Weierstraß-Institut für Angewandte Analysis und Stochastik

WIAS Kolloquium

# X. YAO

# Random Graph and Complex Networks

Leibniz Gemeinschaft

## The Historical Overview

▷ Erdös-Rényi Models for Random Graph

– Probabilistic methods

– phase transition and critical phenomena in random graphs

▷ Small World Phenomena

– six-degree separation of the “world”.(S. Milgram 1967 & F. Karinthy 1929)

– Watts-Strogatz model

– The case of Internet: 19 degrees.(Albert et al. 1999)

▷ Scale-free Networks

– The power-law degree distribution of the real world networks

– Barabasi-Albert Model, the generalized models...

▷ Complex networks

# Birth of the random graph

▷ **Erdös-Rènyi model**
Vertex set: $V = \{1, 2, \cdots, n\}$;

– **The binomial random graph** $\mathcal{G}_{n,p}$

∘ The graph space $\Omega$ = { the graphs with the vertex set $V$ };

∘ The $\sigma$-algebra on $\Omega$: $\mathcal{F}$ = { the subsets of $\Omega$ };

∘ The probability measure $\mathbb{P}$

$$\mathbb{P}(G) = p^{e_G}(1-p)^{\binom{n}{2}-e_G}, \ G \in \Omega. \tag{1}$$

– **The uniform random graph** $\mathcal{G}_{n,M}$

∘ The graph space $\Omega$ = { the graphs with the vertex set $V$ and with $M$ edges };

∘ The $\sigma$-algebra on $\Omega$: $\mathcal{F}$ = { the subsets of $\Omega$ };

∘ The probability measure $\mathbb{P}$

$$\mathbb{P}(G) = \binom{\binom{n}{2}}{M}^{-1}, \ G \in \Omega \tag{2}$$

## The Problems of Random Graphs

▷ Subgraphs problem

–The existence in the $\mathcal{G}_{n,p}$(or $\mathcal{G}_{n,M}$) of at least one copy of a given graph $G$.

▷ Mathchings

–matching: A matching on a graph is a set of edges such that no two of them share a vertex in common

–perfect matching: the matching covers every vertex of the graph.

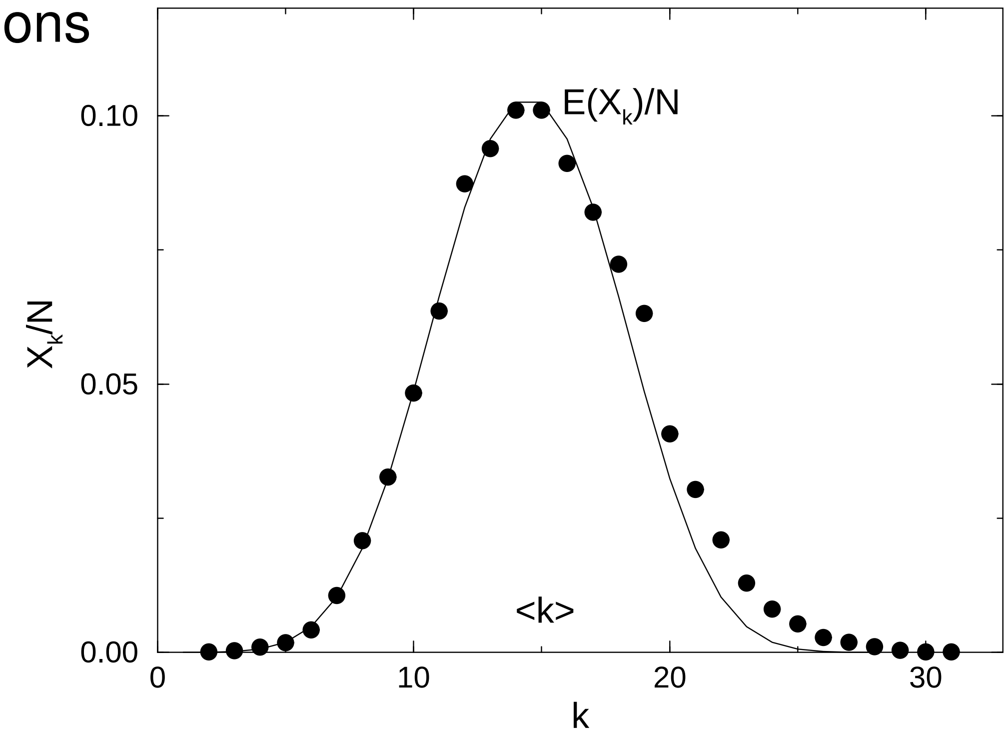
–The existence of perfect matchings in random graph

▷ The Phase transition and the critical behavior

–connectivity problem, largest component size, ...

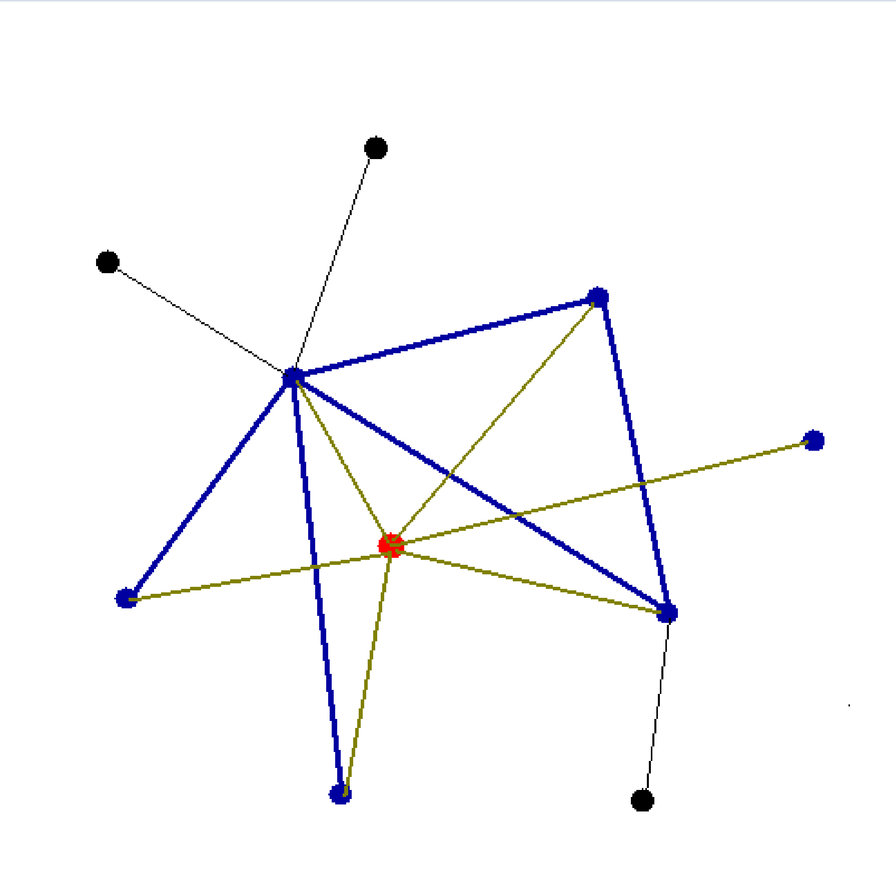
## Some important properties

▷ Degree distributions



## Some important properties

▷ Clustering degree $\Delta$ and Clustering Coefficients $C$



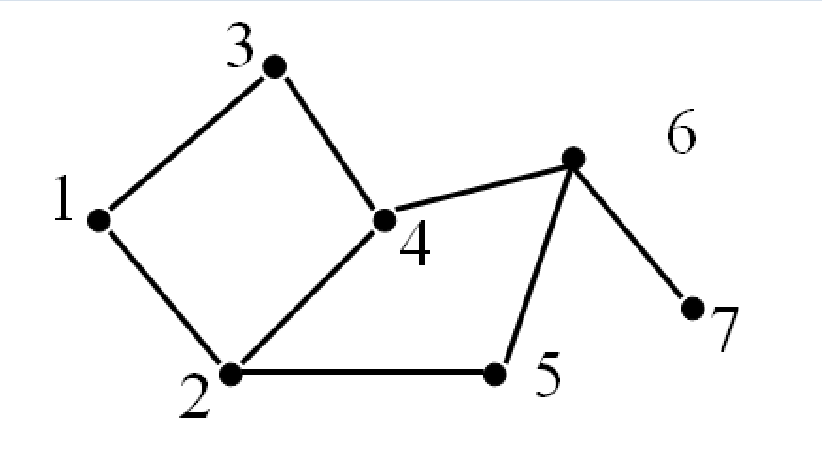
$\Delta_v =$ the number of links between the neighbors of vertex $v$,

$C_v = \Delta_v / \binom{k_v}{2}$,

$C = C_v / \sum_{v \in V} C_v$

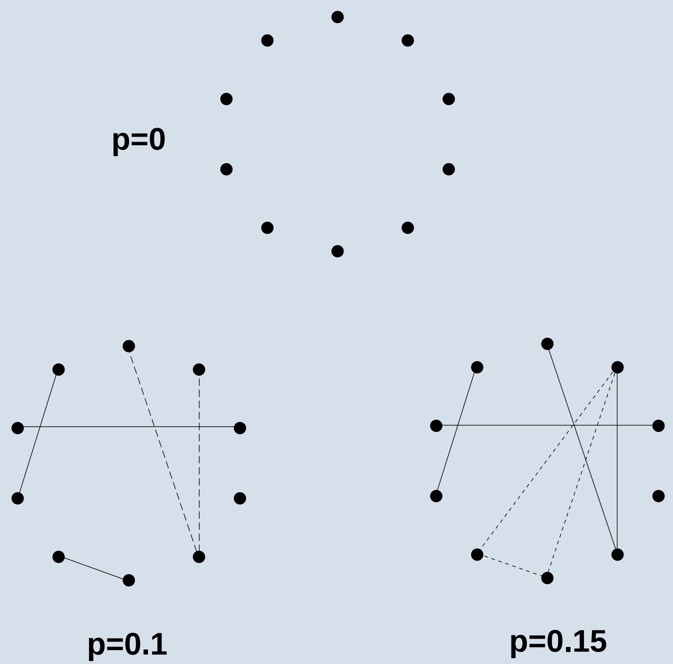
# Some important properties

▷ Mean shortest path $\ell$




$D_{15}, 1 \to 2 \to 5$,
$D_{17}, 1 \to 2 \to 5 \to 6 \to 7$.
To describe the “diameter” of the network, we use the mean value of the shortest paths between any two vertices.

# The evolution of the random graphs




When the connecting probability $p$ increases from 0 to 1, the corresponding random graph become denser and denser. For many graph property, there exists a threshold to decide the whether the random graph possess such property.

## Probabilistic methods

*In order to prove the existence of an object with a specific property, prove that a randomly chosen object satisfies that property with positive probability.*

**The example: the connectivity problem (Erdös and Rényi(1959))**

▷ For the binomial random graphs $\mathcal{G}_{n,p}$, where $p = \frac{\log n + c + o(1)}{n}$ where $c \in \mathbb{R}$, we have

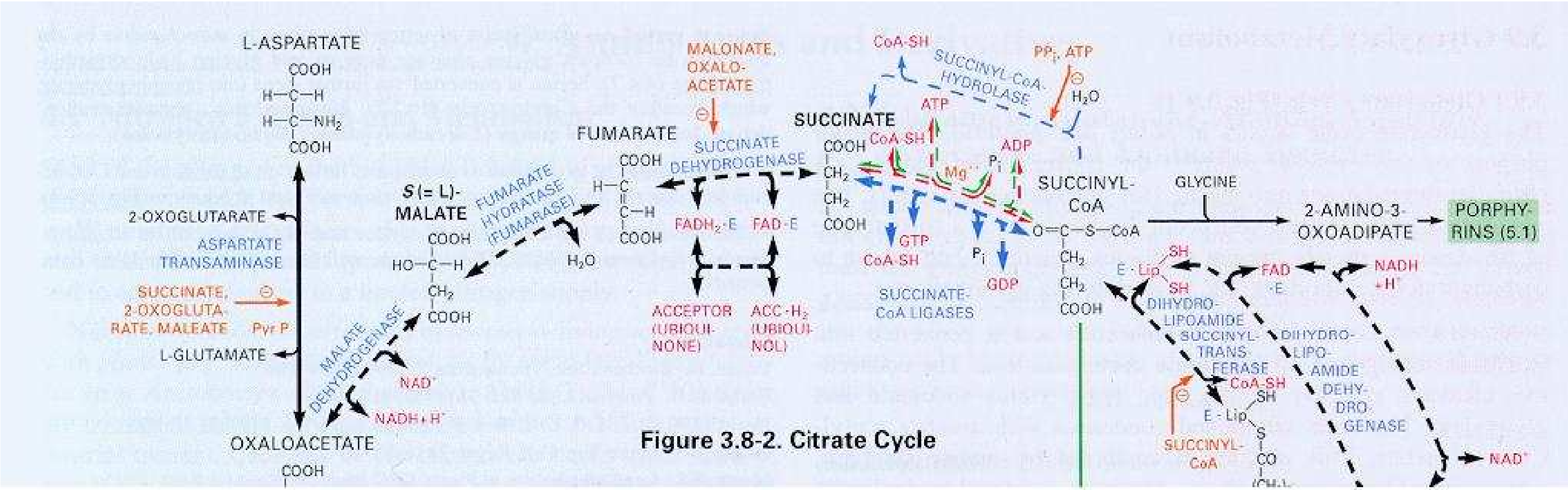
$$\mathbb{P}(\mathcal{G}_{n,p} \text{is connected}) \to \mathrm{e}^{-\mathrm{e}^{-c}} \tag{3}$$

and we have the critical point $p^* = \frac{\log n}{n}$

## A generalized case: Random Intersection graphs

▷ Example 1: metabolic networks

A chemical reaction need several kinds of molecules, and one kind of molecule may involved with many chemical reactions;




(Barabási A.(2005))

# A generalized case: Random Intersection graphs

▷ Example 2: Scientific Collaboration Networks

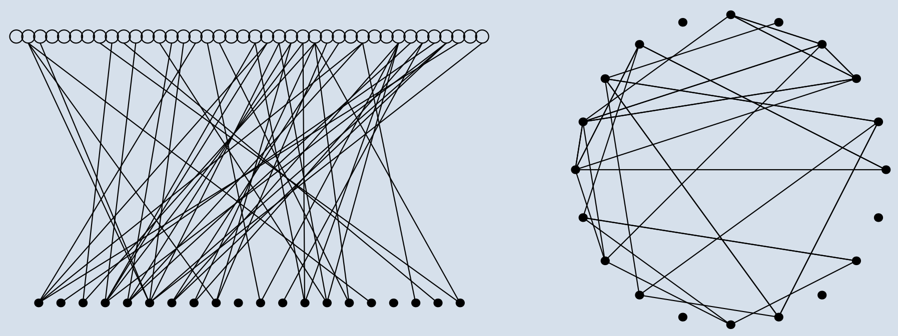
An scientific article may have several authors, and an author may cooperated with many others in other publications.

If the authors are looked as vertices and connected two authors if they have published some same articles together, we also induce a graph.

# A generalized case: Random Intersection graphs

▷ The model of random intersection graph $\mathcal{G}_{N,M,p}$. (Siger-Cohen, K., 1995)

$V = \{v_1, v_2, \cdots, v_N\}$, $W = \{w_1, w_2, \cdots, w_M\}$, $p_{vw} \in [0, 1]$.



The relations among the parameters $M, N, p$ decide almost all the properties of the $\mathcal{G}_{N,M,p}$, let.
$M = N^{\alpha}$, $p = N^{-\beta}$
where
$\alpha \geq 0$ and $\beta \geq 0$

## A generalized case: Random Intersection graphs

▷ Some results of random intersection graphs

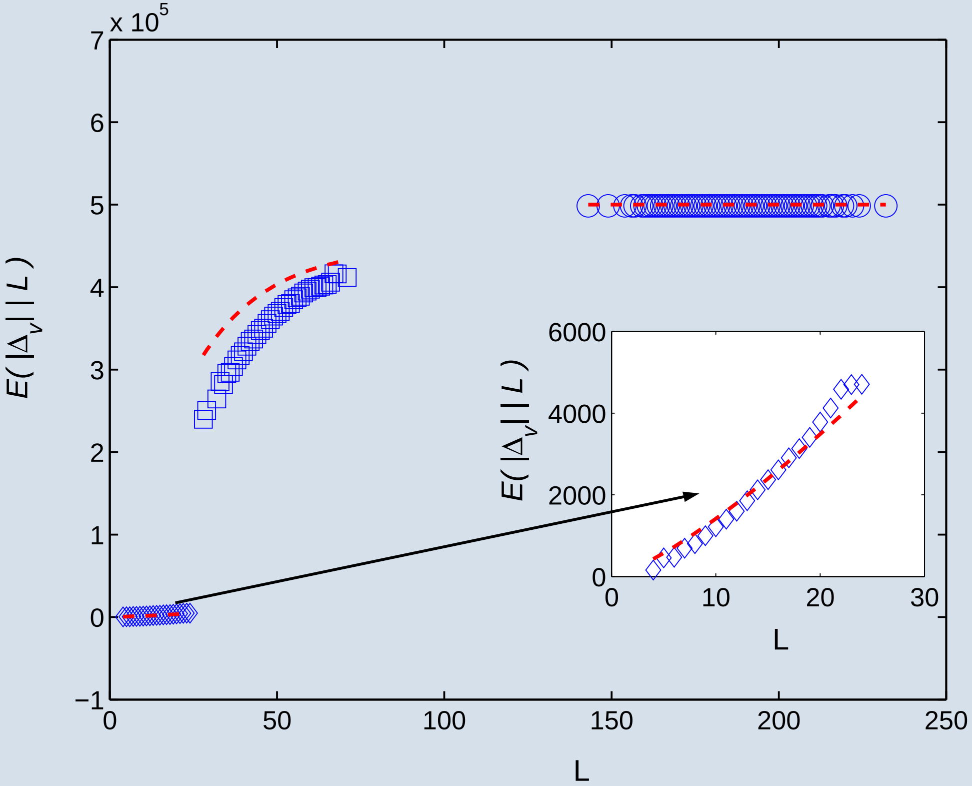
– subgraph problem(Singer-Cohen, K.(1995,1999), Fill(2000))

– Degree distribution(Stark, D. (2005))

– Clustering Coefficients(Yao.X, etc. (2006))

# The critical behavior of the clustering property in $\mathcal{G}_{N,M,p}$

▷ The localized clustering property (1): clustering degree $\Delta_v$



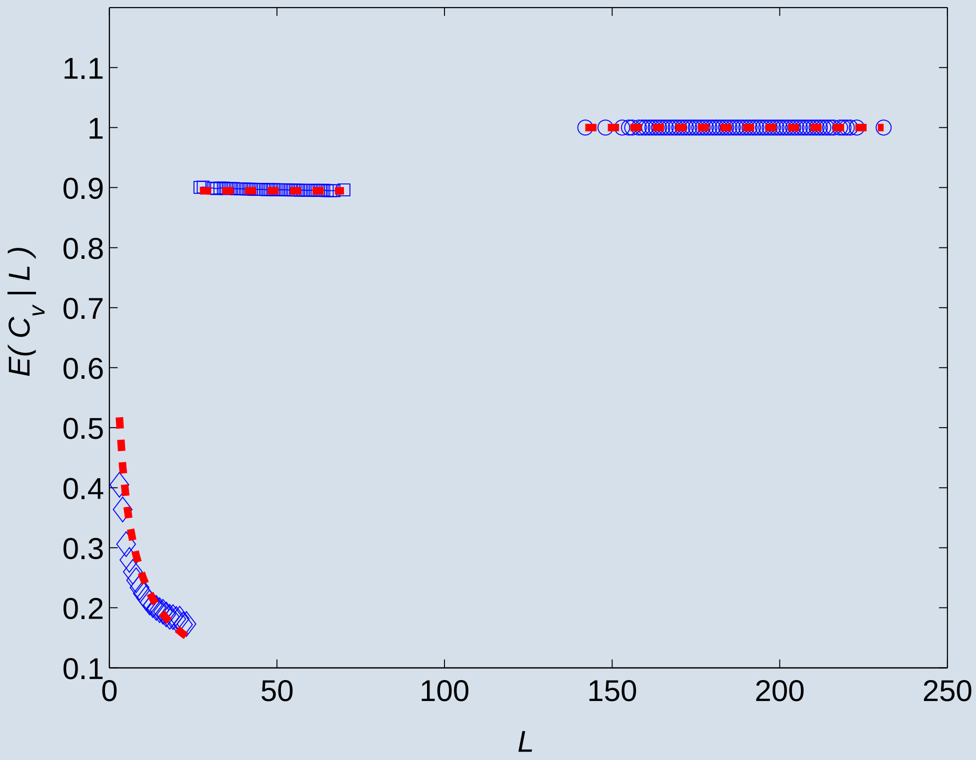

– $L_v$: the number of links between the vertex $v$ and $W$ set.

– $\Delta_v$: the number of links between the neighbors of vertex $v$ in $\mathcal{G}_{N,M,p}$.

$$\mathbb{E}\left[\Delta_v \middle| L\right] \sim \begin{cases} \dfrac{c^2 L}{2} N^{2-2\beta}, & \beta \leq \alpha < 2\beta \\ \dfrac{(1-e^{-Lp})^3}{2} N^2, & \alpha = 2\beta \\ \dfrac{N^2}{2}, & \alpha > 2\beta \end{cases} \tag{4}$$

# The critical behavior of the clustering property in $\mathcal{G}_{N,M,p}$

▷ The localized clustering property (2): Clustering Coefficients $C_v$.



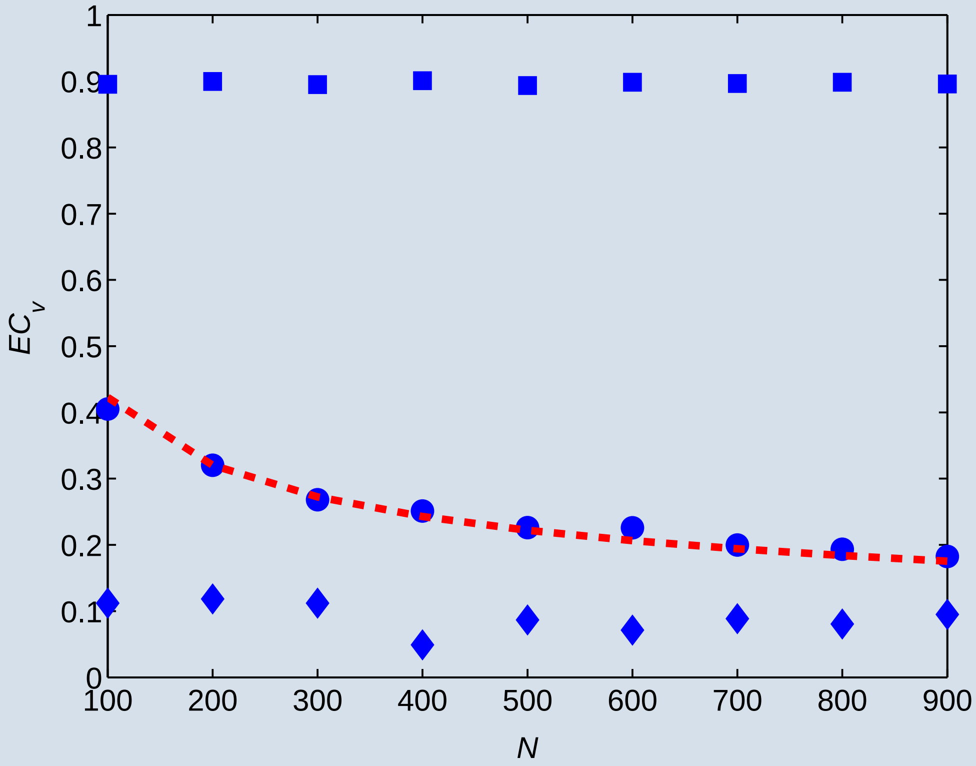

– $L_v$: the number of links between the vertex $v$ and $W$ set.

– $C_v$: the Clustering Coefficient of vertex $v$.

$$\mathbb{E}\left[C_v \middle| L\right] \sim \begin{cases} \frac{1}{L}, & \beta \leq \alpha < 2\beta \\ 1 - e^{-\frac{cL}{N^\beta}}, & \alpha = 2\beta \\ 1, & \alpha > 2\beta \end{cases} \tag{5}$$

# The critical behavior of the clustering property in $\mathcal{G}_{N,M,p}$

▷ The Global clustering property: Clustering Coefficients $C$.



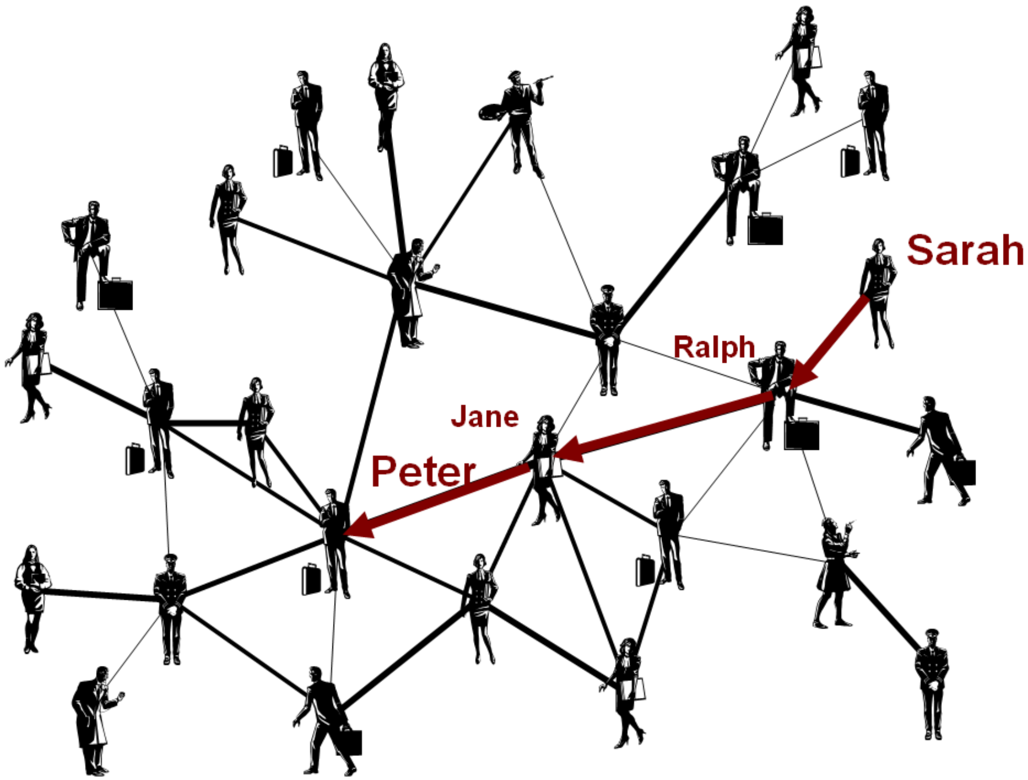

– $L_v$: the number of links between the vertex $v$ and $W$ set.

– $C_v$: the Clustering Coefficient of vertex $v$.

$$\mathbb{E}\left[C_v\right] \sim \begin{cases} c_0, & \alpha = \beta \\ \frac{1}{c} N^{-(\alpha-\beta)}, & \beta < \alpha < 2\beta \\ 1 - e^{-c^2}, & \alpha = 2\beta \\ 1, & \alpha > 2\beta \end{cases} \tag{6}$$

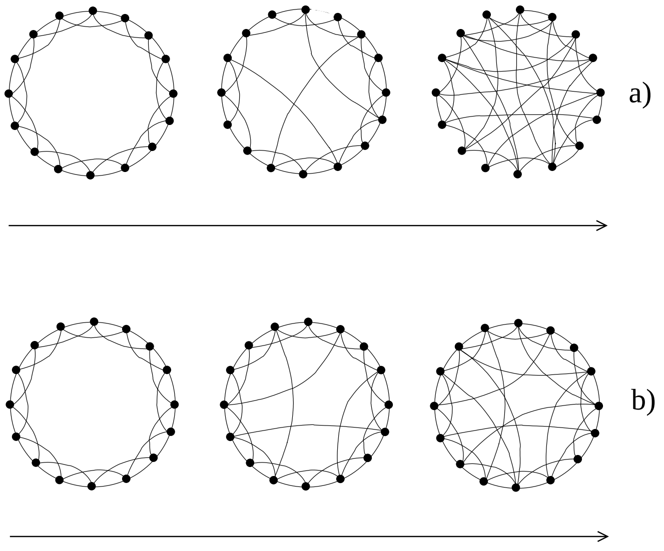
# Small world phenomena

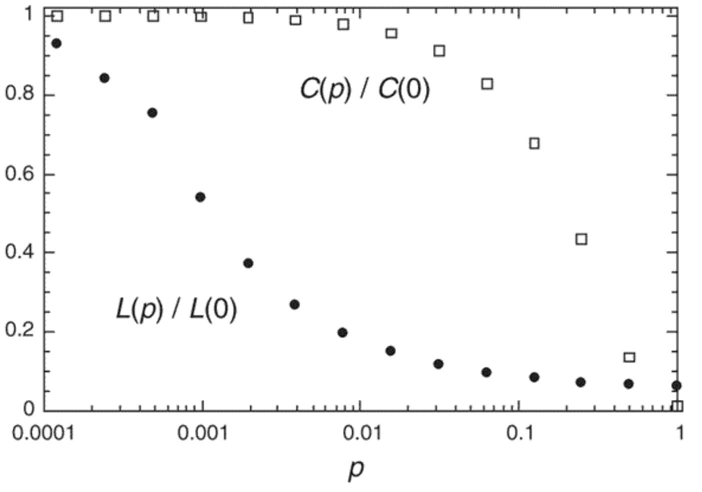
▷ The six degree separation of the world.



# Small world phenomena

▷ Watts-Strogatz model for small world phenomena

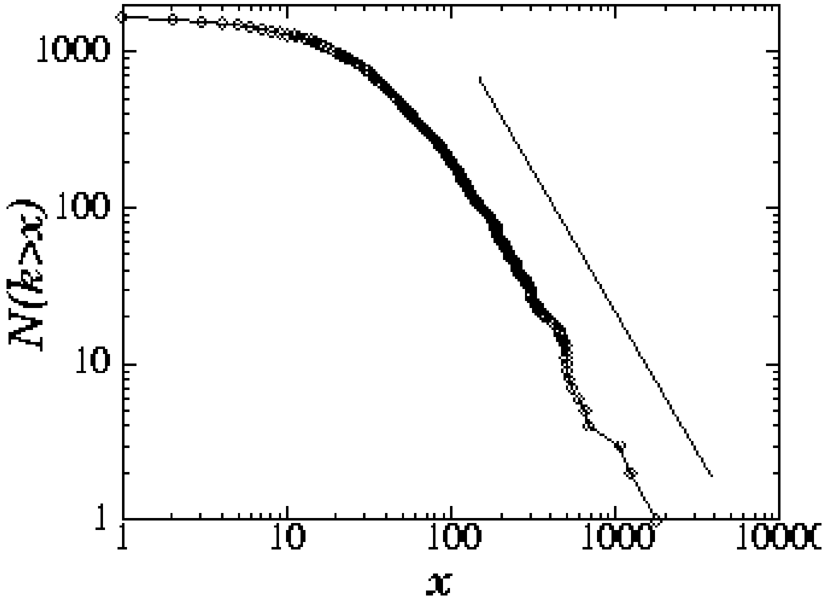



(Watts & Strogatz, 1998)

# Scale-free Networks

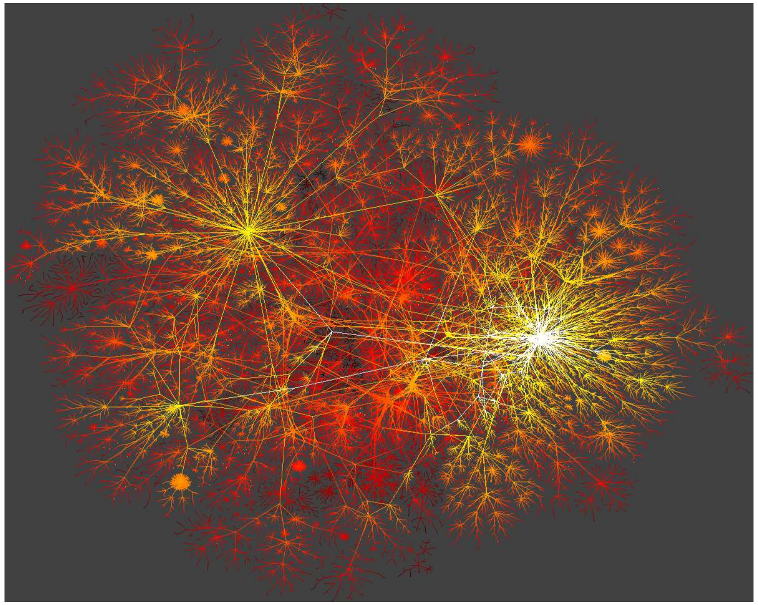
▷ Example 1: Citation Networks

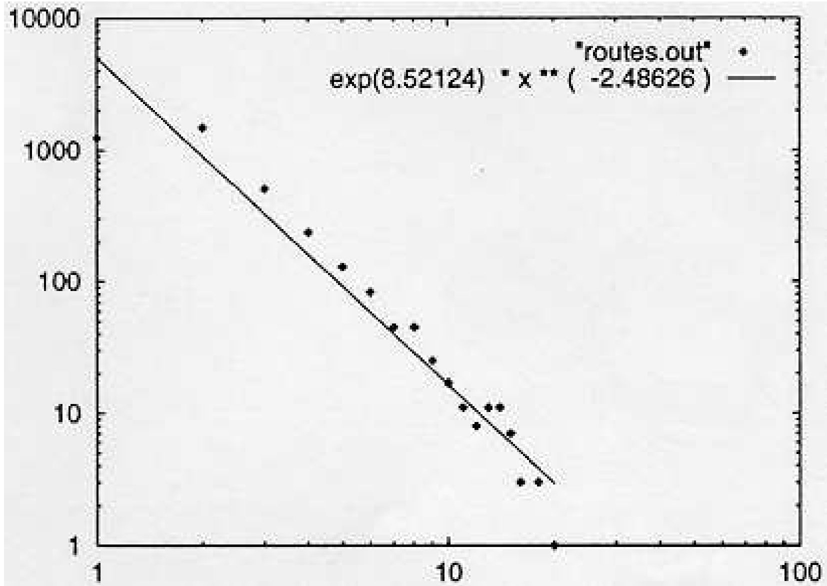



(Barabási(1999))

# Scale-free Networks

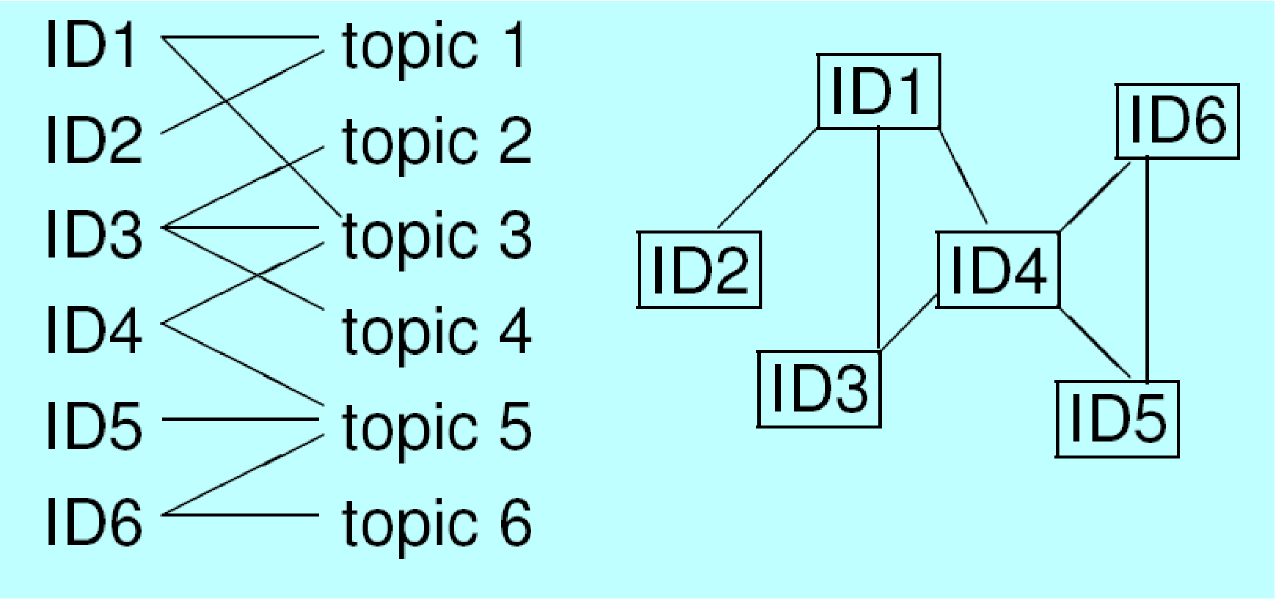
▷ Example 2: Internet



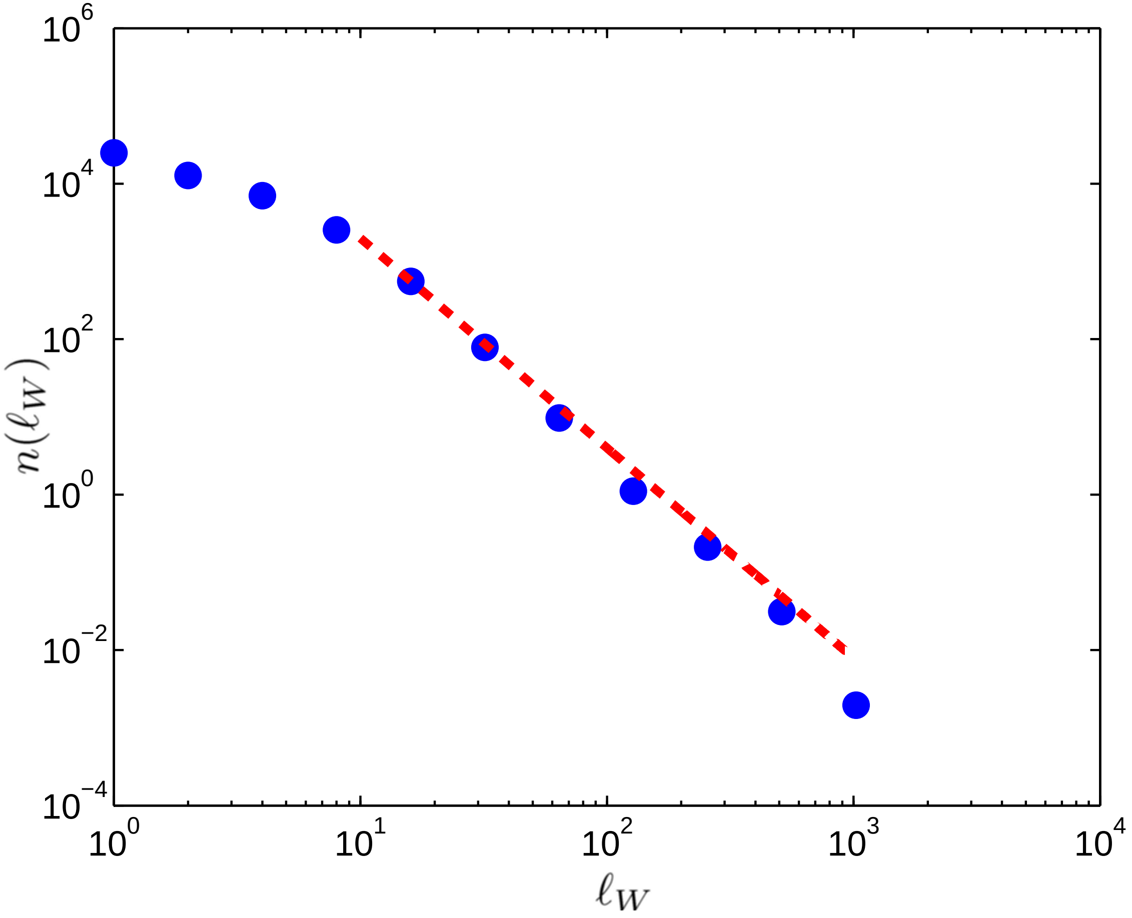


# The real world network: Scale-free Networks

▷ Example 3: Broadcast board System(BBS): the formation of the networks



▷ Example 3: Broadcast board System(BBS): the degree distribution, (Yao,X. etc. (2005))



## Scale-free Networks

▷ The Barabasi-Albert model — linear preferential attachment

– Growth: The network increases with time, by adding a new vertex at each step with $m > 0$ new edges;

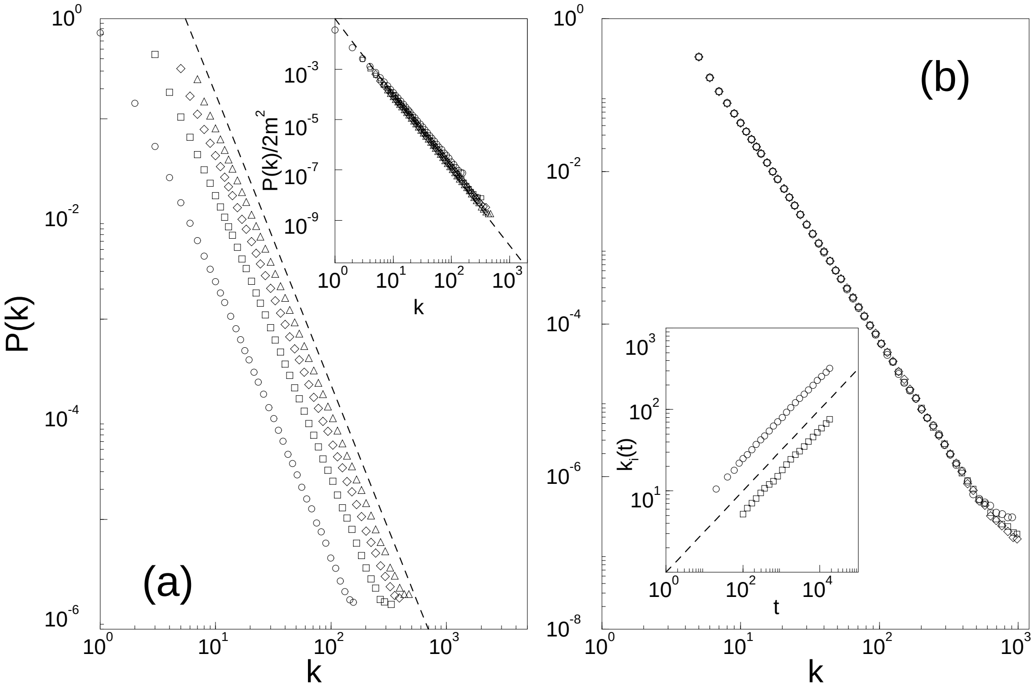
– Preferential attachments: The new added vertices prefer to link to highly connected vertices;

$$\mathbb{P}(v_t \text{ connected with } v_i) = \frac{k_i}{\sum_j k_j} \tag{7}$$

## Scale-free Networks

▷ The Barabasi-Albert model — linear preferential attachment

$-\mathbb{P}(k) \sim k^{-3}$, (Albert & Barabási (1999))



# Scale-free Networks

▷ The Barabasi-Albert model — The generalized case

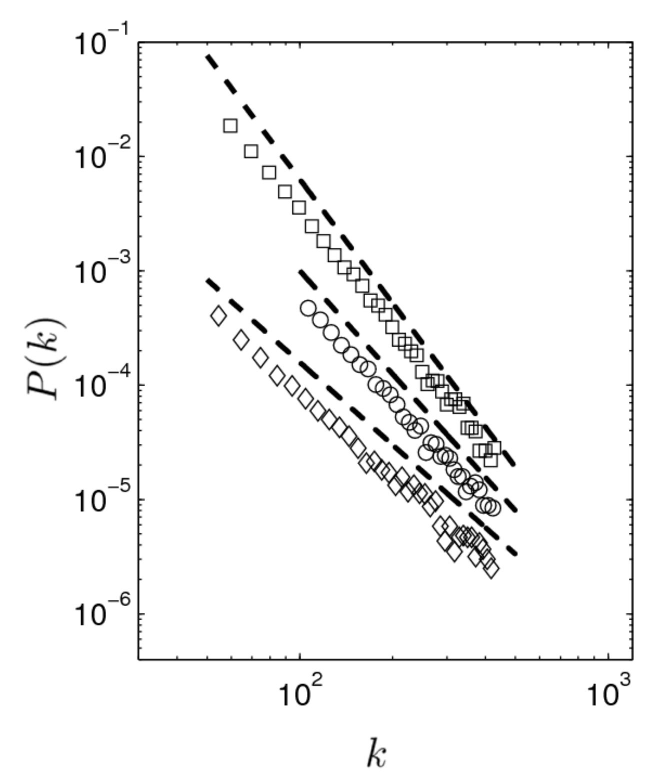
– Growth: The network increases with time, by adding a new vertex with $m > 0$ new edges;

– Preferential attachments: The new added vertices prefer to link to highly connected vertices, but with an offset $k^*$;

$$\mathbb{P}(v_t \text{ connected with } v_i) = \frac{k_i + k^*}{\sum_j (k_j + k^*)} \tag{8}$$

# Scale-free Networks

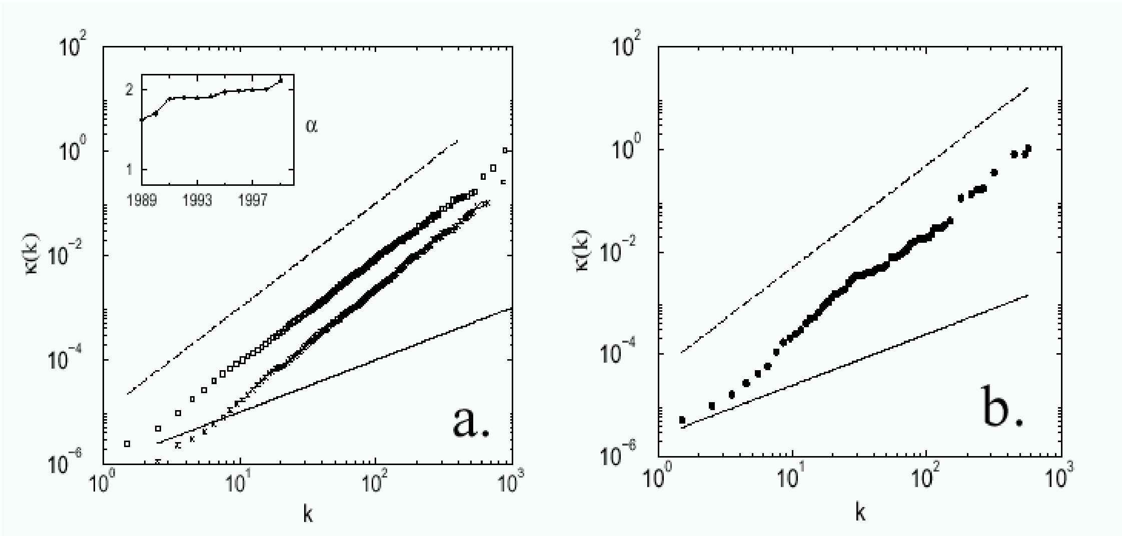
▷ The Barabasi-Albert model — The generalized case




$$P(k) \sim k^{3+\frac{k^*}{m}}$$

## Scale-free Networks

▷ Is the “preferential attachment” true in the real world?



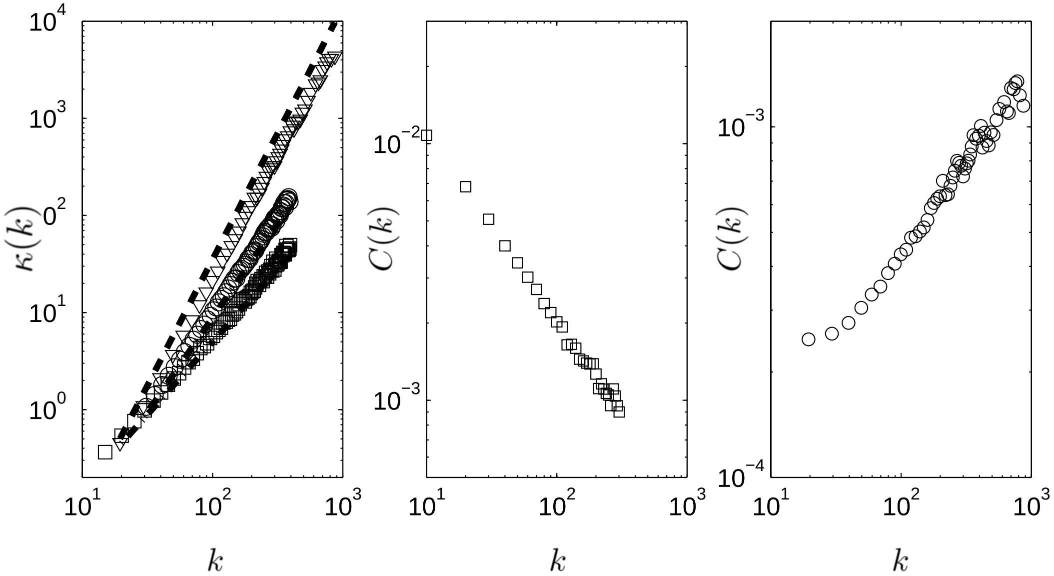

(Albert & Barabasi, (2001))

# What is complex networks

- ▷ Power-law degree distribution;
- ▷ High Clustering property;
- ▷ Short diameter.
  $D \sim \log N$

# The clustering property of Barabási-Albert model(The generalized case)

▷ The clustering degree of Generalzed barabasi model(Yao,X. etc. (2005))




$\kappa(k) \sim \frac{C_1}{N^{1-\beta}} k^{1/\beta}$, where $\beta = \frac{m}{2m+k^*}$.

# The current hot points and future problems

- ▷ Constructing more general and rigorous models;
- ▷ The motif problem—The subgraph problems in complex networks;
- ▷ The community detection—classification problems.
- ▷ The dynamical model on the complex networks, e.g., random walks