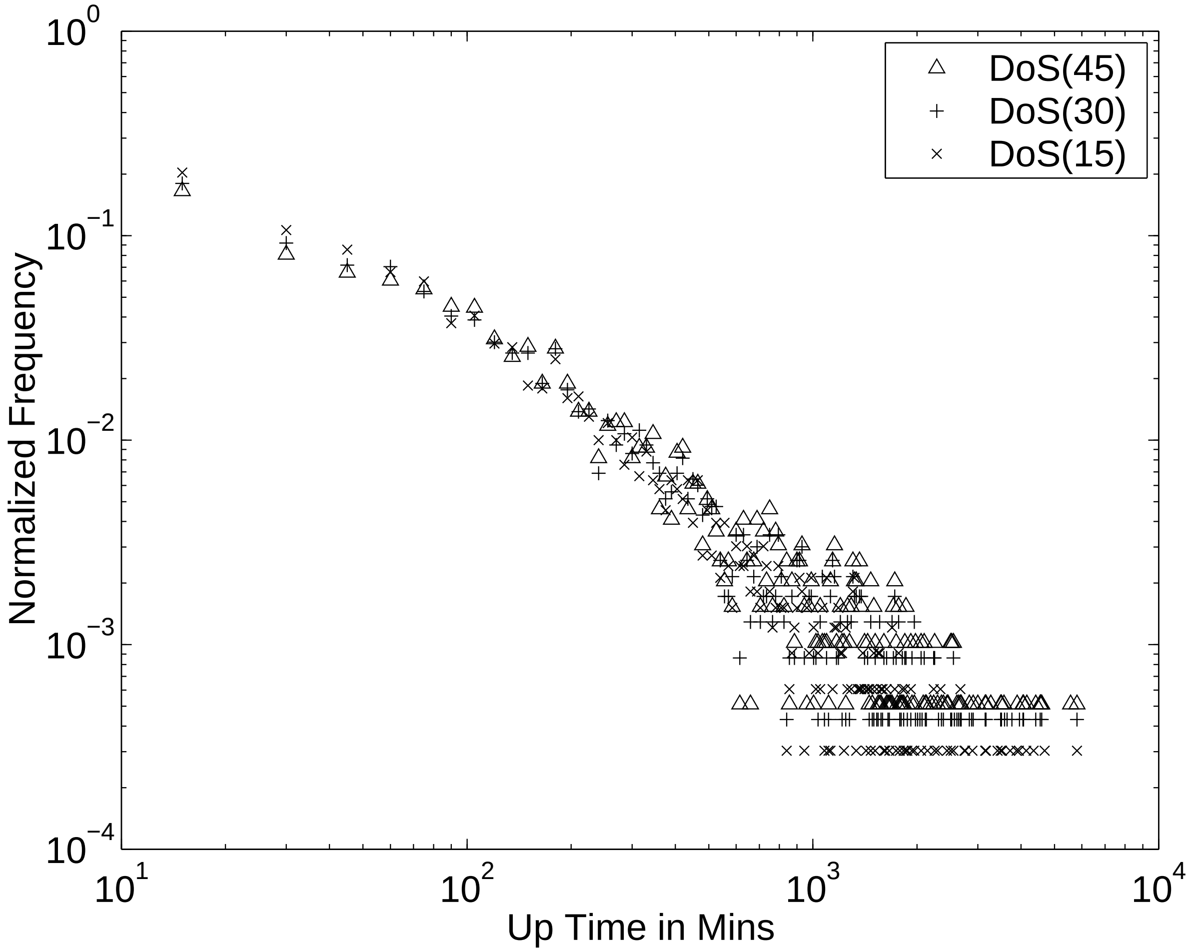
- $k \in [5, 20]$: $P(k) \propto k^{-4}$



August 2003

Session Time – Gamma Distribution

$$P(t) \sim \left(\frac{t}{N}\right)^{-0.82} \exp\left(-\frac{x}{36N}\right). \qquad (23)$$



Dorogovtsev-Mendes Decay Network Model

$$
\begin{aligned}
P(k) &\propto k^{-\gamma} \\
\gamma &\approx 3 + 4(1 - \log 2)\lambda
\end{aligned}
$$

In accordance with the measurement

$$
\begin{aligned}
\lambda_0 &\approx 0.83 \\
\gamma_0 &\approx 4
\end{aligned}
$$

Putting λ_0 into the approximated relation

$$\gamma(\lambda_0) \approx 3 + 4(1 - \log 2)\lambda_0 \approx \gamma_0.$$

Conclusion: *Dorogovtsev-Mendes decay network model might be a possible explanation for the formation of Gnutella P2P.*

Simulation study on using Dorogovtsev-Mendes decay network model

<u>Parameters</u>

$$\begin{aligned} t &= [1, 3000] \\ N &= 8 \\ M &= 5 \\ s &= 2 \\ m &= 3 \end{aligned}$$

t: Number of time steps

N: Scaling factor in $P(t)$

$$\left(\frac{t}{N}\right)^{-0.82} \exp\left(-\frac{t}{36N}\right)$$

M: Number of new nodes being added in each time step

s: Number of seed nodes

m: Number of new connections each new node will add.

<u>Algorithm</u>

- Initial: s nodes fully connected are generated

- Repeat:
  - M new nodes are added
  - Each nodes are connected to m exiting old nodes based on the principle of preferential attachment
  - Nodes are removed randomly following

$$I(\tau) = \begin{cases} \text{Stay alive} & \text{If } \delta_\tau > 0 \\ \text{Remove} & \text{If } \delta_\tau \leq 0 \end{cases}$$

Here

$$\delta_\tau = \frac{1 - P(t \le \tau + 1)}{1 - P(t \le \tau)} - r$$

and r is a random number distributed uniformly in $[0, 1]$. Note that the expectation of δ_τ is a condition probability that

$$\delta_\tau \sim P(\text{Alive at } \tau + 1 | \text{Alives at } \tau).$$

Measurement

- Node degree distribution
- Number of on-line nodes
- Expected degree in terms of time

MatLab Code

```
function [Non,ND,A,lt] = gnu(n,m,s,M,N,rr)

%-------------------------------------
% function [Non,ND,A,lt] = gnu(n,m,s,M,N,rr)
%
% Non : [rr,1] array for on-line nodes number
% ND : [rr,n+1] array node degree distribution
% A : Connection matrix
% N : scaling factors, normally 50*M
% M : Number of nodes being added in each round
% n: number of nodes
% m: number of new connections
% s : number of seed nodes, the seed nodes form a
%     fully connected during initialization
% rr : Number of repeated cycles
%
% Simulator of Gnutella P2P Network
% =================================
%   This simulator will automatically generate a
% scale-free network that mimics the behavior of
% Gnutella P2P network. Node will be generated
% and deleted in accordance with the information
% measured from the actual platform. The network
% generation model is the D-M decay model.
%
% Algorithm
% ---------
%
% Initalization: Generate s (s>m) seed nodes that are
%         fully connected
% Step 1a: Generate 200 online nodes index
% Step 1b: Connect each node to random select m
%          different nodes from the online list
% Step 2: Offline the online nodes
```

```
%

ND = zeros(n+1); % Node degree distribution
Non = zeros(ceil(rr/N),1);  % Number of on-line nodes
NonK = 1;
A = zeros(n,n);
Bdown = zeros(rr,1);
v = rand(n,2);
I = zeros(n,1);
lt = zeros(n,1);       % Life time vector
nindex = [1:1:n];

% ---------------------------------
% Simulate the actual life time
%
kc = 36;
x = [1:1:100*N];
ff=power(x/N,-0.82).*exp(-x/N/kc);
ff = ff/sum(ff);
pp(1) = ff(1);
for k=2:100*N,
    pp(k) = pp(k-1) + ff(k);
end


% ----------------------------------
% Seed nodes on-line
% Initiate on-line node list 'Son'
%

Seed = mod(round(rand(s,1)*100000),n) +1 ;
A(Seed,Seed) = ones(s,s);
Son = Seed;

for repeat=1:rr;
```

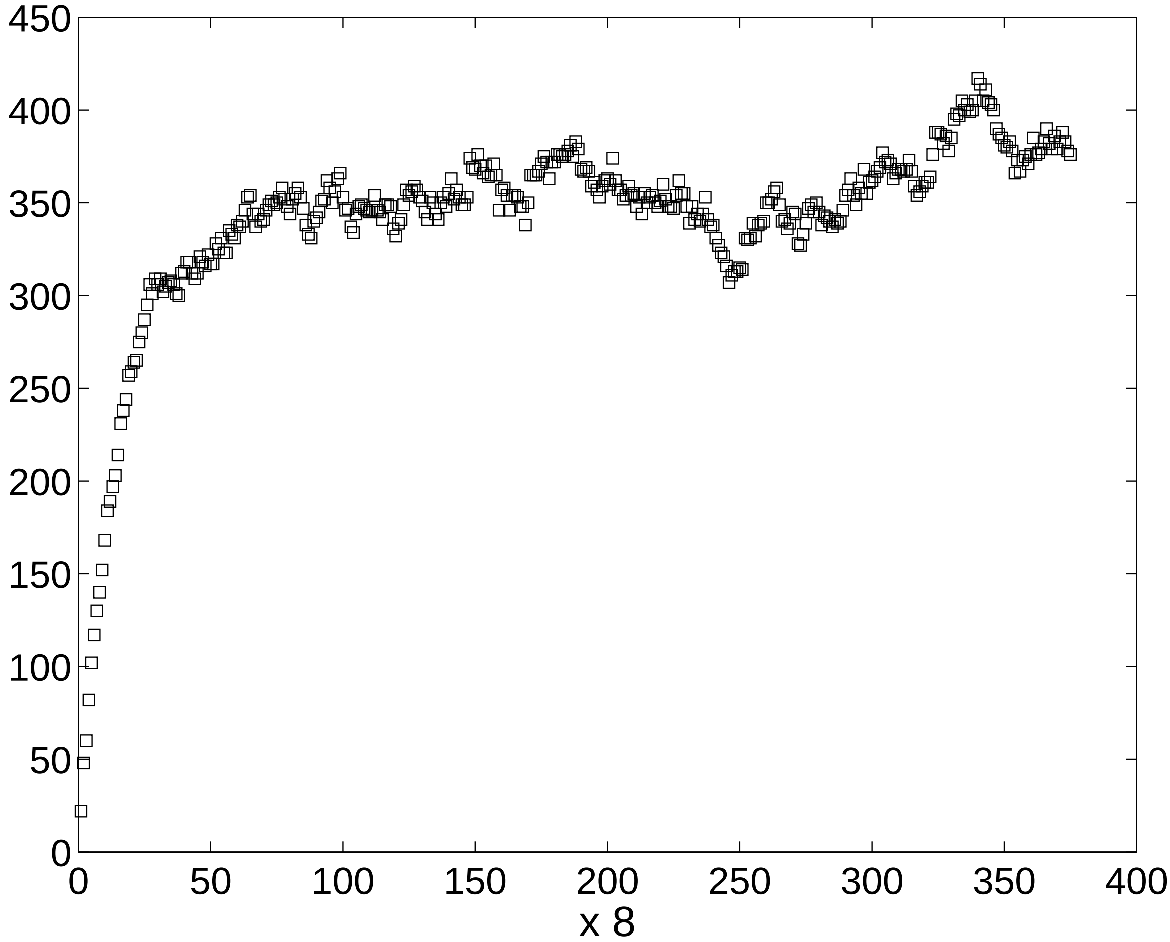
```
% ---------------------------------------
% Random off line according to power law
% Update online nodes list Son
% Update connection matrix A
% Update life time index 'lt'
%

if repeat > 1,
    Nol = length(Son);
    tt = rand(1,Nol);
    Palive = (1-pp(lt(Son)+1))./(1-pp(lt(Son)));
    SignOn = sign(Palive-tt);
    OnFlag = nindex(Son).*SignOn;
    Son = intersect(Son,OnFlag);
    Soff = setdiff(nindex,Son);
    ndel = length(Soff);
    A(:,Soff) = zeros(n,ndel);
    A(Soff,:) = zeros(ndel,n);
    lt(Soff) = 0;
end

Sontmp = Son;

% ------------------------------------------
% Display the result immediately
%
%

if (mod(repeat, N) == 0),
[MaxF, I] = sort(n-sum(A));
Non(NonK) = sum(diag(A));
NonK = NonK + 1;

figure(1);
gplot(A,v);
hold off;
```

```
plot(v(:,1),v(:,2),'g+');
plot(v(I(1:10),1),v(I(1:10),2),'rs');
drawnow

figure(2);
hold off
ND = hist(max(0,sum(A)-1),[0:1:n]);
loglog([1:1:n],ND(2:n+1),'^');
drawnow

figure(3);
hold off
plot(Non,'s');
drawnow;
end

% -------------------------------------------
% Increment the life time of online
% nodes by 1b
%

lt(Son) = lt(Son) + 1;

% -------------------------------------------
% Random on-line M nodes from 'Soff'
% If nodes number on Soff < M
% Select all nodes in Soff to online
% Update corresponding life time
%

Soff = setdiff(nindex,Son);
LSoff = length(Soff);
if LSoff > M,
    olindex = Soff(1:M);
elseif LSoff > 0
    olindex = Soff;
```

```
else
    olindex = 1;
end

lt(olindex) = 1;

% -------------------------------------------
% Make random connection on the new nodes
% Calcuate the total node degrees 'sumfk'
% Making 'm' new connects from each new node
%  to Son
%

Mon = min(M, length(olindex));
for newnode = 1:Mon,

    fk = sum(A(:,Son)) - 1;
    sumfk = sum(fk);

    j=olindex(newnode);
    A(j,j) = 1;
    for jj = 1:m,
        ii = mod(round(rand*111111),sumfk)+1;
        tmp=0; kk=1;
        while (tmp < ii),
            tmp = tmp + fk(kk);
            kk = kk+1;
        end
        kk=kk-1;
        A(j,Son(kk)) = 1;
        A(Son(kk),j) = 1;
    end

    Son = union(Son, j);

end
```

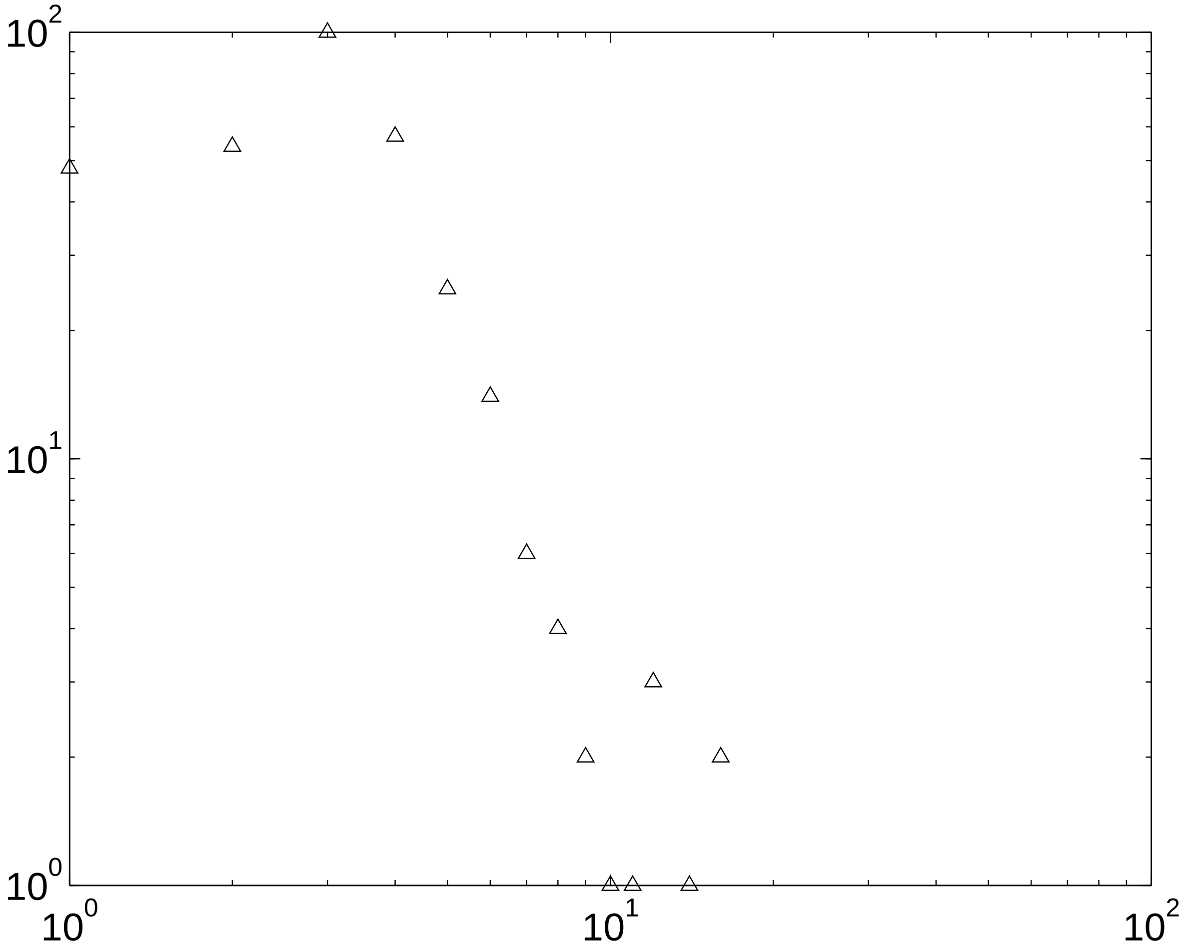
```
    Nnew = length(olindex);
    tnew = rand(1,Nnew);
    Pnew = 1-pp(lt(olindex));
    SignNewOn = sign(Pnew - tnew);
    NewOnFlag = olindex.*SignNewOn;
    olnewindex = intersect(olindex, NewOnFlag);

    OFFnew = setdiff(olindex,olnewindex);
    Ndel = length(OFFnew);
    A(:,OFFnew) = zeros(n,Ndel);
    A(OFFnew,:) = zeros(Ndel,n);
    lt(OFFnew) = 0;
    Son = union(Sontmp, olnewindex);

end
```

Number of nodes on-line

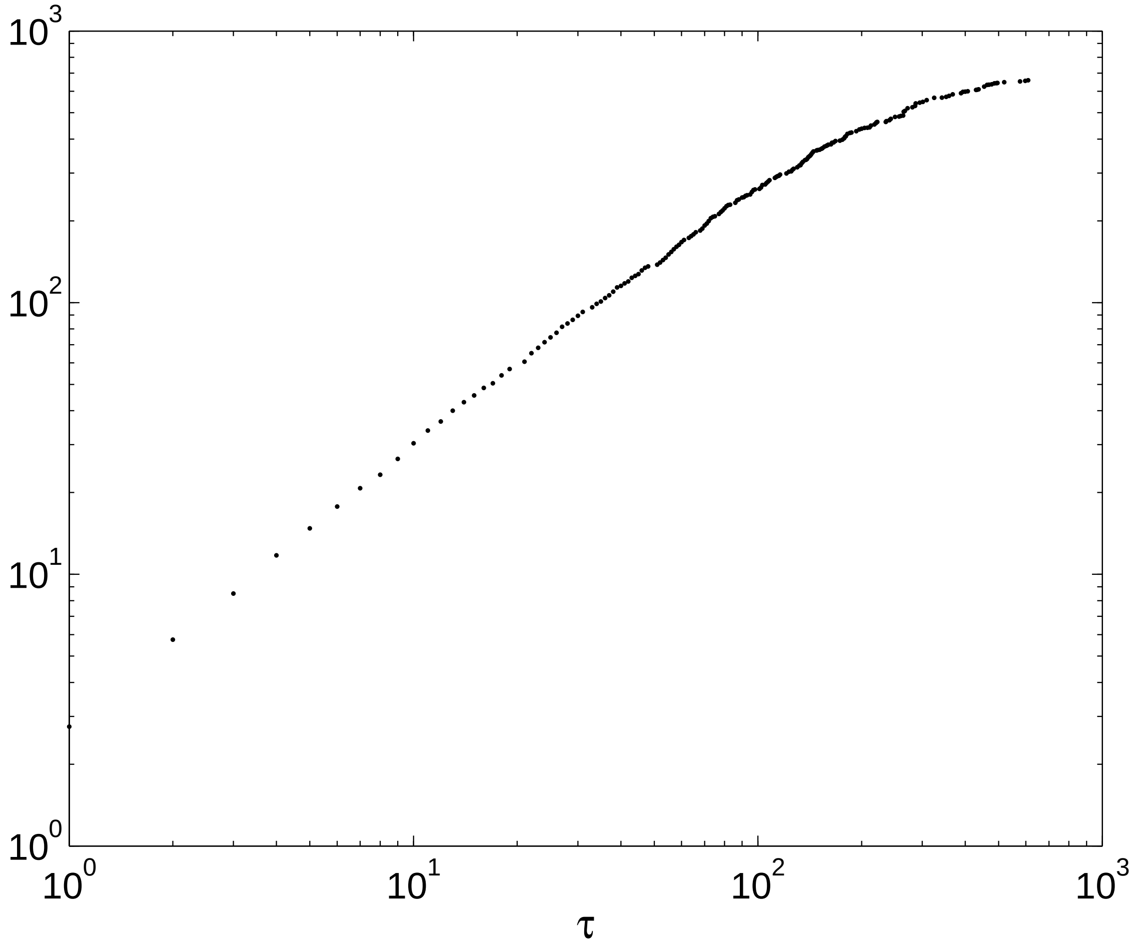


Data is taken at every 8 time steps.

Node degree distribution at $t = 3000$, $\gamma \approx 3$ ($\neq 4$?)



Expected degree of connection versus τ, $\beta \approx 1$



It should be noted that

- node degree generated from the simulated model seems not the same as the one measured from Gnutella and

- the slope of expected degree and the slope of node degree distribution do not fit the scaling relation
$$\gamma = \frac{1}{\beta} + 1.$$
Since $1 + 1/\beta \approx 2 \neq \gamma \approx 3$.

So, *what should be the underlying evolving model for Guntella P2P ?*

Conclusion

- Historical background of small world networks, scale-free networks

- Importance of scale-free network

- Application of the power law distribution

- Modeling Gnutella

References

Reka Albert and Albert-Laszlo Barabasi, Statistical mechanics of complex network, *Review of Modern Physics*, Vol. 74, 47-97, 2002.

Albert-Laszlo Barabasi, *Linked: The science of networks*, Persus, Cambridge, MA 2002.

Stefan Bornholdt and Heinz Geog Schuster (eds), *Handbook of Graphs and Networks: From the genome to the Internet*, Wiley-Vch, 2003.

S.N. Dorogovtsev and J.F.F. Mendes, Evolution of networks, *Advances in Physics* Vol.51, 1079-1187, 2002.

S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of networks: From biological nets to the Internet and WWW*, Oxford, 2003.

Marc E.J. Newman, The structure and function of complex networks, *SIAM Review*, Vol. 45, 167-256, 2003.