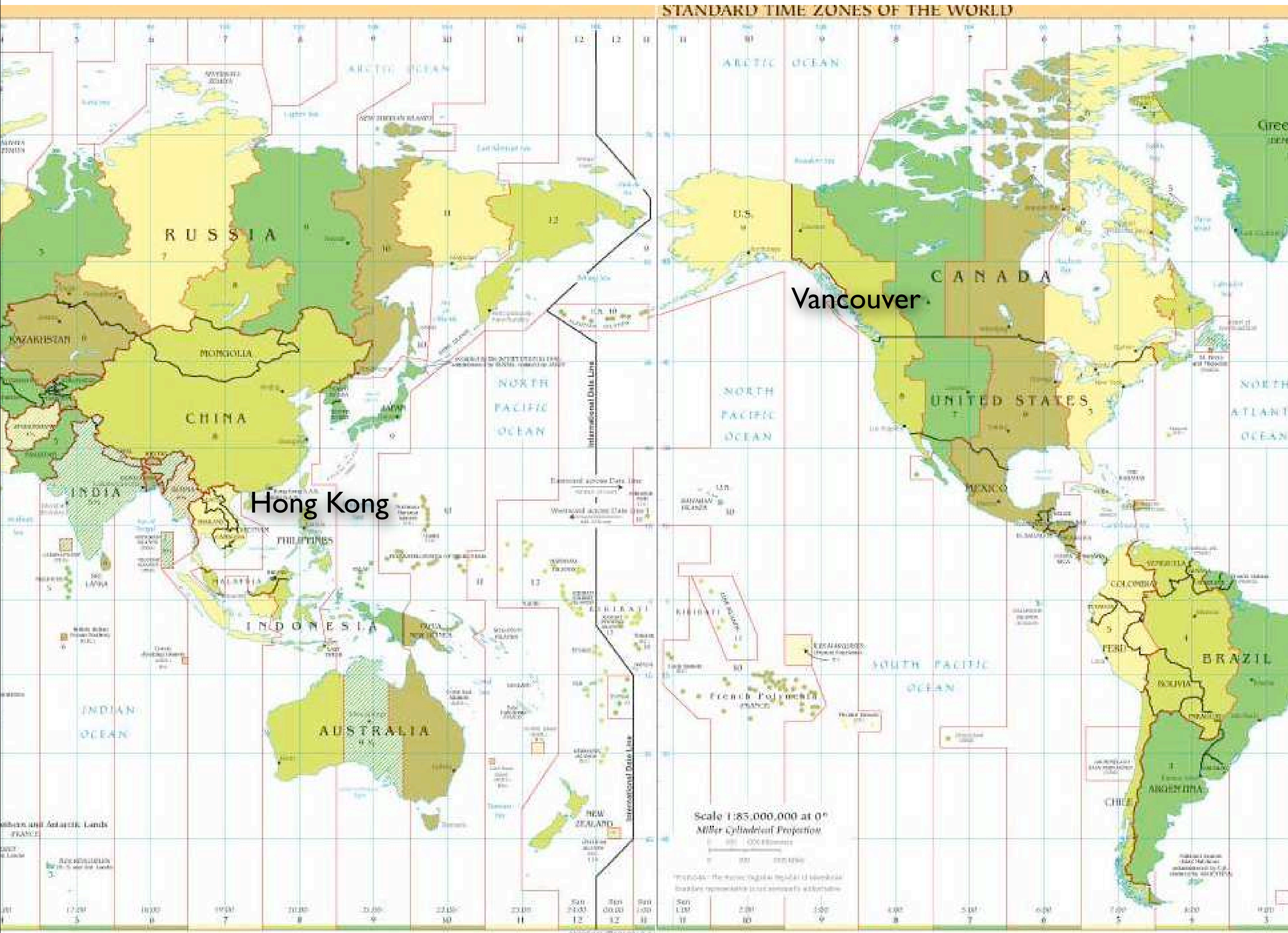
Applications of Complex Networks in Traffic Analysis of Communication Networks

C.K. Michael Tse
Hong Kong Polytechnic University
Email: cktse@ieee.org


STANDARD TIME ZONES OF THE WORLD
Vancouver
Hong Kong
ARCTIC OCEAN
RUSSIA
KAZAKHSTAN
MONGOLIA
CHINA
INDIA
JAPAN
PHILIPPINES
INDONESIA
AUSTRALIA
NEW ZEALAND
NORTH PACIFIC OCEAN
INDIAN OCEAN
International Date Line
Eastward across Date Line
Westward across Date Line
KIRIBATI
French Polynesia
U.S.
CANADA
UNITED STATES
MEXICO
COLOMBIA
PERU
BRAZIL
BOLIVIA
CHILE
ARGENTINA
SOUTH PACIFIC OCEAN
Scale 1:85,000,000 at 0°
Miller Cylindrical Projection

Nonlinear landing at the chaotic old Kai Tak Airport (before 1997)

Hong Kong Airport now


YUE HWA

Applications of
Complex Networks
in Traffic Analysis of
Communication Networks

C.K. Michael Tse
Hong Kong Polytechnic University
Email: cktse@ieee.org

Acknowledgment

The work described in this lecture has been the results from collaboration with the following people:

Wai Man Tam, Yongxiang Xia, Francis Lau and Michael Small.

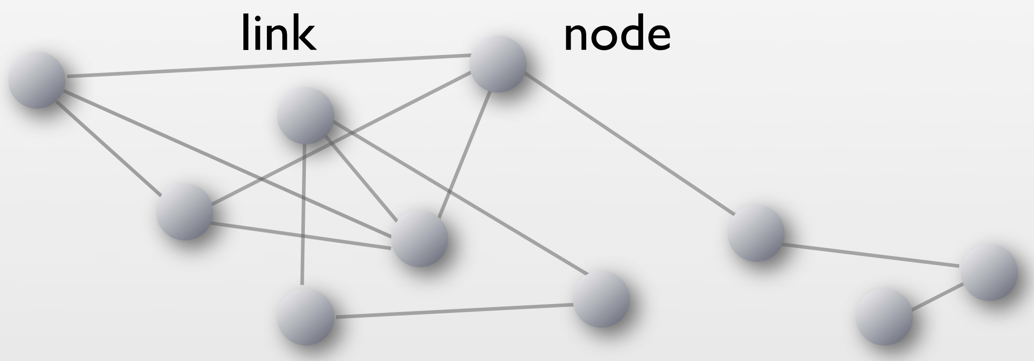
Special thanks to Prof. Ljiljana Trajkovic for organizing this workshop in Vancouver and for her kind invitation.

Outline

- Networks
  - Regular (uniform), Small World, and Random
  - Scale-free networks
- Telephone User Network Model
  - Scale-free user network model
  - Call process, call blockings, channel capacity threshold effect
  - Traffic analysis and simulations
- Extension to Cellular Networks
  - Scale-free user network model and mobility model
  - Handoff failures and call drop probabilities
- Extension to Internet
  - Basic model with multiple power laws
- Conclusions

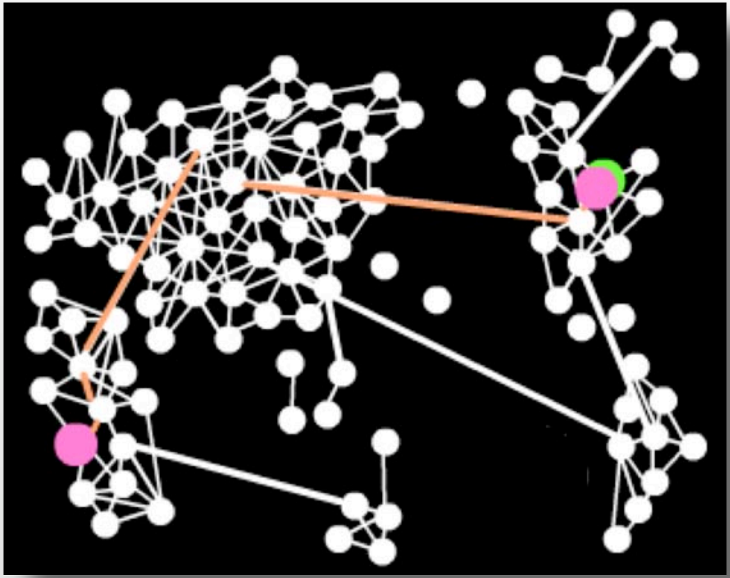
What is a network?

Networks



- Parameters:
  - number of nodes = *n*
  - average number of links per node = *k*
  - average shortest path lengths between nodes = *L*
  - clustering coefficient (probability that nodes form groups of 3) = *C*

“6-Degrees of Separation”

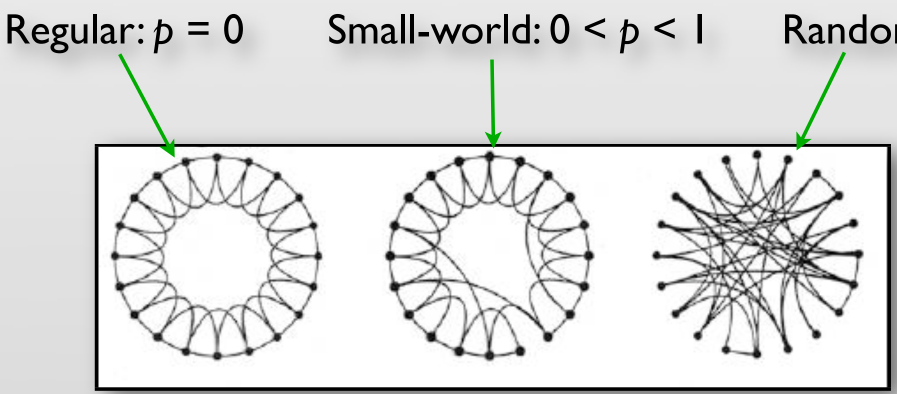


- Early study of “distance” between 2 persons in USA by Stanley Malgram in 1967 (published in Psychology Today)
  - Average distance between any two persons in the USA is SIX.

Small World Networks

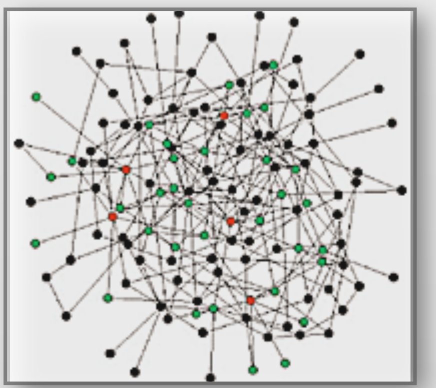
Regular versus Random Networks

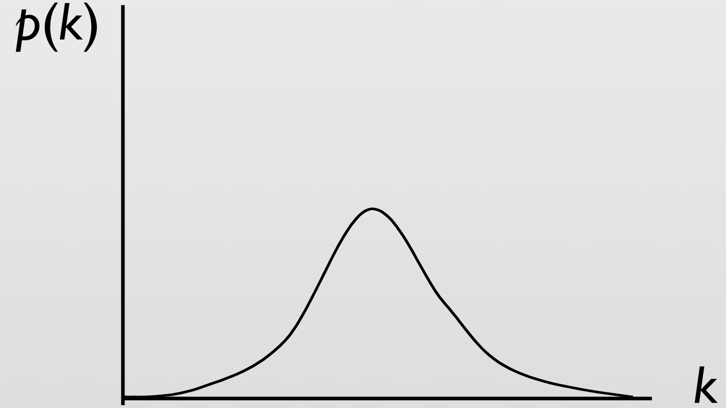
- Begin with a regular network, e.g., $n = 20$, $k = 4$
- Remove links randomly.
- Unlink each pair of linked nodes with probability p and re-link two randomly chosen nodes.



Random Network

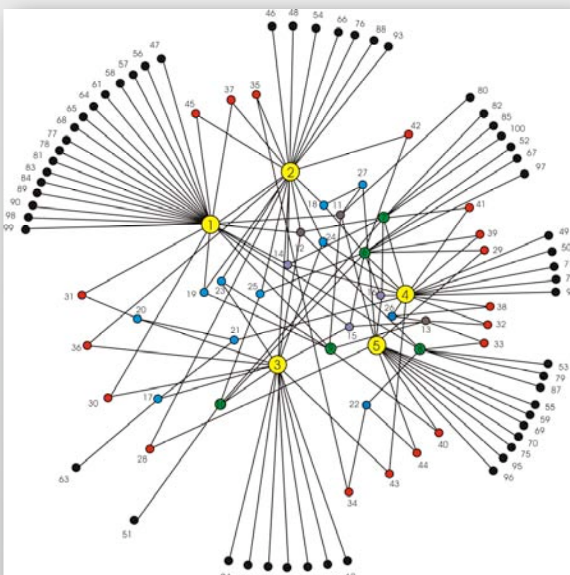
- The links are completely democratic.
    - Probability of a node having k links = $p(k)$
        - $p(k)$ follows a **bell curve** distribution

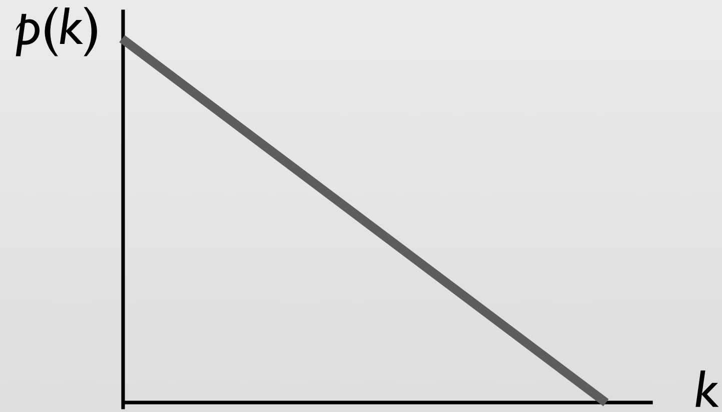




Scale-free (Power Law) Network

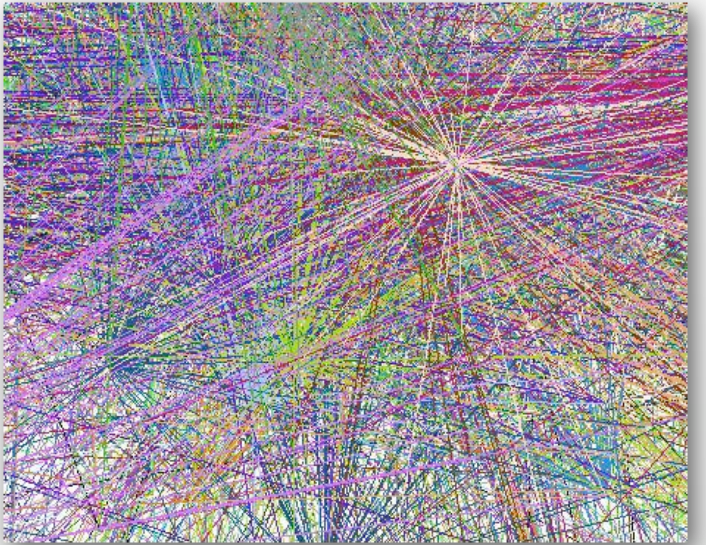
- A minority of nodes have a majority of links
  - *p*(*k*) follows a **power law** distribution





The 80/20 Universal Rule

- Growth and "preferential attachment"
- Probability that node A will connect to node B depends
  - number of links incident on node B (older nodes are more likely to be attached by incoming links)
  - "fitness" or relevance of node B (better nodes are more likely to be attached by incoming links)



Examples:

Barabási* proved that the Internet is scale-free and fewer than 20% of the web pages have at least 80% of incoming links (e.g., Yahoo, Amazon, eBay, Google ...)

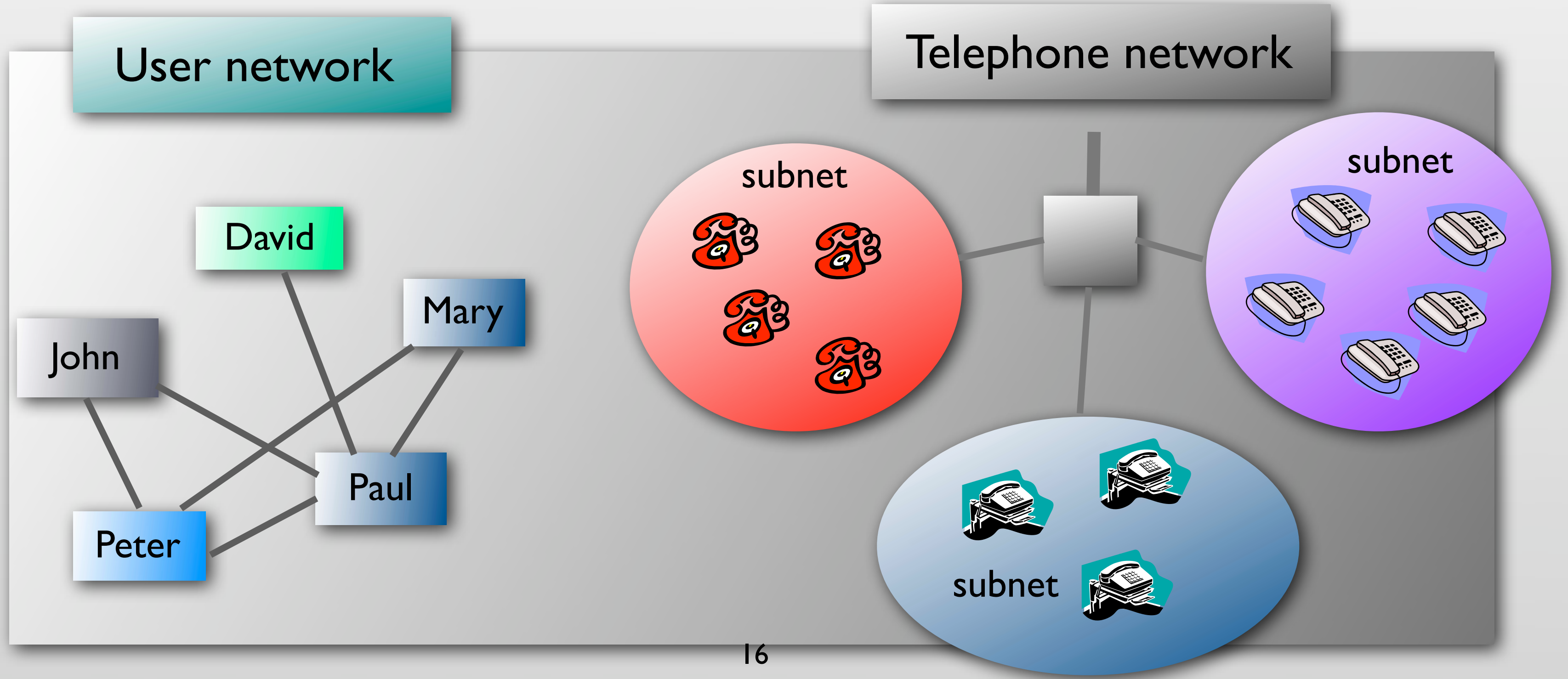
In economics, less than 20% of people own more than 80% of the wealth.

*A. Barabási, *Linked,* 2nd edition, New York: Plume Books, 2002.

Telephone network

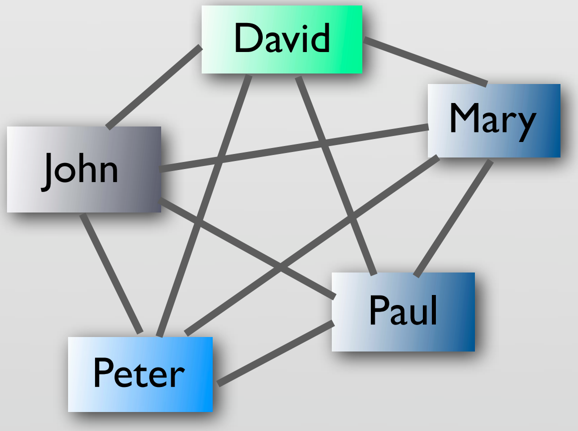
Telephone Network

- The telephone consists of ***two*** networks:
  - User connection — ??
  - Telephone exchange — fixed topology, e.g., cellular, hierarchical, etc.

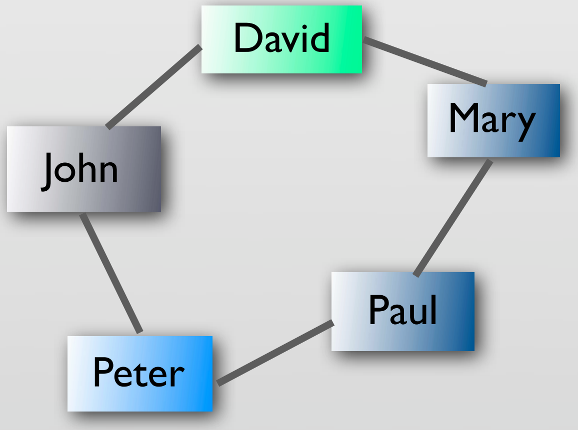


Traditional User Network Model

- Regular/uniform user network
  - Each user can call a fixed number of users and each with equal probability
  - Fully democratic!



k = 4

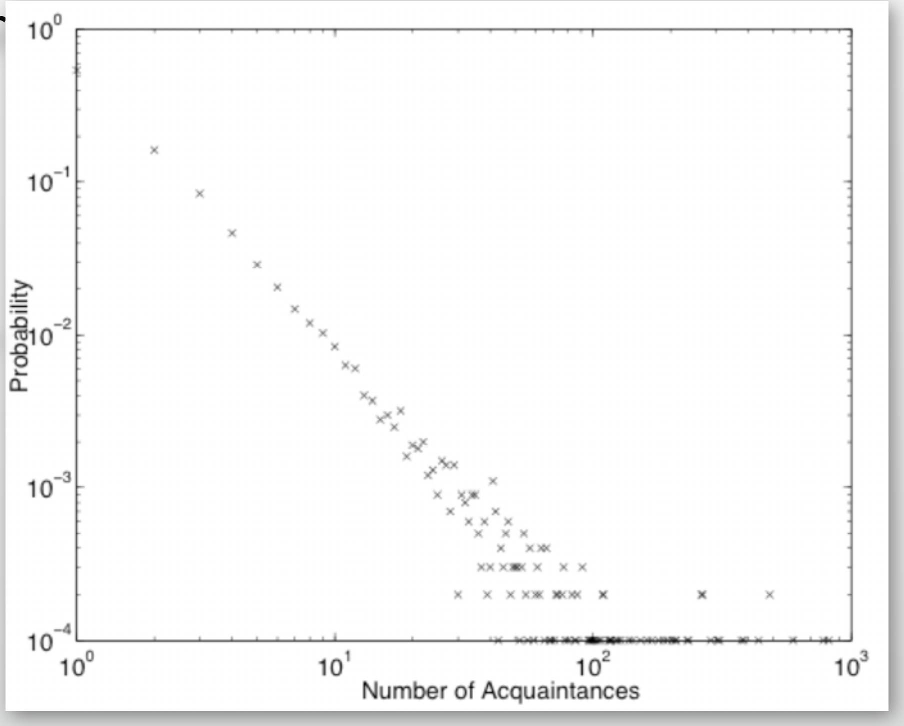


k = 2

Scale-free User Network

- Scale-free network is assumed for the user connection
- $p(n_i)$ follows a power law distribution

$$p(n_i) \sim n_i^{-\gamma}$$



where

n_i = number of links of user *i*

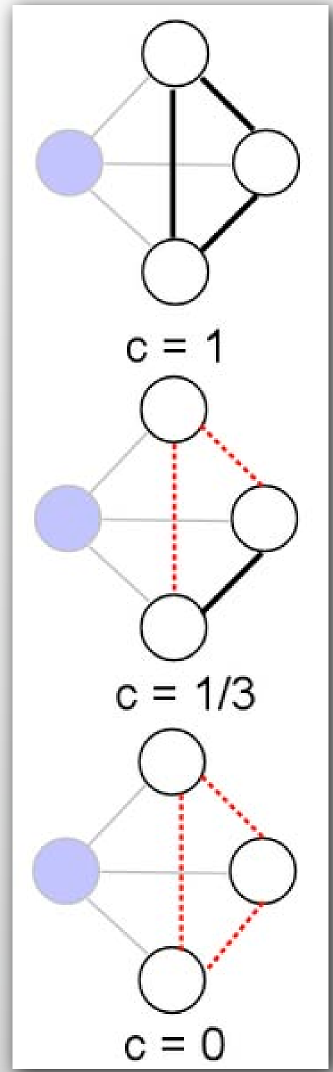
γ = characteristic exponent

Small-world network: Clustering

- One criterion used to determine if a network is small-world is the interrelatedness of the nodes within the network. Basically we use the parameter called **clustering coefficient, C,** which is the probability of any two nodes connected to a given node are themselves also connected.

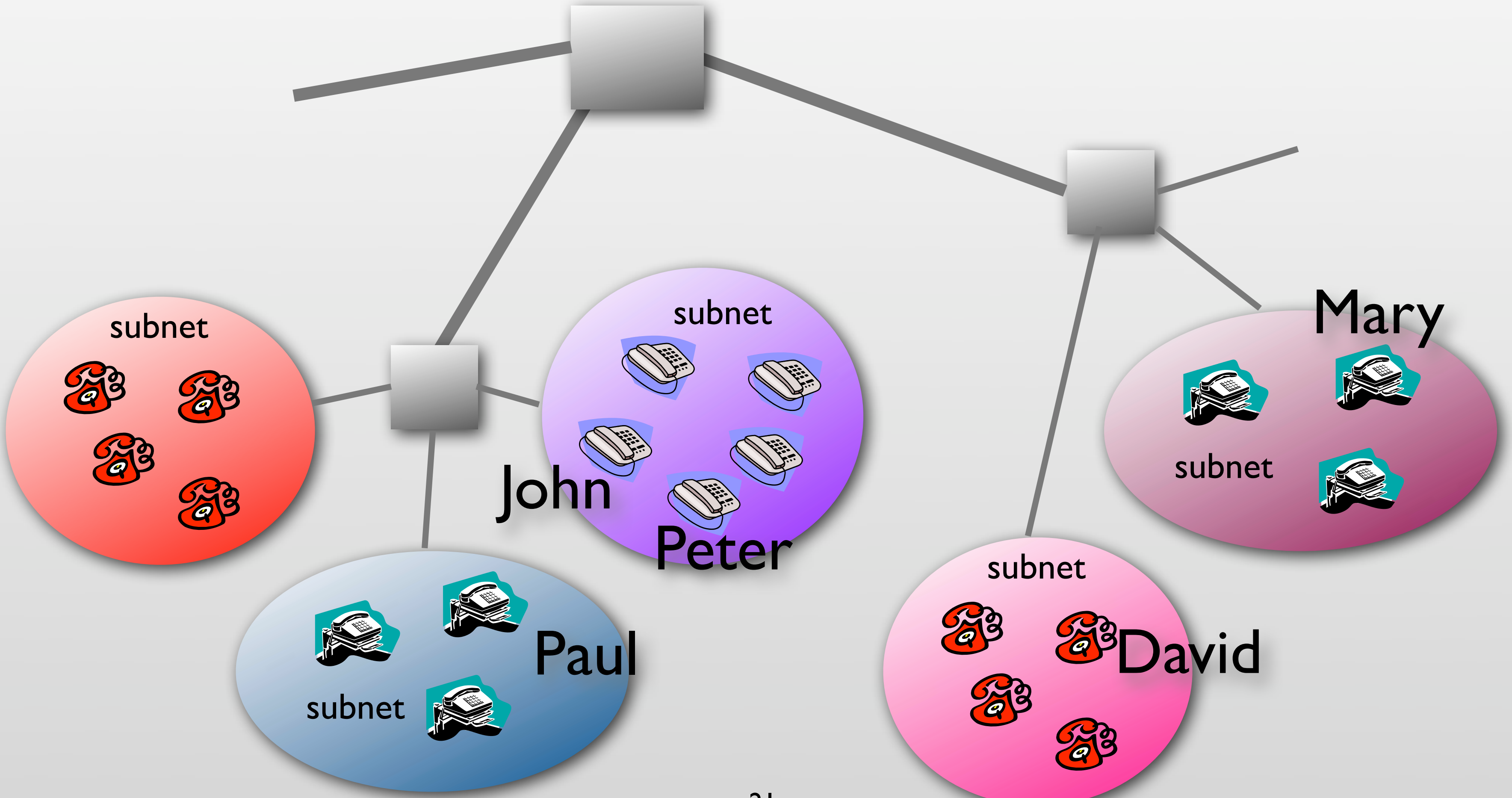
- If C = 1, the clustering is extreme. If C = 0, there is no clustering.

- The clustering coefficient of the whole network is just the average of all nodes.



Telephone network meets user network

Telephone Network meets User Network



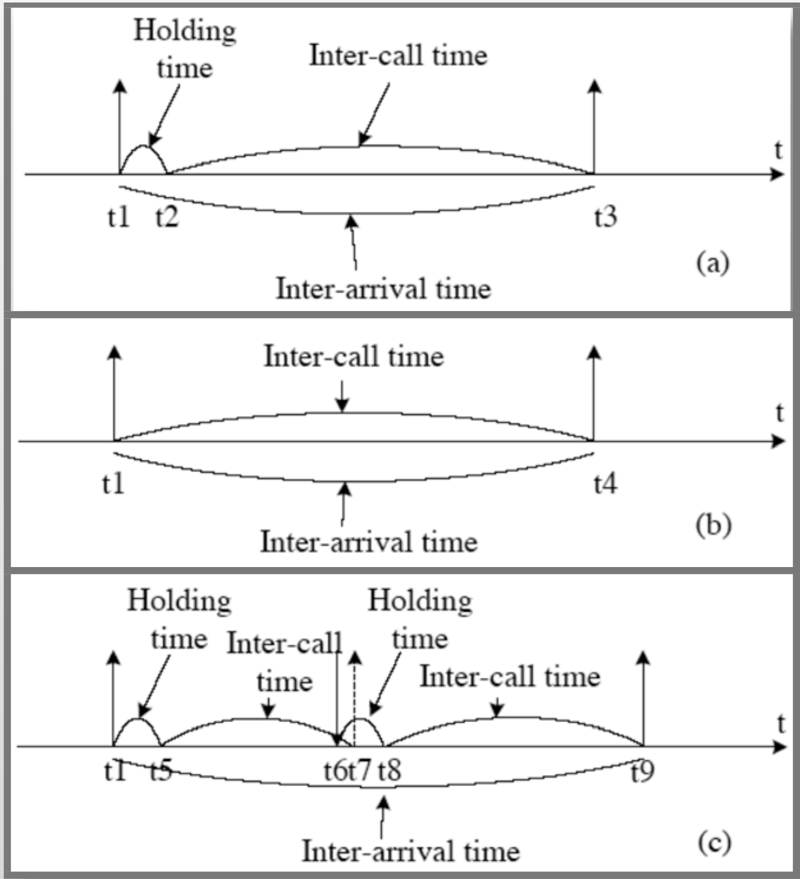
Call processes

- Usual call process
  - holding time t_m and inter-call time $1/\mu_i$ are assumed to exponentially distributed:

$$\text{PDF of holding time} = \frac{1}{t_m} e^{-t/t_m}$$

$$\text{PDF of intercall time} = \mu_i e^{-\mu_i t}$$

- Call blocking
  - call blocked due to receiver being engaged
- Call cancelling
  - Outgoing call occurs during an incoming call. Then, this outgoing call will be cancelled.



Traffic

- Traffic crudely (in layman language) refers to the amount of call establishments in a telephone network that goes through a particular (set of) point(s) per unit time in the physical network. Literature has more precise definitions.*
  - ***carried traffic*** (call per min)
  - ***call blockings/failures*** (call per min)
  - ***channel capacity*** (allowable number per min; related to investment/resources)

*K. Anurag, D. Manjunath, and J. Kuri, *Communication Networking: An Analytical Approach,* Amsterdam: Elsevier/Morgan Kaufmann, 2004.

Traffic Intensity

- If we observe the traffic for a long time, we get an average impression of the call arrivals.
- Suppose user *i* has an average number of call arrivals per minute = λ_i
- The average number of call arrivals per min in the network is

$$\lambda = \sum_{i=1}^{N} \lambda_i \quad \text{where } N \text{ is the number of users}$$

- Offered traffic intensity (average activity over a period of time)

$$\lambda_{\text{offered}} = 2\lambda t_m$$

- Carried traffic intensity

$$A_{\text{carried}} = A_{\text{offered}}(1 - p_{\text{blocking}}) = 2\lambda t_m(1 - p_{\text{blocking}})$$

Simulations

Simulations

- Parameters:
  - Number of users = N = 10000
  - Number of subnetworks = M = 4
  - Average number of acquaintances per user = $\overline{n}$ = 5
  - Average inter-call rate = $\overline{\mu}$ = 0.01 call/min
  - Average holding time = $\overline{t_m}$ = 4 min

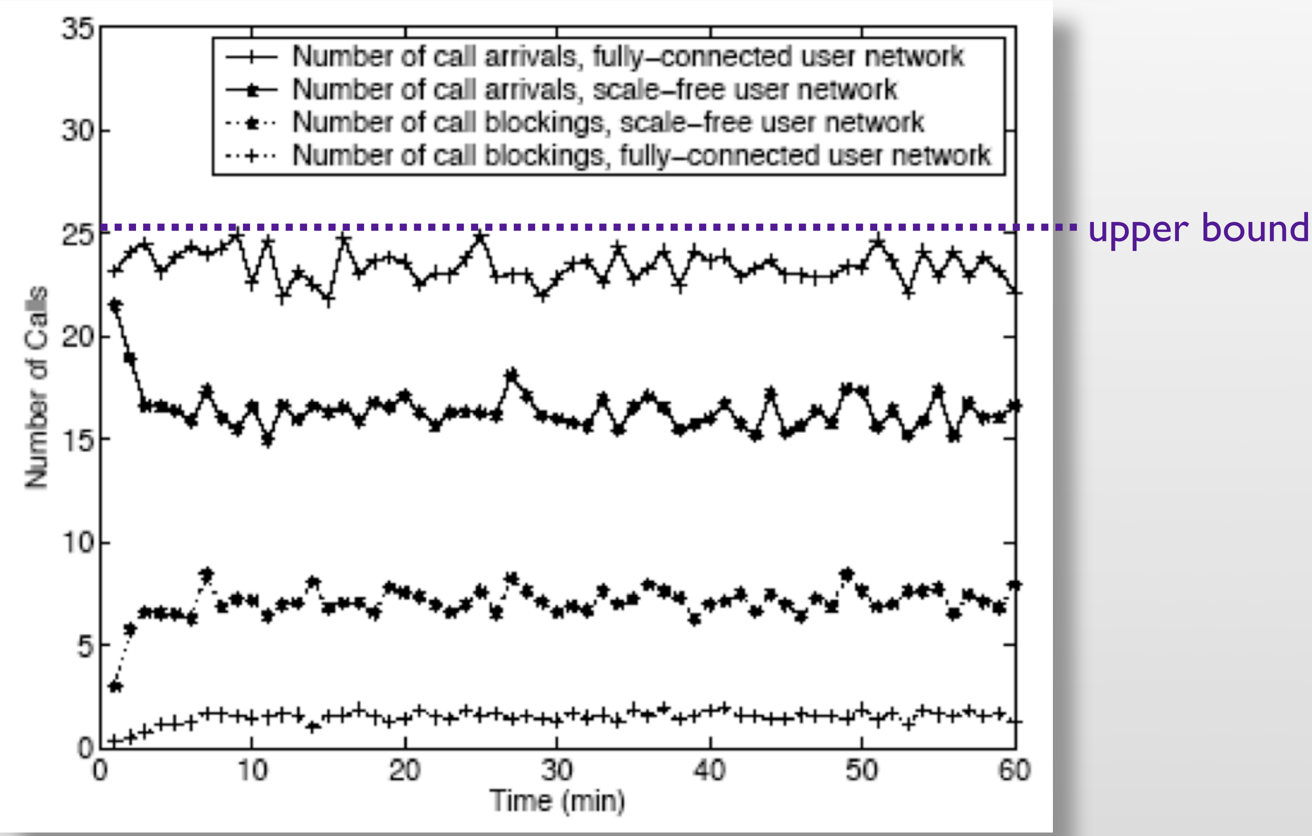
  - Our basic assumption is that $\mu_i = p_0 n_i$ (which means that the inter-call time is shorter if a user has more acquaintances) where p_0 is a proportionality constant.
  - Typically, each user contributes 0.025 to 0.05 erlang.
  - So, p_0 ≈ 1/500 call/min/acquaintance
  - Upper bound of λ is

$$\lambda = \sum_{i=1}^{N} \lambda_i < \sum_{i=1}^{N} \mu_i = p_0 \sum_{i=1}^{N} n_i = N p_0 \bar{n} = 25 \text{ call/min/subnetwork} = 100 \text{ call/min}$$

Simulations

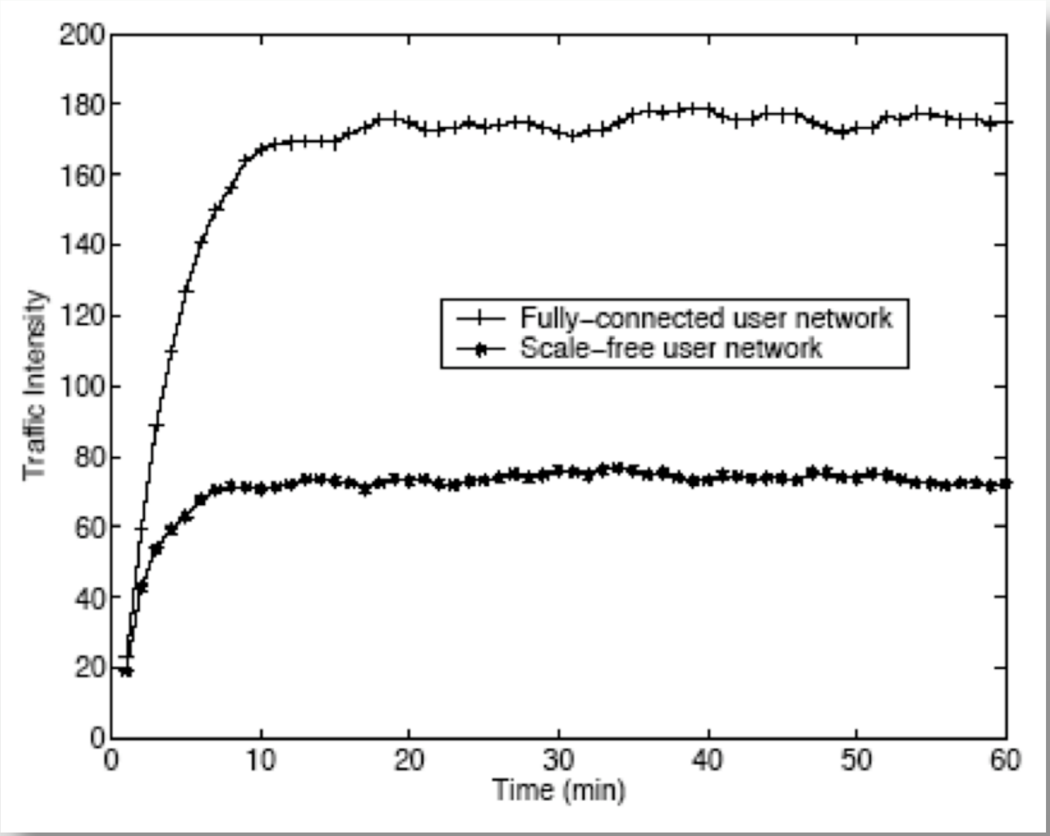
- We compare
  - fully-connected (regular) user network (each user can call any one with equal probability)
  - scale-free user network (each user has a fixed set of acquaintances and the number of acquaintances follow a power-law distribution)

Call arrivals/blockings



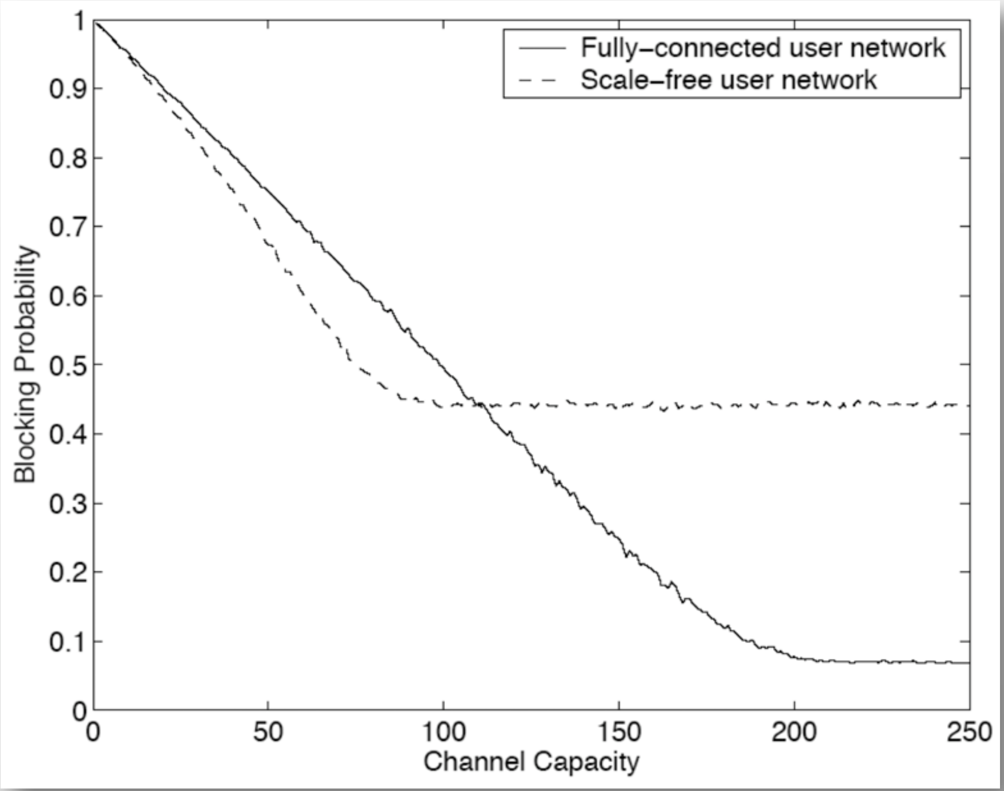
Scale-free user network has lower call arrivals and higher blockings.

Traffic intensity

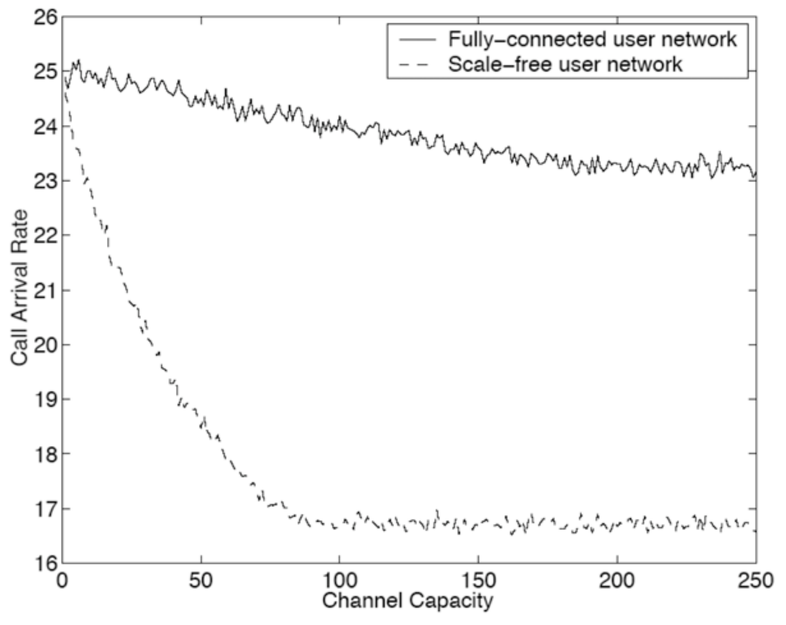


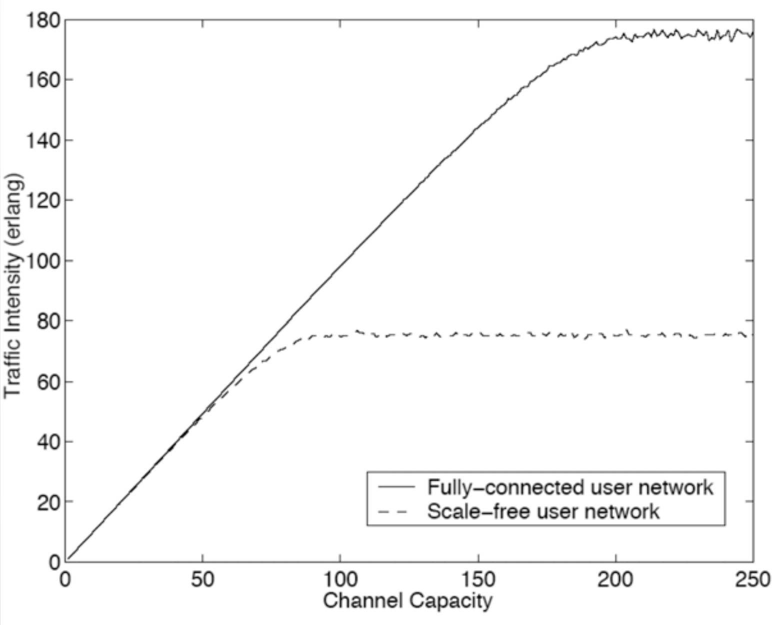
Scale-free user network has much lower traffic intensity, i.e., more call cancellings and blockings.

Threshold effect

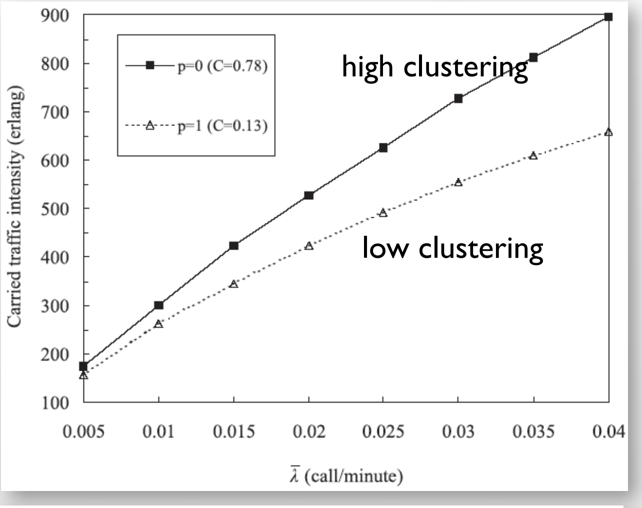


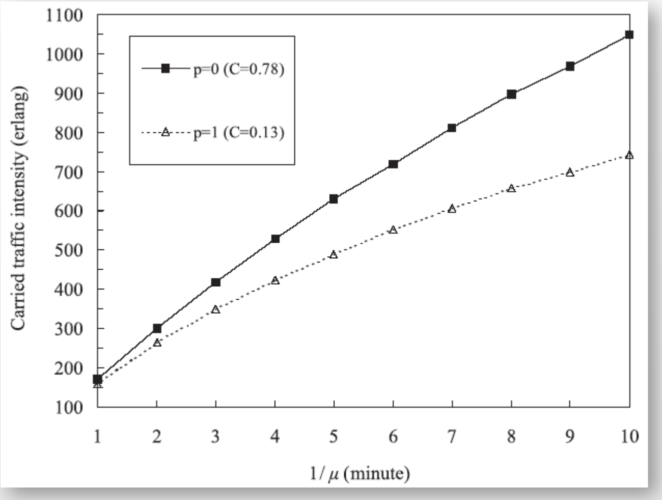
More threshold effects

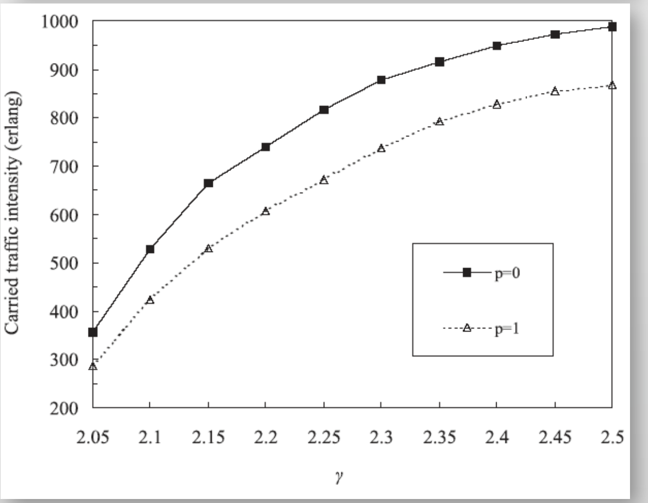




Clustering: small-worldness







Small-worldness has a profound effect on the traffic.

More highly clustered networks generally can carry more traffic.

Can it be improved?

Increase the network capacity (invest more).

Yes and no.

Yes, when the threshold is not reached.

No. Beyond the threshold, expanding network capacity won't help.

Other solutions:

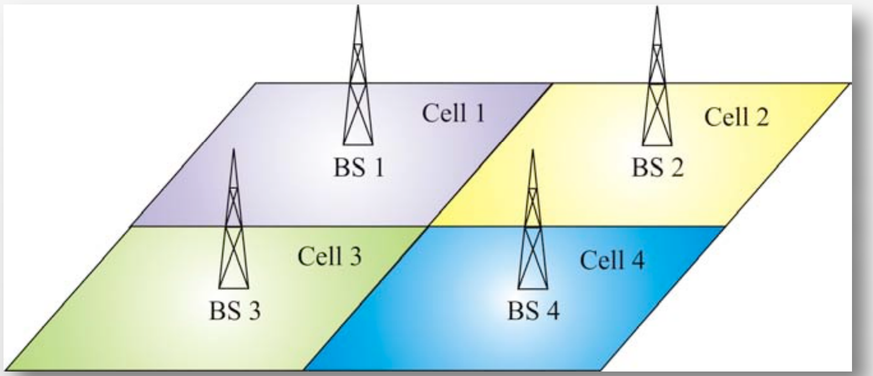
- Altering the user network
- Pricing strategy

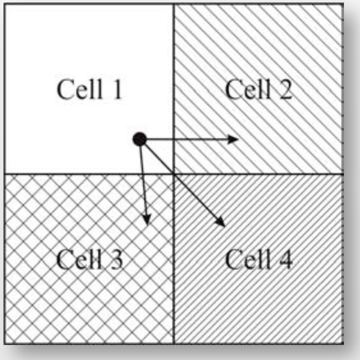
Some messages

- The user network configuration plays an important role in determining the traffic intensity.
- The "scale-free-ness" of the network is accounting for most of the blocked calls, and is the only reason when network capacity has reached the threshold.
- Small-worldness (clustering) is also having significant effects on the carried traffic
- Increasing capacity (intensifying investment) is not going to help!
- Telephone network providers should implement other policies that can alter the user network parameters in order to ease traffic,e.g., through pricing plans.

Extension to cellular networks

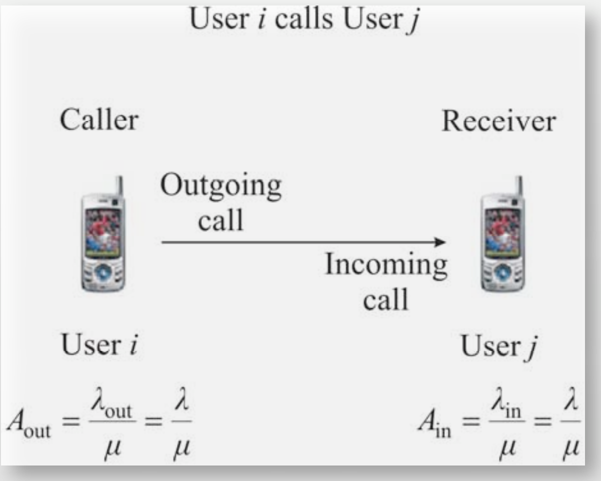
Cellular networks





- In addition to the scale-free user network, mobility model has to be considered. We may consider
  - zero mobility
  - identical mobility
  - power-law (scale-free) mobility

Traffic model



Arrival rates of incoming and outgoing calls: λ_{in} and λ_{out}

Call departure rate (completion rate): μ

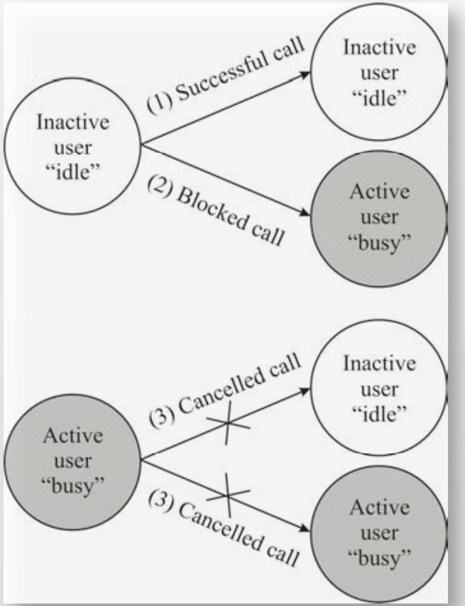
Offered traffic intensity: A

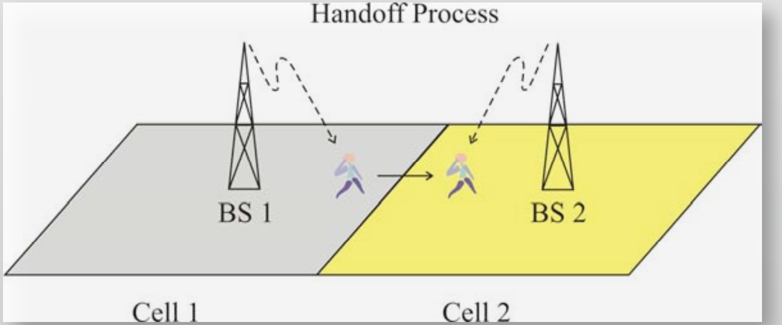
$$A = \frac{\lambda_{\text{in}} + \lambda_{\text{out}}}{\mu} = \frac{2\lambda}{\mu}$$

A traffic intensity of 1 erlang implies that a channel will be occupied for 60 minutes during an one-hour period.

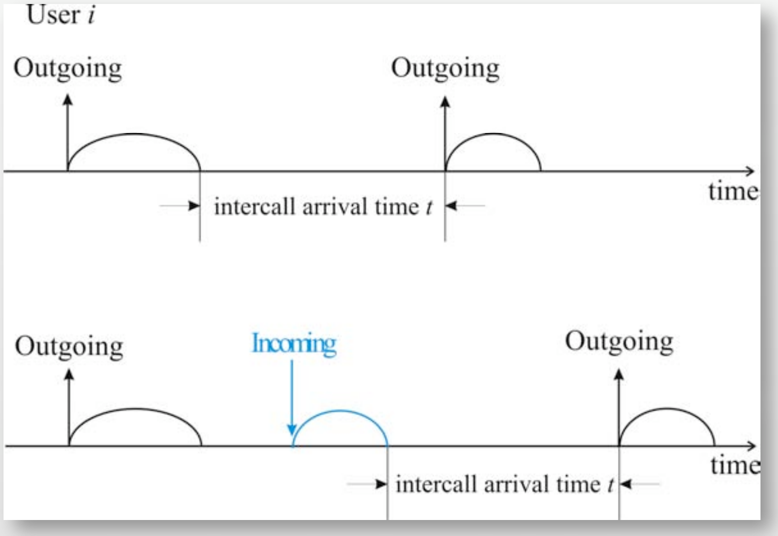
Other considerations

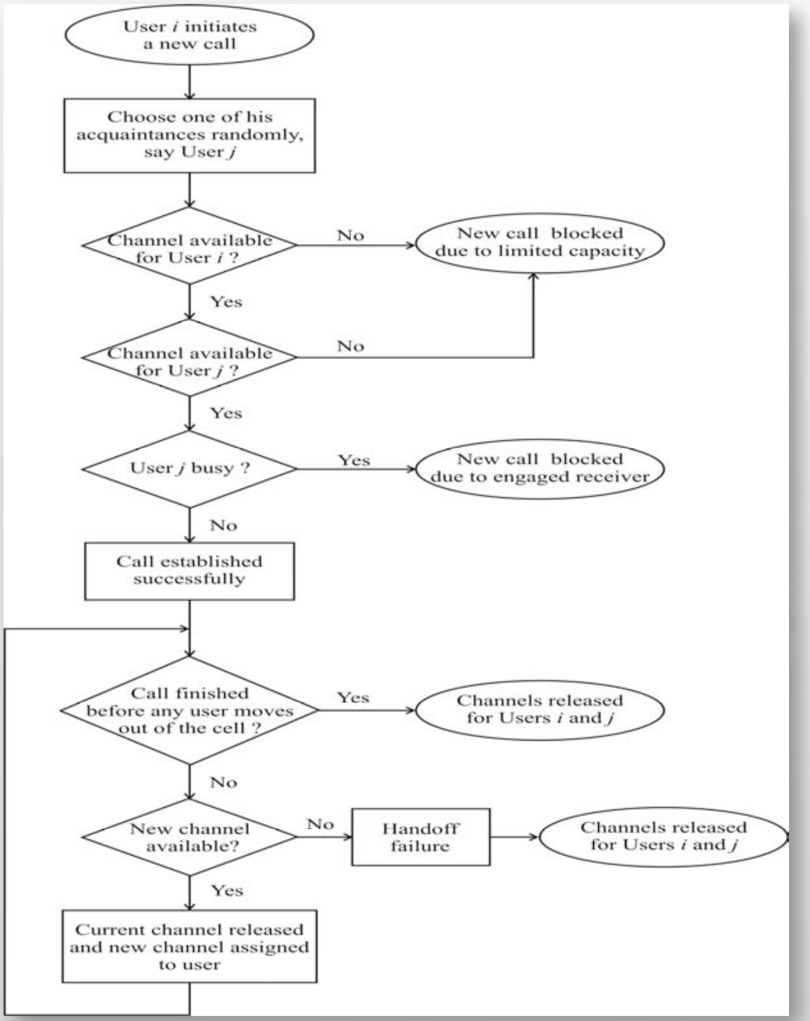
- Call establishments
  - States of callers and callees
    - caller busy — call cancellation
    - callee busy — call blocking
  - Channel availability
    - a pair of channels required for call establishment
  - Handoff processes





Traffic flow





The calls can be classified into two types:

Complete call — established and wholly finished call.

Incomplete call — blocked call during call initiation or terminated call during handoff.

Model assumptions

Telephone call process is similar to the conventional telephone network.

User network: power law distribution $\Pr(n_i) \sim n_i^{-\gamma}$, λ can follow other distributions

Call holding time: exponential distribution $f(T_c) = \mu e^{-\mu T_c}$

Mobility: The time that a user stays in the same cell (called ***cell residence time***), denoted by τ_m, is exponentially distributed with mean $1/\theta_i$ for the *i*th user. The pdf of τ_m thus equals

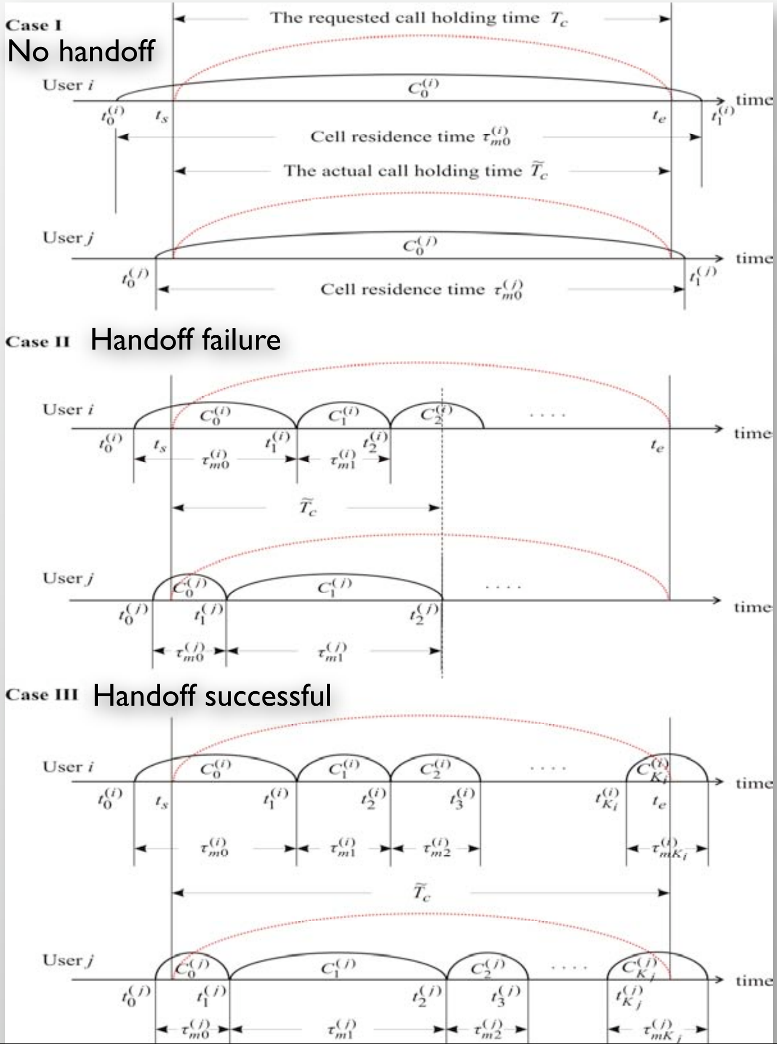
$$f(\tau_m) = \theta_i e^{-\theta_i \tau_m}$$

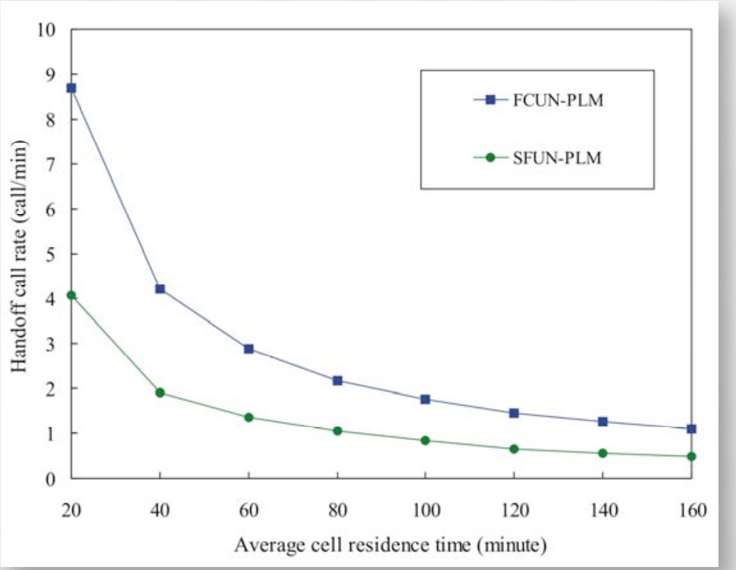
Zero mobility: Users stationary

Identical mobility: $\theta_i = \frac{1}{N}\sum_{i=1}^{N}\theta_i = \overline{\theta}$

Power-law mobility: $f(\theta_i) \sim \theta_i^{-\nu}$

Handoff call arrival vs average cell residence time





Comparison of models

	User network	Mobility model
Conventional model	Fully-connected	Identical mobility
Scale-free model	Scale free 1. The majority of the users have a few acquaintances. 2. A small number of users (sometimes referred to as super users) have a large number of acquaintances.	Power law mobility 1. A large number of users (e.g., people working in an office) stay in the serving cell for a long period of time. 2. A small number of users (e.g., people on the bus or train) move from cell to cell frequently.

Performance indicators of cellular networks

$\tilde{A}$ carried traffic, which is the summation of the service time durations within a given time period (e.g., 1 min);

p_{b1} new call blocking probability due to limited capacity, or

$$p_{b1} = \Pr[\text{new call blocking due to limited capacity} \mid \text{new call arrival}] = \frac{\text{Number of new call blockings due to limited capacity}}{\text{Number of new call attempts}};$$

p_{b2} new call blocking probability due to user engagement, or

$$p_{b2} = \frac{\text{Number of new call blockings due to user engagement}}{\text{Number of new call attempts}};$$

p_b overall new call blocking probability, or

$$p_b = \frac{\text{Number of new call blockings}}{\text{Number of new call attempts}} = p_{b1} + p_{b2};$$

p_f handoff failure probability, or

$$p_f = \Pr[\text{handoff failure} \mid \text{handoff call arrival}] = \frac{\text{Number of handoff failures}}{\text{Number of handoff call attempts}};$$

p_{nc1} call incompletion probability due to new call blockings, or $p_{nc1} = p_b$;

p_{nc2} call incompletion probability due to handoff failures, or

$$p_{nc2} = \frac{\text{Number of handoff failures}}{\text{Number of new call attempts}} (\neq p_f);$$

p_{nc} overall call incompletion probability, or $p_{nc} = p_{nc1} + p_{nc2}$.

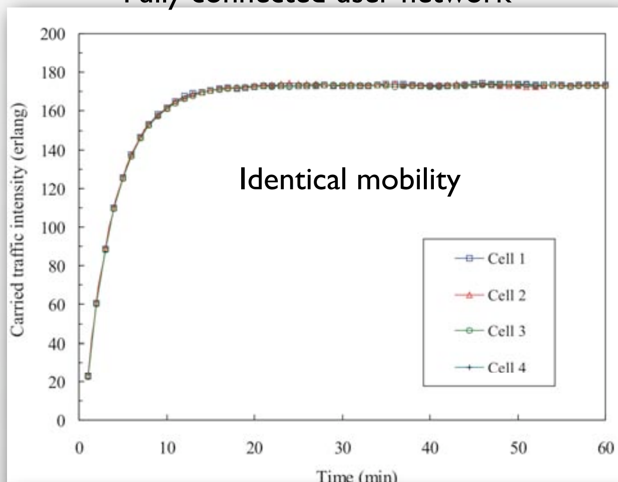
To be analyzed under

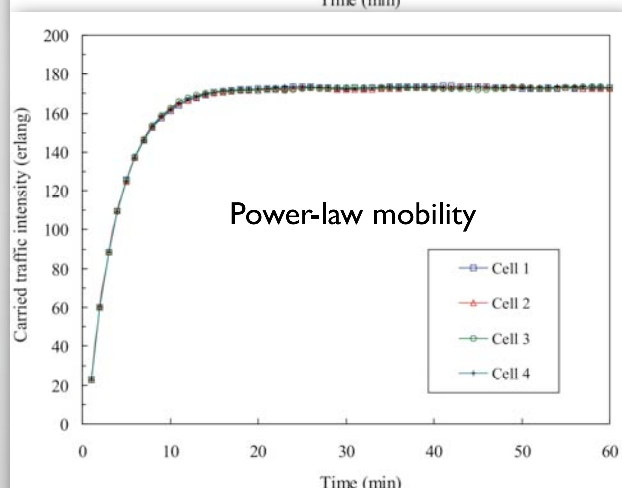
1. Unlimited channel capacity
2. Finite channel capacity

Simulations

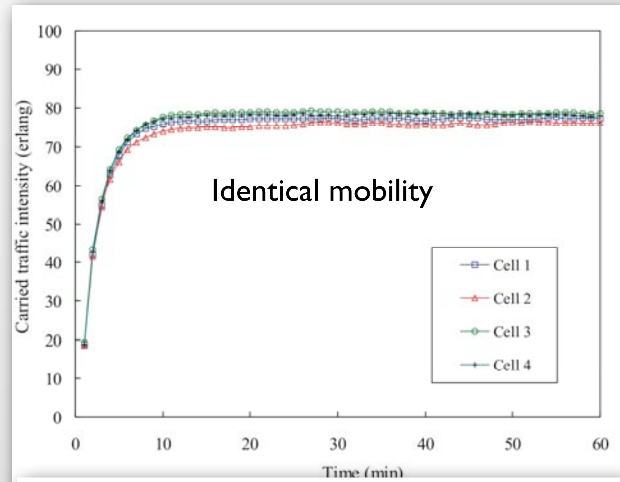
Performance under unlimited channel capacity

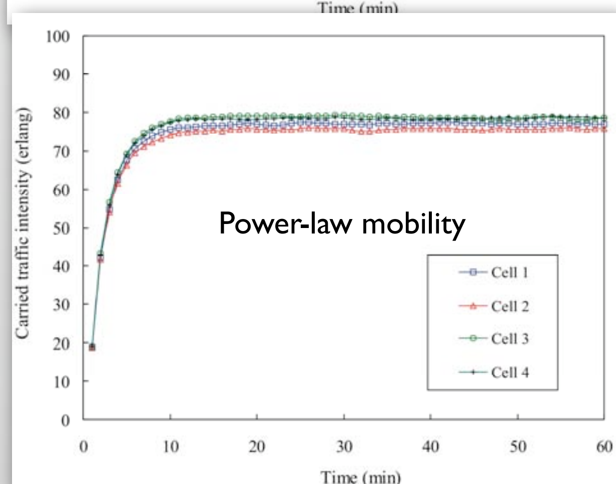
Fully connected user network



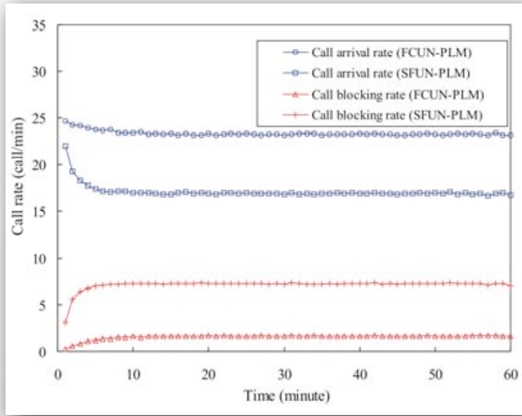


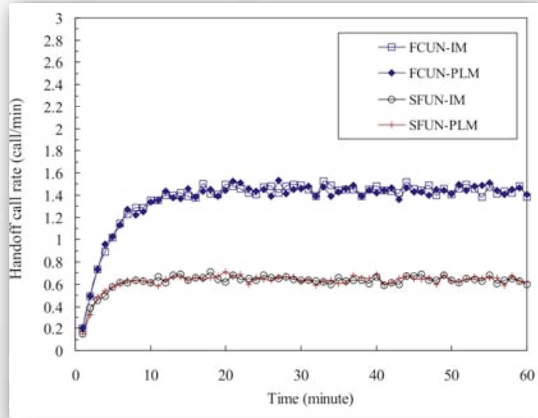
Scalefree user network



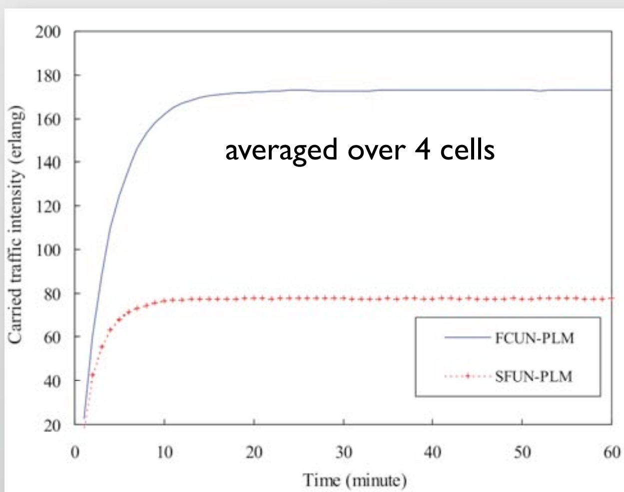


Performance under unlimited channel capacity





New call cancellation +
New call arrival rate –
Handoff call rate –

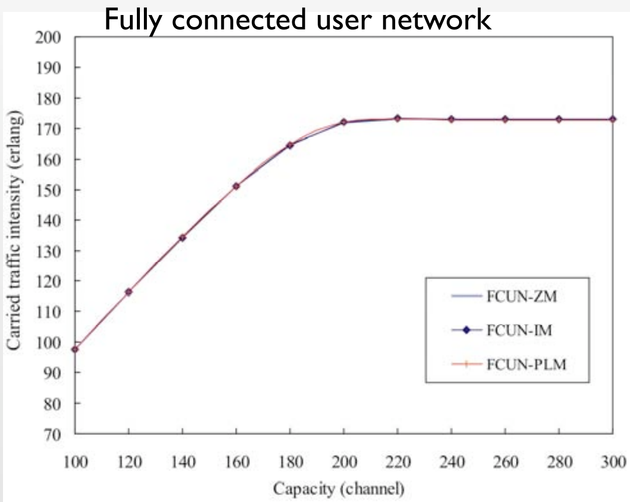


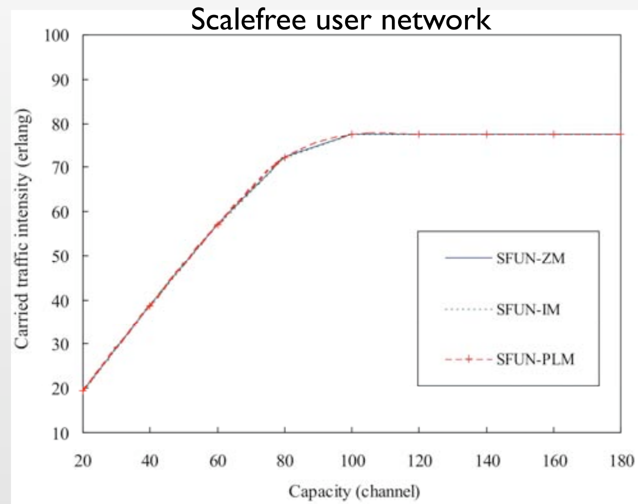
New call cancellation +
Carried traffic intensity –

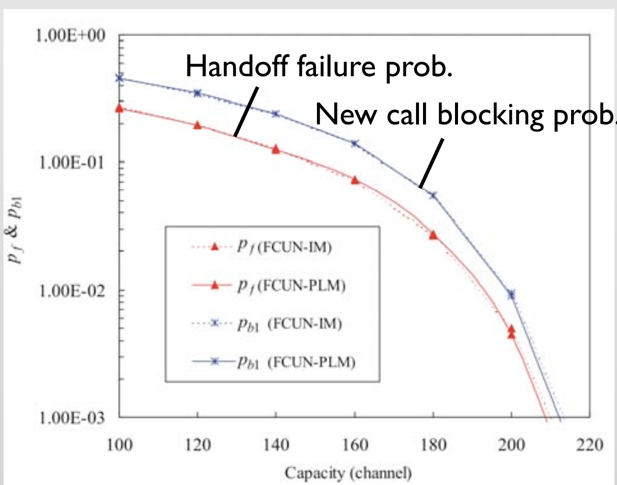
and

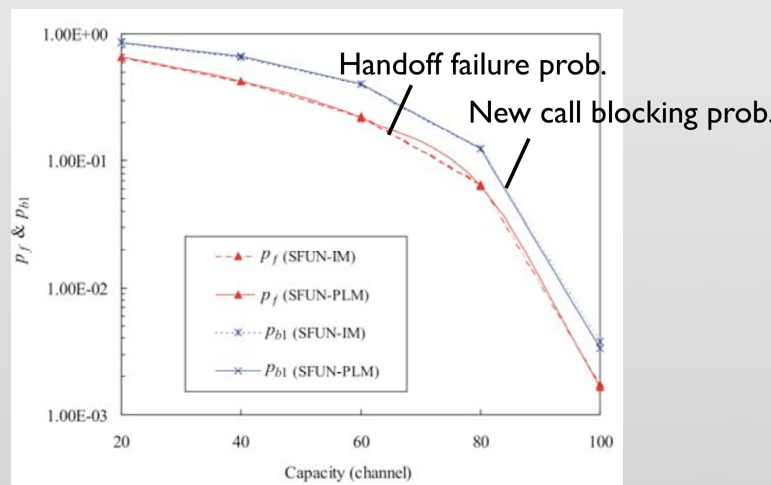
New call blocking +
Carried traffic intensity –

Performance under limited channel capacity









Some messages

- The user network configuration plays an important role in determining the traffic intensity.
- The handoff rates are much lower for scale-free user networks because of the less number of established calls.
- Compared to the user network, mobility models seem to play a less crucial role in determining the traffic. More investigations needed.
- Questions:
  - what if call times become power law distributed?
  - what about independent power laws for call times and call probability?
  - how about effects of clustering?

Extension to Internet

The Scale-free Internet

http://www.bearcave.com/bookrev/linked/wwwnet.jpg

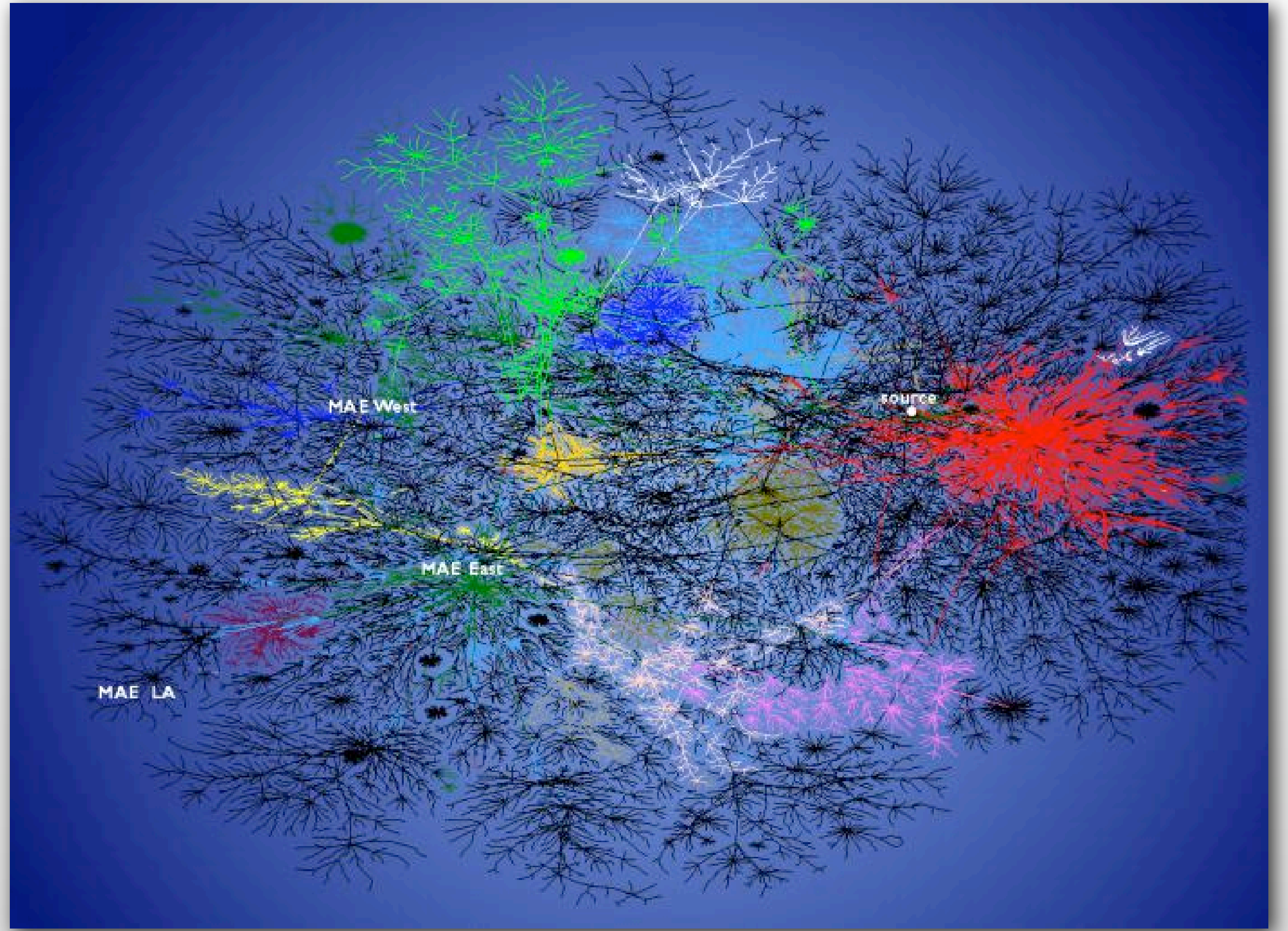
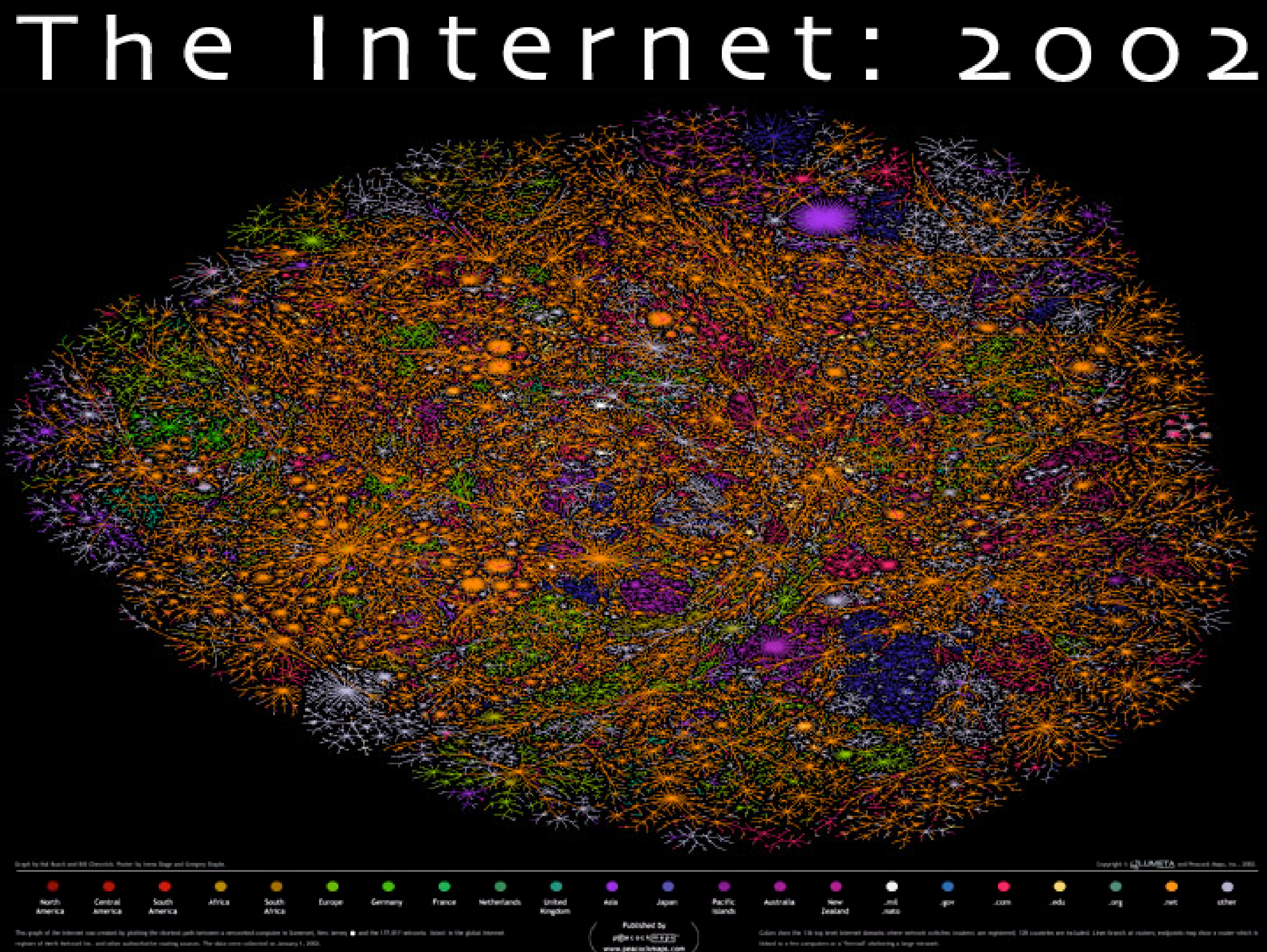


Diagram of Internet connections, showing the major Metropolitan Area Exchanges (MAE), by K.C. Claffy, republished on Albert-László Barabási's Self-Organized Networks Gallery web page.

Peacock Maps 2002



Basic model

with multiple power laws

- The Internet is scale-free following a power-law degree distribution

$$\Pr(k_i) \sim k_i^{-\gamma}$$

- Each node in the network has a buffer size $B_i = [1 + \delta(k_i)^{\gamma_3}]$
- At each time step, new packets are generated in node *i* with probability $\lambda_i = \alpha(k_i)^{\gamma_1}$
- The packet destination is randomly chosen, and the packet is placed at the end of the buffer of the node.
- If the buffer is full, the packet is dropped.
- At each step, node i sends first μ_i packets one step closer to the destination.

$$\mu_i = [1 + \beta(k_i)^{\gamma_2}]$$

- The packet is removed when it reaches the destination.

Simulations

Parameters

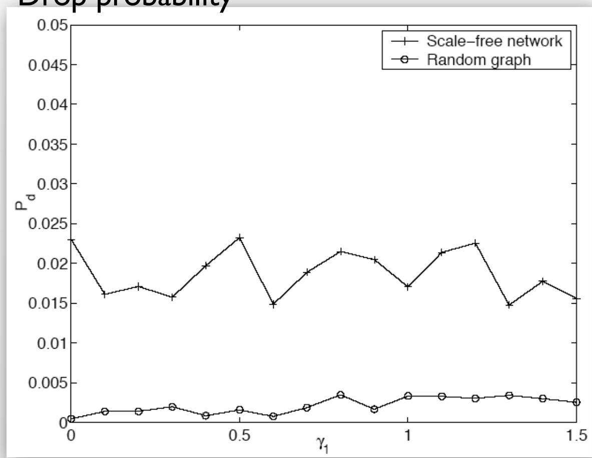
- Network: $N = 10000$ and $\gamma = 2.4$
- Packet generation: $\sum_{i=1}^{N} \lambda_i \approx 100, \sum_{i=1}^{N} \mu_i \approx 2100$ and $\sum_{i=1}^{N} B_i \approx 10000$

  - γ_1 : more new packet generation in higher degree nodes
  - γ_2 : more packets sent in higher degree nodes
  - γ_3 : bigger buffer size in higher degree nodes

Note: If $\gamma_1 = 1$, for example, the degree (connection) power law and the packet generation power law is directly proportional. If $\gamma_1 > 1$, the packet generation power law has a higher exponent, i.e., a more severe disparity among nodes.

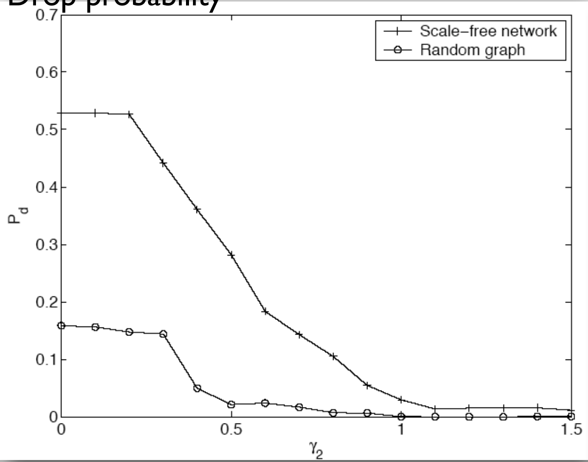
Effects of parameters

Drop probability



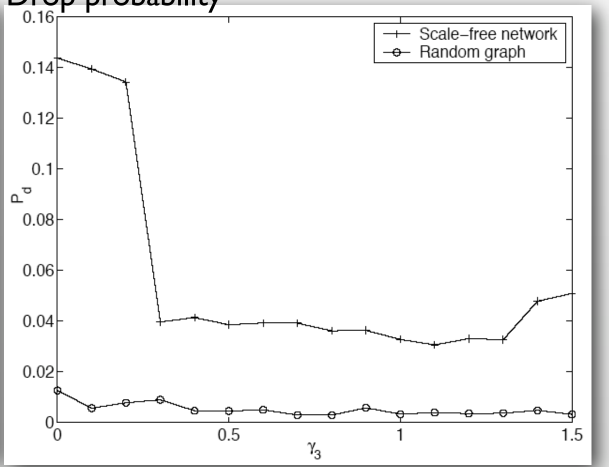
Performance relatively unaffected by making more or less hubs in the network

Drop probability



Assigning more delivery capacity to hubs benefits the traffic performance.

Drop probability



Assigning appropriate buffer benefits the performance, but excess buffer will not help.

In general, scale-free network model reflects a higher drop probability than random graph model.

Table 1: Values of λ_i for different k_i and γ_1. γ_1 increases from 0 to 1.5. A scale-free network is considered, where $N = 1000$ and $\gamma \approx 2.4$.

	$\gamma_1 = 0$	$\gamma_1 = 0.1$	$\gamma_1 = 0.5$	$\gamma_1 = 1$	$\gamma_1 = 1.5$
$k_i = 30$	0.100	0.132	0.362	1	1
$k_i = 10$	0.100	0.118	0.209	0.350	0.443
$k_i = 1$	0.100	0.094	0.066	0.035	0.014

Further Thoughts

Key messages

- Complex network theory offers a totally new vista to the analysis of traffic dynamics in communication networks.
- The key issue is the user network, which is ***not uniform or random***, but is ***scale-free.***
- The traffic volume can be severely limited by this very property of user connection.
  - Philosophical question: ***Why, then, would we naturally evolve to this scale-free connection style in our society?***

More work is needed

- Study of
  - the power-law properties of the parameters and their dependence.
  - effects of other important network parameters, e.g., clustering coefficient.
  - directed graph and weighted graph models, e.g.,
    - incoming and outcoming calls are not necessarily the same in terms of rate of generation,
    - some links are more frequent in occupying traffic.
- Verification of simulation data with real industrial data.

Last philosophical question

- Why would nature like to be connected in a scale-free manner? Optimization of performance!
    - In most places, a small number of people naturally do most of the work (power law distribution). Unequal workload is the natural choice!
    - In artificial systems, optimization often leads to unequal workload in the system. (In LDPC coding, we found that optimization of the decoder actually gives very similar results as scale-free decoder.)
- ***Can this be an inspiration for optimizing artificial systems?***

References

Journal papers:

Y. Xia, C.K. Tse, W.M. Tam, F.C.M. Lau and M. Small, "Scale-free user-network approach to telephone network traffic analysis," *Physical Review E,* vol. 72, 026116-1-7, August 2005.

Y. Xia, C.K. Tse, W.M. Tam, F.C.M. Lau and M. Small, "Analysis of telephone network traffic based on a complex user network," *Physica A,* vol. 368, Issue 2, pp. 583-594, August 2006.

Y. Xia, C.K. Tse, F.C.M. Lau, W.M. Tam and X. Shan, "Telephone traffic analysis based on scale-free user network and scale-free load distribution," *Dynamics of Continuous, Discrete and Impulsive Systems, Series B: Applications and Algorithms,* to appear.

W.M. Tam, F.C.M. Lau, C.K. Tse, Y. Xia and X. Shan, "Effect of clustering in a complex user network on the telephone traffic," *Physica A,* to appear.

Conference papers:

Y. Xia, C.K. Tse, F.C.M. Lau, W.M. Tam and X. Shan, "Traffic congestion analysis in complex networks," *IEEE International Symposium on Circuits and Systems*, (ISCAS'06), Kos Island, Greece, pp. 2625-2628, May 2006.

Y. Xia, C.K. Tse, F.C.M. Lau, W.M. Tam and M. Small, "Telephone traffic analysis based on a scale-free user network," *International Symposium on Nonlinear Theory and Its Applications* (NOLTA'05), Bruges, Belgium, pp. 110-113, October 2005.

W.M. Tam, F.C.M. Lau, Y. Xia, C.K. Tse and M. Small, "Traffic analysis of a mobile communication system based on a scale-free user network," *International Symposium on Nonlinear Theory and Its Applications* (NOLTA'05), Bruges, Belgium, pp. 130-133, October 2005.

Applied Nonlinear Circuits and Systems Group
Department of Electronic and Information Engineering
Hong Kong Polytechnic University
http://chaos.eie.polyu.edu.hk