Modeling of Complex Networks

Lecture 1: Basic Concepts

(Semester A, Fall 2019)

Instructor: Aaron, Haijun Zhang, Office: Rm 1601, Bld L

Info.

Tutor: YU Qiong, REN Lang

Lab: Rm 701, Bld G

Email: 1579224995@qq.com, leonren1997@gmail.com

PPT, homework and project assignment can get from our QQ discussion group:

683010511

Public Email:

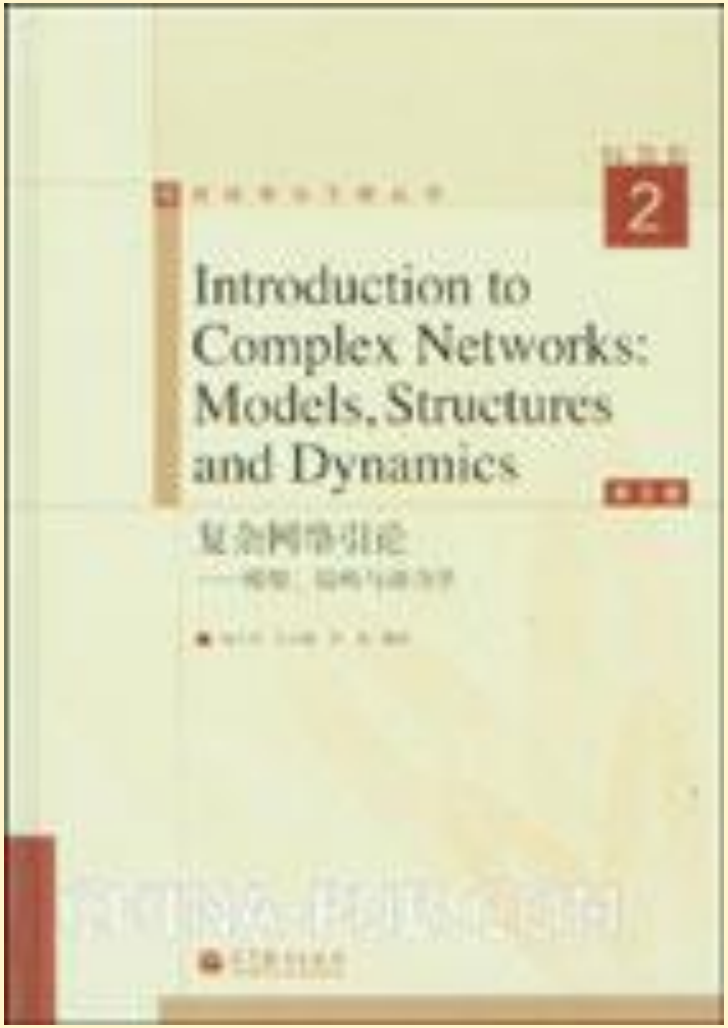
Username: complexnetwork19@sina.cn

2019复杂网络建模

扫一扫二维码，加入群聊。

Textbook for This Course

Textbook:
Introduction to Complex Networks: Models, Structures and Dynamics by G.R. Chen, X.F. Wang and X. Li, Higher Education Press, Beijing, May 2012.



References (recommended but not required) :
Introduction to Complex Dynamical networks (in Chinese复杂网络理论及其应用) by X. F. Wang, X. Li and G. R. Chen, Tsinghua University Press, Beijing, 2006.

Other Recommended Books

- Networks: An Introduction by M. E. J. Newman, Oxford University Press, New York, 2010.
- Lecture Notes on Graph Theory by Tero Harju, http://users.utu.fi/harju/graphtheory/graphtheory.pdf, 2007.
- Networks, Crowds, and Markets: Reasoning about a Highly Connected World by David Easley and Jon Kleinberg, 2010.
- Top-notch journal and conference papers.

Course Coverage

- Basic Concepts
- Introduction to Graph Theory
- Network Topologies - Basic Models and Properties
- Internet – Topology and Modeling
- Spreading Dynamics
- Cascading Reactions
- Synchronization
- Human Opinion Dynamics

Assessment Pattern and Requirements

- Pre-requisite : Nil but one programming language (for the project implementation).

- Lecture, Homework and Project Development

- Final grade will be determined by:
  - 20% homework (4 or 5 assignments)
  - 30% project
  - 50% Final Exam
- Examination duration 2 hours, within two weeks after finishing the lecture (in the middle of this semester).

Intended Learning Outcomes

- Upon completion of this course, you will be able to :
  - apply basic knowledge of graph theory to various engineering problems
  - apply basic concepts to build representative network models
  - analyze the effects of network structures on dynamical behaviors
  - employ the learned techniques to solve some practical problems
  - ……

Lecture 1: Basic Concepts

S8101003Q (Sem A, Fall 2019)

Instructor: Aaron, Haijun Zhang

1.1 Some Background and Motivation

- Between two randomly selected persons in the world, how many friends are there connecting them together?
- When searching from one Web page to another through the World Wide Web (WWW), how many clicks are needed in average?
- How can computer viruses propagate so fast and so widely through the Internet?
- How are people infected by diseases such as AIDS, SARS, and bird flu all over the world?
- How do rumors spread out in human societies?
- How can traffic jams in metropolitan cities be regulated effectively?
- How does electric power blackout emerge from local system failures through the huge power grid?

We are living in a networked world today

- The influence of various complex and dynamical networks is currently pervading all kinds of sciences, ranging from physical to biological, even to social sciences. Its impact on modern engineering and technology is prominent and will be far-reaching.
- On one hand, networks bring us with convenience and benefits, improve our efficiency of work and quality of life, and create tremendous opportunities that we never had before.
- On the other hand, however, networks also generate harms and damages to humans and societies, typically with epidemic spreading, computer virus propagation, and power blackout, to name just a few.

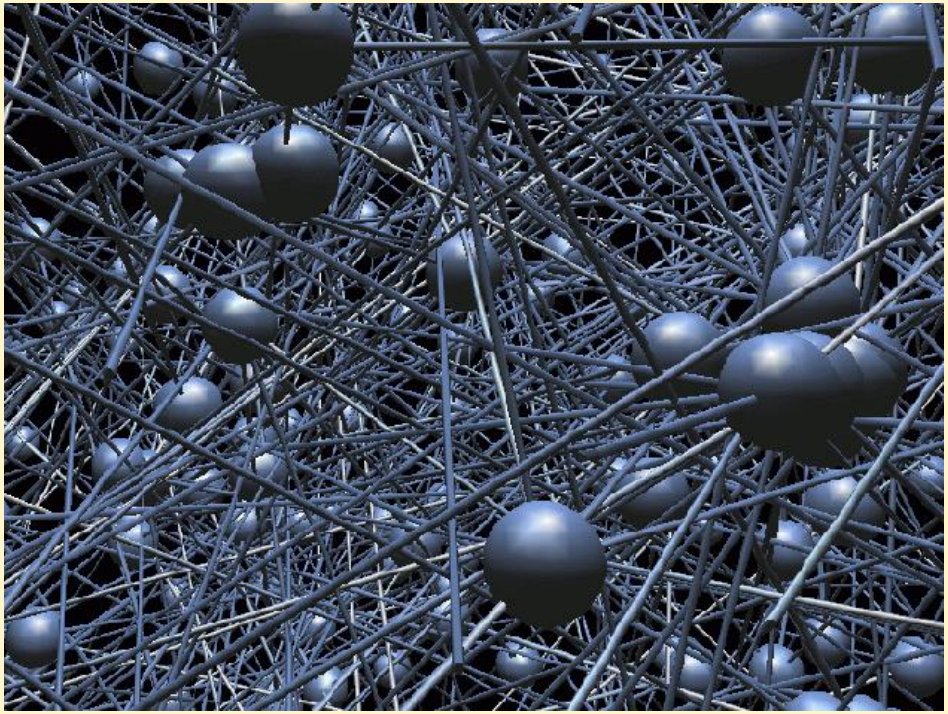
Development of network studies

- For a long time in the history, studies of communication networks, power networks, biological networks, economic networks, social networks, etc., were carried out separately and independently.
- Recently, there were some rethinking of the general theory of complex dynamical networks towards a better understanding of the intrinsic relations, common properties and shared features of all kinds of networks in the real world.
- The new intention of studying the fundamental properties and dynamical behaviors of complex networks, both qualitatively and quantitatively, is important and timely, although very challenging technically.
- The current research along this line has been considered as "network science and engineering", and has indeed become overwhelming.

Network Complexity (I)

Structural complexity:

- A network typically appears complicated in structure, which may even be seemingly messy and disordered.
- The network topology (i.e., structure) may vary in time.
- The connections (i.e. edges) among nodes may be weighted or directed.

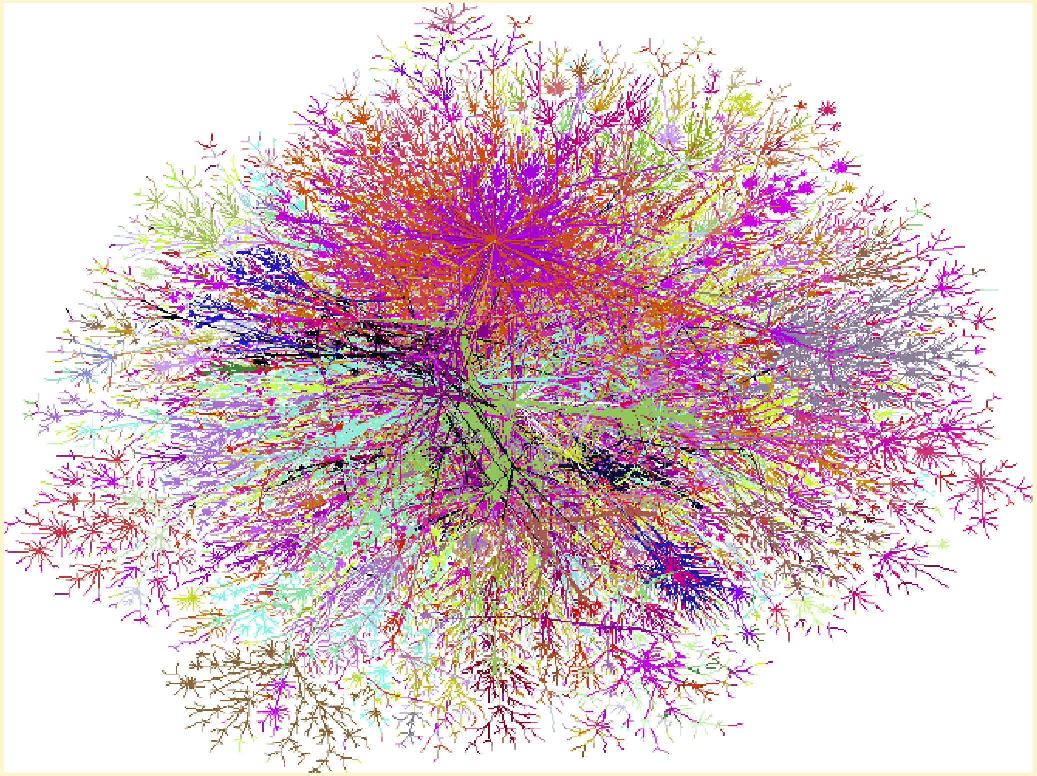


Illustrative graph of a social relationship structure in Canberrra, Australia [A.S. Klovdahl, ANU]

Network Complexity (II)

Node-dynamics complexity:

- A node in the network can be a dynamical system, which may have bifurcating and even chaotic behaviors.
- A network may have different kinds of nodes.

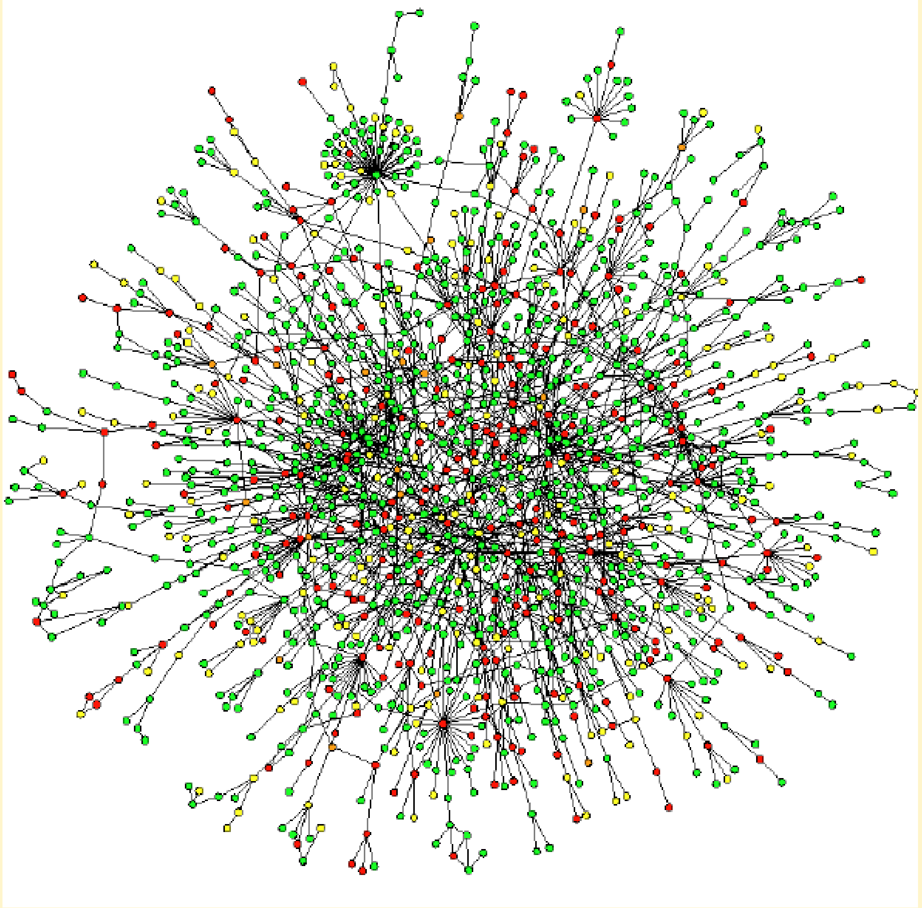


Illustrative graph of some IP addresses in the Internet [W. R. Cheswick, Lumeta Corporation, USA]

Network Complexity (III)

Mutual interactions among various complex factors:

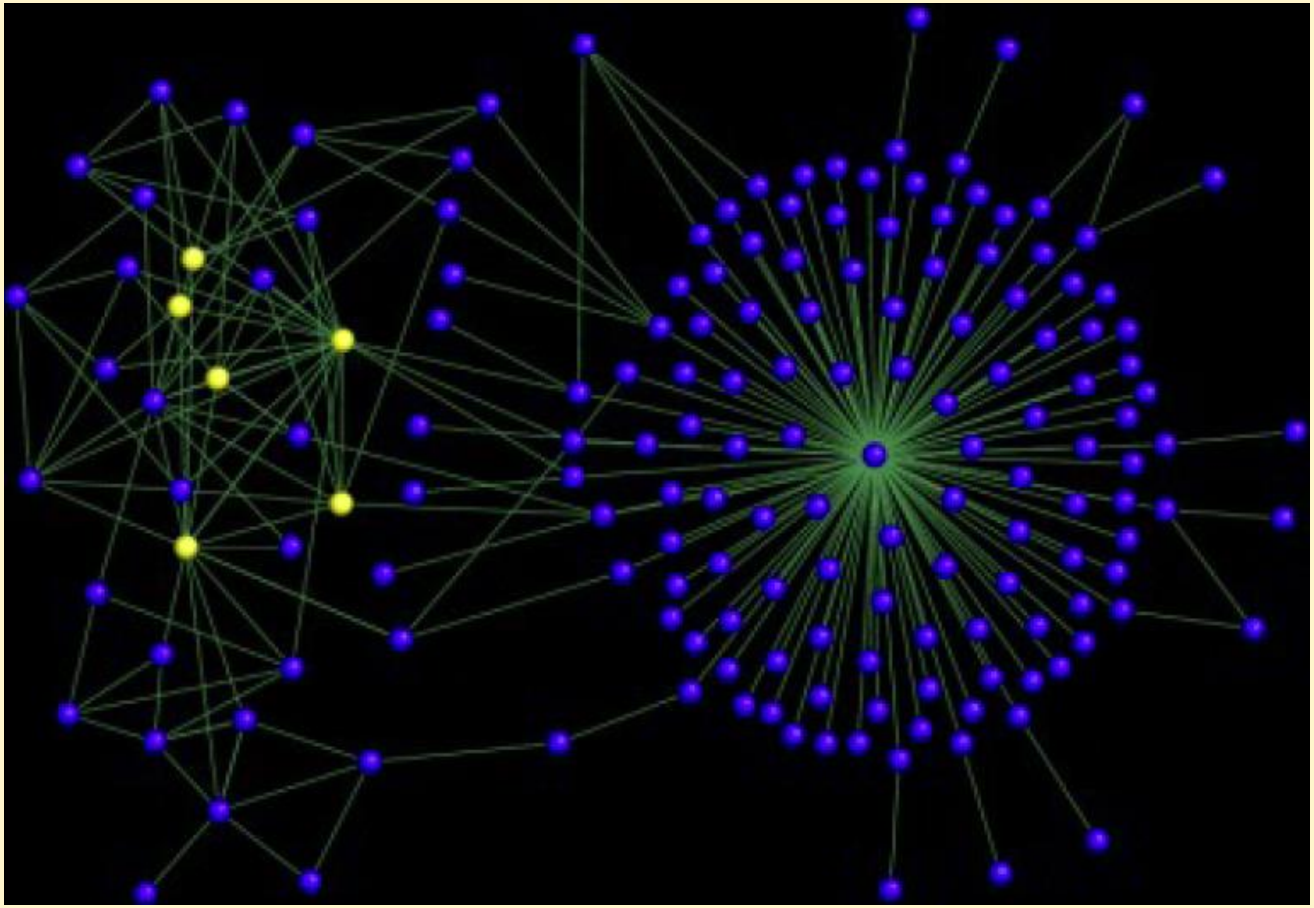
- A real-world network is affected by many internal and external factors.
- The close relations between networks or sub-networks make the already-complicated behaviors of each of them become more complex.



Illustrative graph of interactions among proteins [H. Jeong, Korea Advanced Institute of Sci. and Tech.]

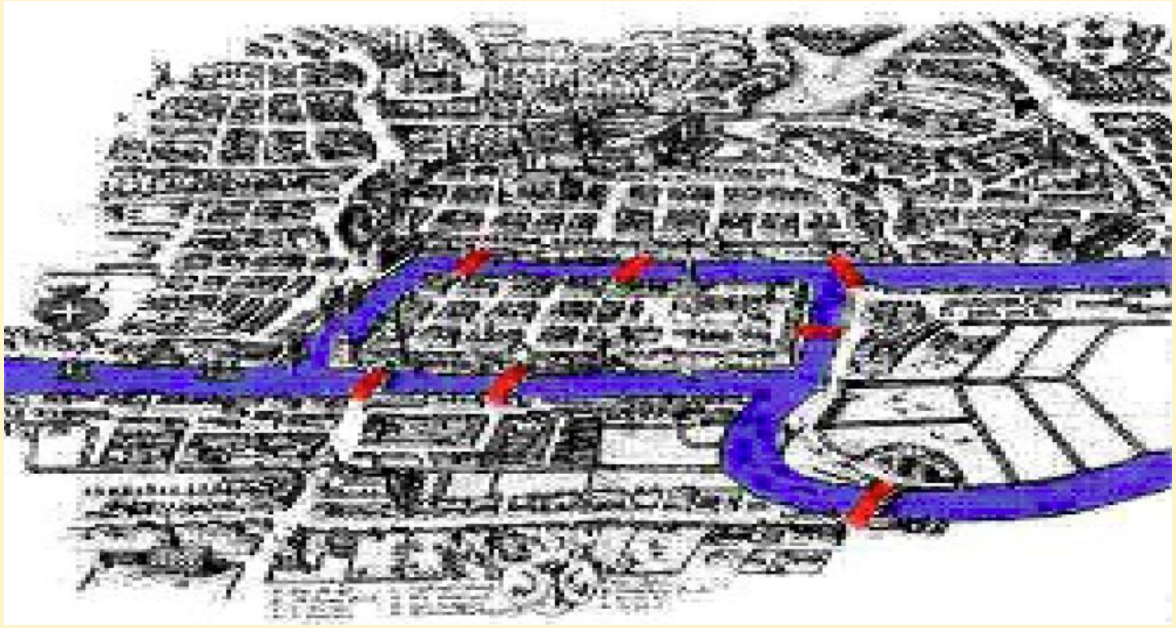
1.2 Brief History of Network Research

Let's start with the Graph Theory …



Seven-bridges problem of Königsberg

- Konigsberg is now in the territory of Russia.
- There are two small islands in the river Pregel passing through the town konigsberg, and there are seven bridges over the river. In the old days, the residents were always amazed as if someone could walk through all the seven bridges and return to the starting point without going over any bridge for more than once.

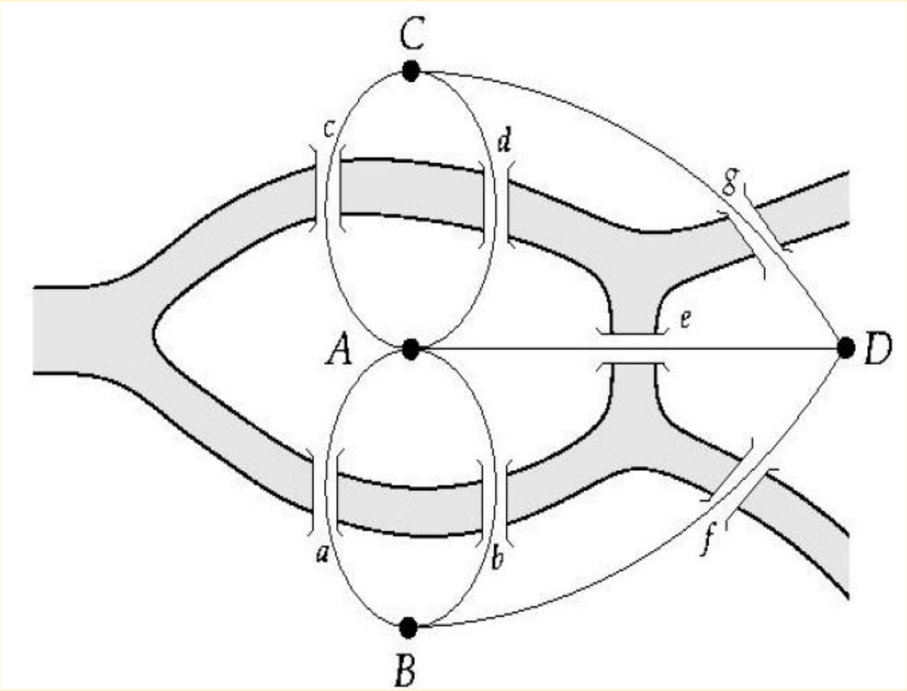


The town Königsberg and the seven bridges in year 1736

Q: Can one walk across all the seven bridges, once and once only on each bridge, and then return to the starting point?

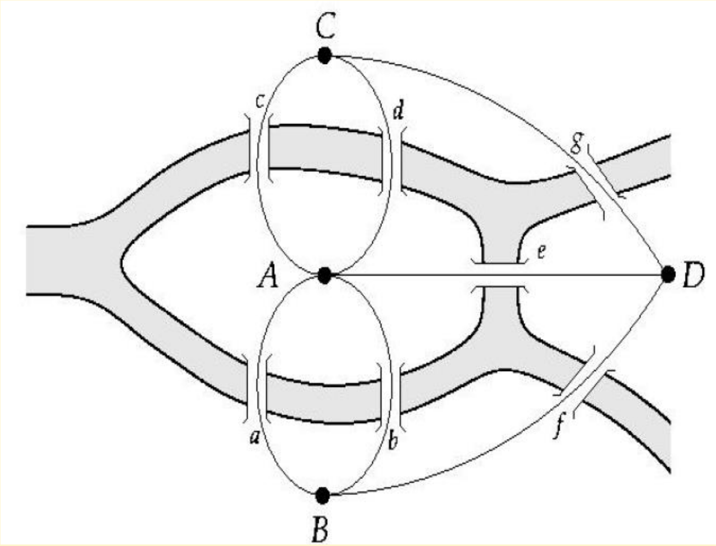
Leonhard Eüler (1707-1783)

"Father of graph theory"



Viewed the problem as a graph with 4 nodes and 7 edges

Solution



Terminology:
A network is a set of nodes (vertices) interconnected by edges (links)

- A path can only has one starting point and one end point -> nodes with an odd number of edges must be either the starting point or the end point -> to have a solution, all points must have an even number of edges
- Thus, the Konigsberg graph has no solution ! – proved by Eüler in 1736 (29 years old)
- By the way, a new bridge was built between B and C in 1875

Eüler's Contributions to Complex Networks

- Eüler initiated the mathematical graph theory
- Graph theory is the basis for the investigation of complex networks
- The construction and evolution of various networks are key to understanding the complex world today

Network Topology

- A network is a set of nodes interconnected via edges

Examples:

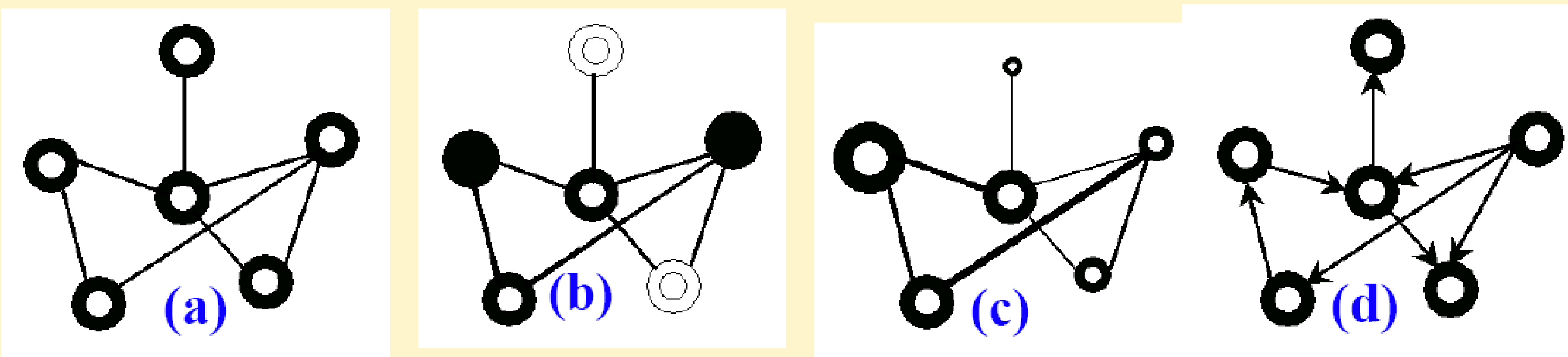
- Internet: Nodes – routers; Edges – optical fibers
- WWW: Nodes – document files; Edges – hyperlinks
- Scientific Citation Network:

  Nodes – papers; Edges – citation
- Social Networks:

  Nodes – individuals; Edges – relations

A Brief Introduction to the Graphs Theory

Some typical graphs:



(a) Undirected and unweighted network with the same type of nodes

(b) Undirected and unweighted network with different types of nodes and links

(c) Undirected but weighted network with weights on both nodes and links

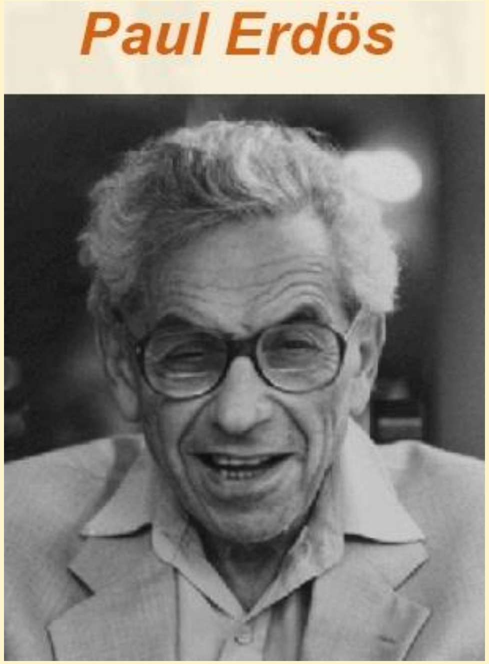
(d) Directed but unweighted network with same type of nodes

(e) … …

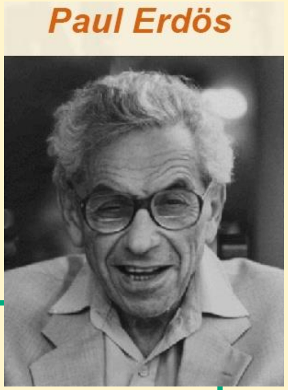
In this course, only undirected and unweighted networks are studied

Random Graph Theory

-- A revolution in the 1960s



- The simplest possible model for the most complex networks
- A single rigorous framework: *Random Graph Theory*

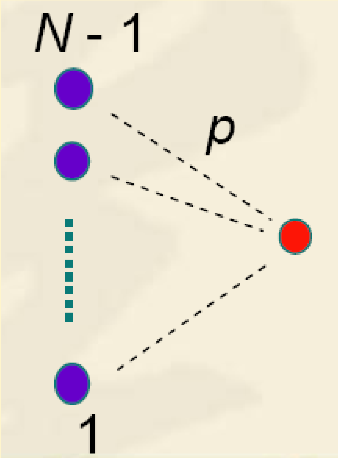


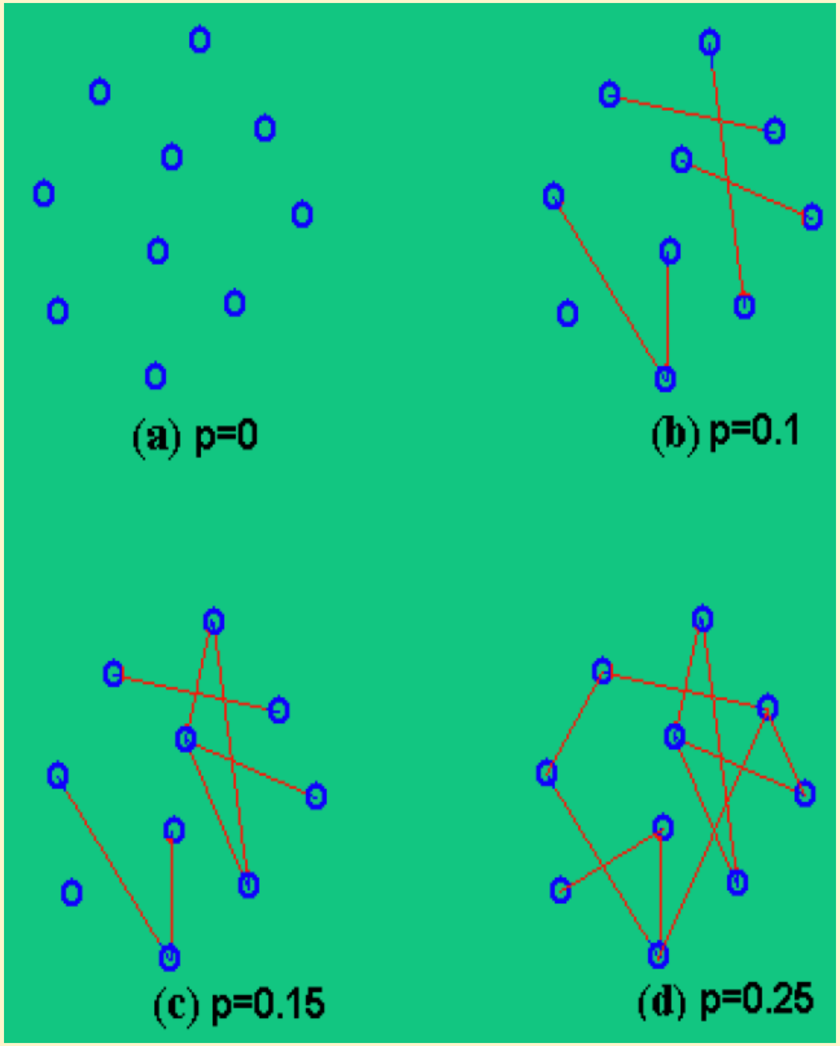
Story of Paul

- As a historical remark, Paul is one of the most distinguished leading mathematicians of the twentieth century, "the man who loves only numbers".
- Paul as legendary; he published more than 1600 research papers with more than 500 co-authors, and he made very fundamental contributions in modern mathematics.
- Paul was always excited when he met and worked with other mathematicians. When he met a colleague, he often said "My brain is open"; when he left one coworker to meet with another, he used to say "another roof, another proof".
- He had devoted his entire life to his beloved mathematics.

Erdös-Rényi Random-Graph Networks

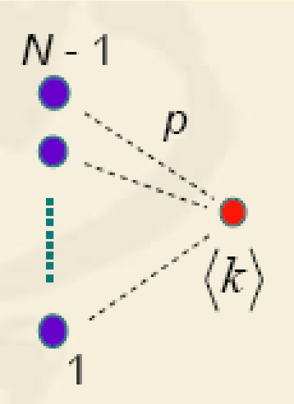
- Given N isolated nodes
- For every pair of nodes, with probability p add an edge between the two nodes
- There will be a total of $pN(N-1)/2$ edges





Random-Graph Networks

- Average degree: $\langle k \rangle = p(N-1) \approx pN$



- Clustering coefficient: (to be proved later)

$$C = \langle k \rangle / N = p$$

- Degree distribution *P(k)-probability that a node has degree k*: (proof on the next page)

Poisson

$$P(k) = \frac{\mu^k}{k!} e^{-\mu}$$

Proof of Poisson Distribution: $P(k) = \frac{\mu^k}{k!} e^{-\mu}$

- Among *N* pairs of nodes, the probability of having exactly *k* edges is $P(k \mid N) = \binom{N}{k} p^k (1-p)^{N-k}$
- Expectation value is the average degree $\mu = pN$
- Viewing the above distribution as a function of μ rather than *N* and for a fixed probability *p*, one can rewrite the distribution as

$$P_\mu(k \mid N) = \frac{N!}{k!(N-k)!} \left(\frac{\mu}{N}\right)^k \left(1 - \frac{\mu}{N}\right)^{N-k}$$

- Thus,

$$P(k) = \lim_{N\to\infty} P_\mu(k \mid N)$$

$$= \lim_{N\to\infty} \frac{N(N-1)\cdots(N-k+1)}{N^k} \frac{\mu^k}{k!} \left(1 - \frac{\mu}{N}\right)^N \left(1 - \frac{\mu}{N}\right)^{-k}$$

$$= 1 \cdot \frac{\mu^k}{k!} \cdot e^{-\mu} \cdot 1$$

Random Graph and Poisson Degree Distribution

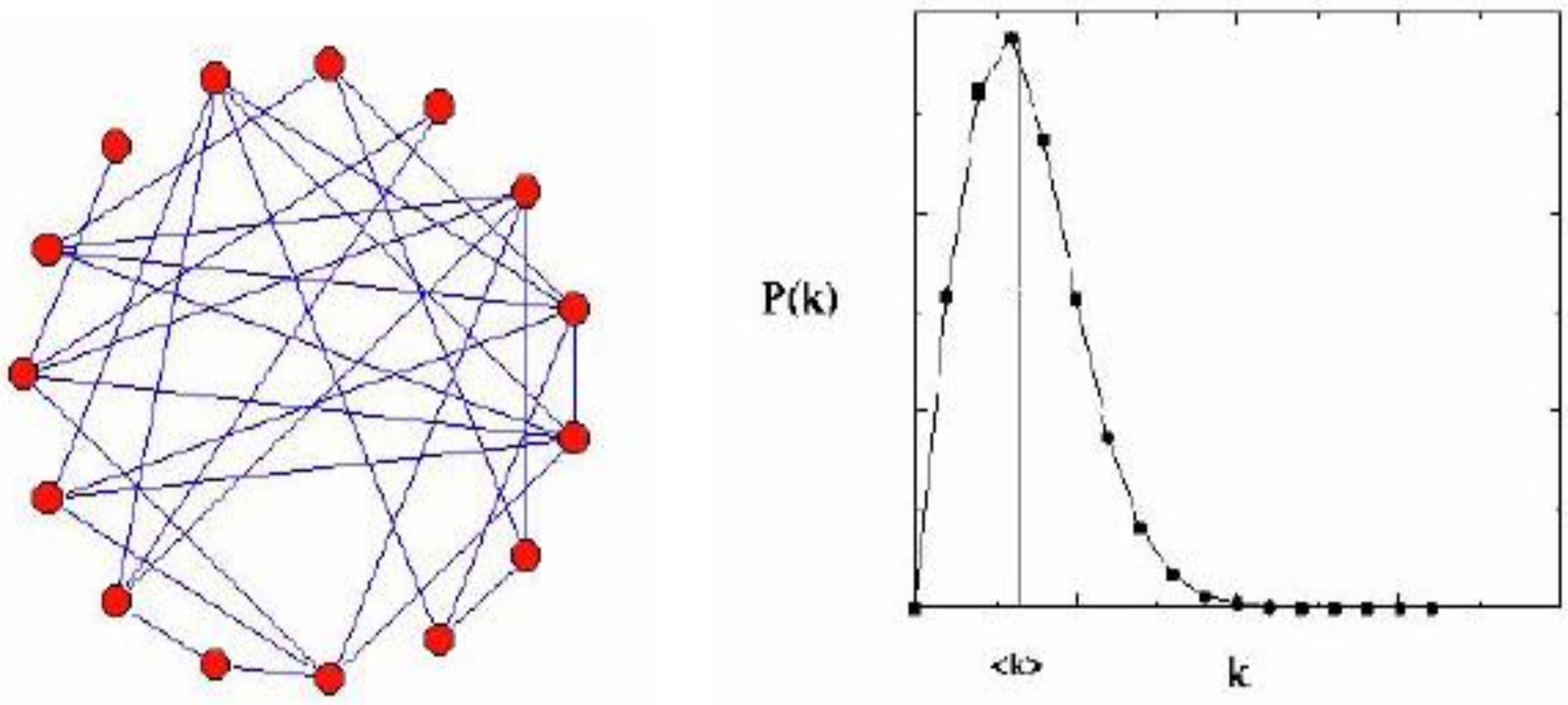


Illustration of Erdös-Rényi randon-graph network model

A Legitimate Question

- Erdös and Rényi were far more intrigued by the mathematical beauty of random-graph networks than by the model's ability to faithfully capture natural networks' properties.
- Actually, what do real networks look like?

Statistical Properties of Some Real-World Complex Networks

	Network	Type	N	M	$<k>$	L	γ	C
Social science	Film actors	undirected	449913	25516482	113	3.48	2.3	0.78
	Company directors	undirected	7673	55392	14.4	4.6	–	0.88
	Math coauthorship	undirected	253339	496489	3.92	7.57	–	0.34
	Physics coauthouship	undirected	52909	245300	9.27	6.19	–	0.56
	Biology coauthorship	undirected	1520251	11803064	15.5	4.92	–	0.6
	Telephone call graph	undirected	47000000	80000000	3.16			
	E-mail messages	undirected	59912	86300	1.44	4.95	1.5/2.0	0.16
	E-mail addresses books	undirected	16881	57029	3.38	5.22	–	0.13
	Student relationships	undirected	573	477	1.66	16	–	0
	Sexual contacts	undirected	2810				3.2	

N – number of nodes; *M* – number of edges

<k> average degree;

L average path length;

r the exponent of a power-law distribution;

C clustering coefficient.

Statistical Properties of Some Real-World Complex Networks

	Network	Type	N	M	$<k>$	L	γ	C
Information Science	WWW nd.edu	directed	269504	1497135	5.55	11.3	2.1/2.4	0.29
	WWW Altavista	directed	203549046	2.13E+09	10.5	16.2	2.1/2.7	
	Citation network	directed	783339	6716198	8.57		3.0/–	
	Roget's Thesaurus	directed	1022	5103	4.99	4.87	–	0.15
	Word co-occurrence	undirected	460902	1.7E+07	70.1		2.7	0.44
Technology	Internet (AS-level)	undirected	10697	31992	5.98	3.31	2.5	0.39
	Power grid	undirected	4941	6594	2.67	19	–	0.08
	Train routes	undirected	587	19603	66.8	2.16	–	0.69
	Software packages	directed	1439	1723	1.2	2.42	1.6/1.4	0.08
	Software classes	directed	1377	2213	1.61	1.51	–	0.01
	Electric circuits	undirected	24097	53248	4.34	11.1	3	0.03
	Peer-to-peer network	undirected	880	1296	1.47	4.28	2.1	0.01

Statistical Properties of Some Real-World Complex Networks

	Network	Type	N	M	$<k>$	L	γ	C
Biology	Metabolic network	undirected	765	3686	9.64	2.56	2.2	0.67
	Protein network	undirected	2115	2240	2.12	6.8	2.4	0.07
	Marine food web	directed	135	598	4.43	2.05	–	0.23
	Freshwater food web	directed	92	997	10.8	1.9	–	0.09
	Neural network	directed	307	2359	7.68	3.97	–	0.28

Common Features:

Small L and Small C → Random-Graph Networks

Small L but Large C → Small-World Networks

Power-Law $\sim k^{-\gamma}$ → Scale-Free Networks

Basic Studies of Complex Networks

- ***Discovering:*** Trying to reveal the global statistical properties of a network and to develop measures for these properties.
- ***Modeling:*** Trying to establish a mathematical model of a given network, enabling better understanding of the network statistical properties and the causes of their appearance.
- ***Analysis:*** Trying to find out the basic characteristics and essential features of nodes, edges, and the whole network in a certain topology, to develop fundamental mathematical theories that can describe and predict the network dynamical behaviors.
- ***Control:*** Trying to develop effective methods and techniques that can be used to modify and improve network properties and performances, suggesting new and possibly optimal network designs and utilizations, particularly in the regards of network stability, synchronization, and data-traffic management.
- ***Applications:*** Trying to apply and utilize some special and fundamental properties and characteristics of complex networks to facilitate the design and applications of network-related problems, such as data-flow congestion control on the Internet and traffic control for city transportations, optimal integrated circuit design for chip design for chip fabrication, better decision-making of policy and strategy for commercial trading and financial management, etc.

Some Basic Concepts

- Distance and Average Path Length
- Clustering Coefficient
- Degree and Degree Distribution
- Other Concepts (Betweenness, Assortativeness, …)

  will be introduced in latter lectures

(Stephen G. Eick)

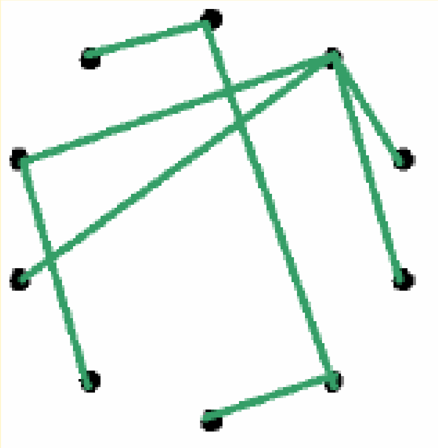
Network as a Graph

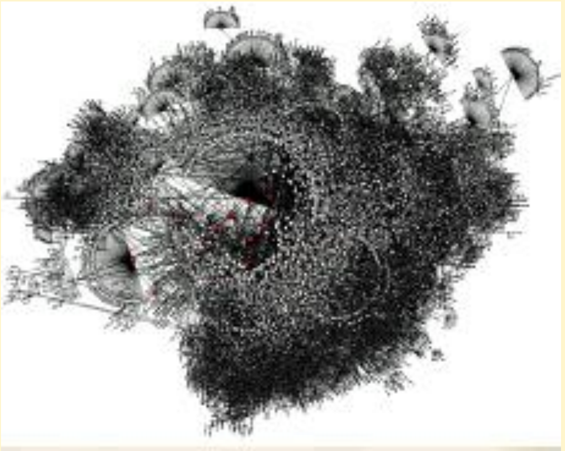
A given network can be represented by a Graph $G=(N,E)$

N – *set of nodes*

E – *set of edges*

Only undirected and unweighted networks are studied:





Distance and Average Path Length

- Distance $d_{i,j}$ between two nodes i and j

  = the number of edges along the shortest path connecting them

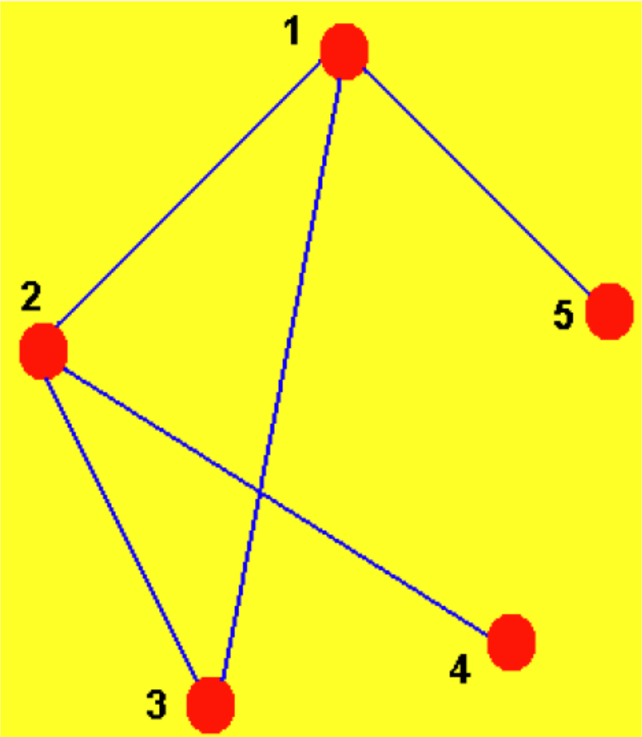
- Diameter D = max$\{d_{i,j}\}$
- Average path length L = average over all $d_{i,j}$
- Many large and complex networks:

  small L -> small-world feature

Distance and Average Path Length

Example:

A network having N = 5 nodes and 5 edges



$$L = \frac{1}{\frac{1}{2}N(N-1)} \sum_{i<j} d_{ij}$$

$$\begin{array}{llll}
d_{12} = 1 & d_{13} = 1 & d_{14} = 2 & d_{15} = 1 \\
 & d_{23} = 1 & d_{24} = 1 & d_{25} = 2 \\
 & & d_{34} = 2 & d_{35} = 2 \\
 & & & d_{45} = 3
\end{array}$$

Total = 16

Average: L = 16/10 = 1.6

Clustering Coefficient

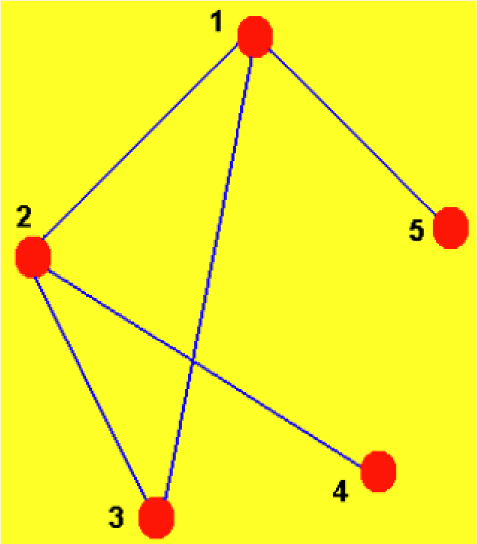
- Clustering Coefficient C of a network
- How many of your friends are also friends themselves?
  - Consider a node, i
  - Suppose that node i has $k(i)$ edges (so it has $k(i)$ neighbors)
  - $E(i)$ = number of existing edges of these $k(i)$ neighbors actually have
  - $T(i)$ = number of possible edges of these $k(i)$ neighbors can have

    $= k(i)\,[k(i)-1]/2$
  - $C(i) = E(i)/T(i)$ – clustering coefficient of node i
  - C= average over all $C(i)$ – clustering coefficient of the network
- Usually, $0<C<1$
  - $C=1$ if all neighbors of a node are connected pair-wise
  - $C=0$ if no neighbors of a node are connected pair-wise
- Many large and complex networks:

  large C -> small-world feature

How to compute the clustering coefficient ?

- Example:
  - Node *i* has *k*(i) edges (so it has *k*(i) neighbors)
  - *E*(i) = number of edges of these *k*(i) neighbors actually have
  - *T*(i) = number of edges of these neighbors can possibly have
    = *k*(i) [*k*(i) -1]/2
  - *C*(i) = *E*(i)/ *T*(i) = 2*E*(i)/[*k*(i)(*k*(i) -1)]
  - *C*= average over all *C*(i)



Node-1 has 3 neighbors, $E(1) = 1$, so $C(1) = 2\text{x}1/(3\text{x}2) = 1/3$

Node-2 has 3 neighbors, $E(2) = 1$, so $C(2) = 2\text{x}1/(3\text{x}2) = 1/3$

Node-3 has 2 neighbors, $E(3) = 1$, so $C(3) = 2\text{x}1/(2\text{x}1) = 1$

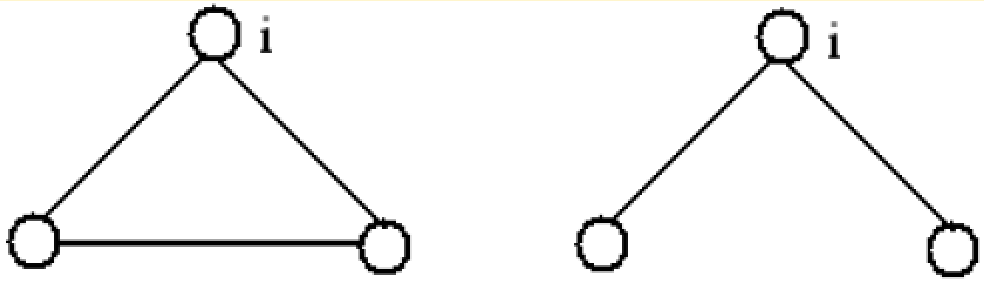
Node-4 has 1 neighbor, $E(4) = 0$, so $C(4) = 0$

Node-5 has 1 neighbor, $E(5) = 0$, so $C(5) = 0$

Average $C = (1/3+1/3+1+0+0) / 5 = 1/3$

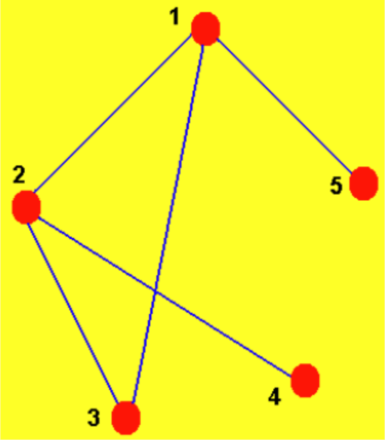
Another definition from the geometric viewpoint

$$C(i) = \frac{number\ of\ complete\ triangles\ with\ corner\ i}{number\ of\ all\ triangular\ graphs\ with\ corner\ i}$$



Left – complete triangle is counted in both numerator and denominator

Right – incomplete triangular graph is counted only in denominator



Node-1 has 1 **complete triangle and** 3 **triangular graphs, so** $C(1) = 1/3$

Node-2 has 1 **complete triangle and** 3 **triangular graphs, so** $C(2) = 1/3$

Node-3 has 1 **complete triangle and** 1 **triangular graph, so** $C(3) = 1$

Node-4 has 0 **complete triangles, so** $C(4) = 0$

Node-5 has 0 **complete triangles, so** $C(5) = 0$

Average $C = (1/3+1/3+1+0+0) / 5 = 1/3$

Degree and Degree Distribution

- Degree k(i) of node i

  = the total number of its edges

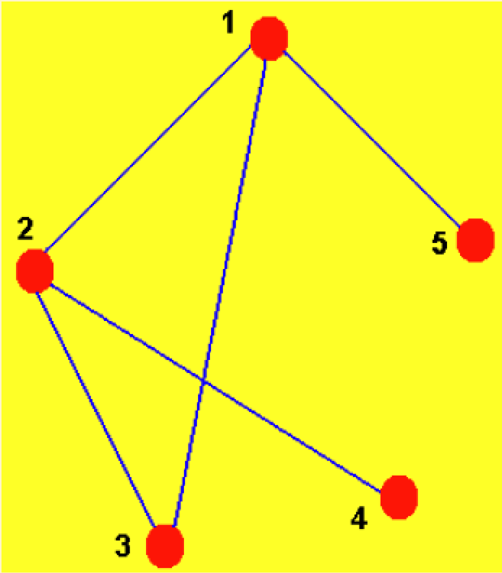
- Average Degree $\langle k \rangle$ over the network

- The spread of node degrees over a network is characterized by a distribution function:

  $P(k)$ = probability that a randomly selected node has exactly degree k

Degree and Degree Distribution

Example:



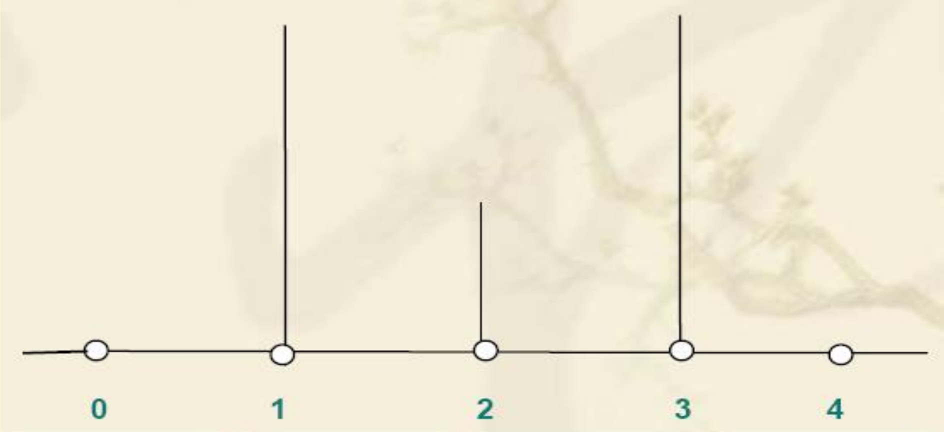
- Degree:

$$\text{node1} = 3, \quad \text{node2} = 3, \quad \text{node3} = 2,$$
$$\text{node4} = 1, \quad \text{node5} = 1$$

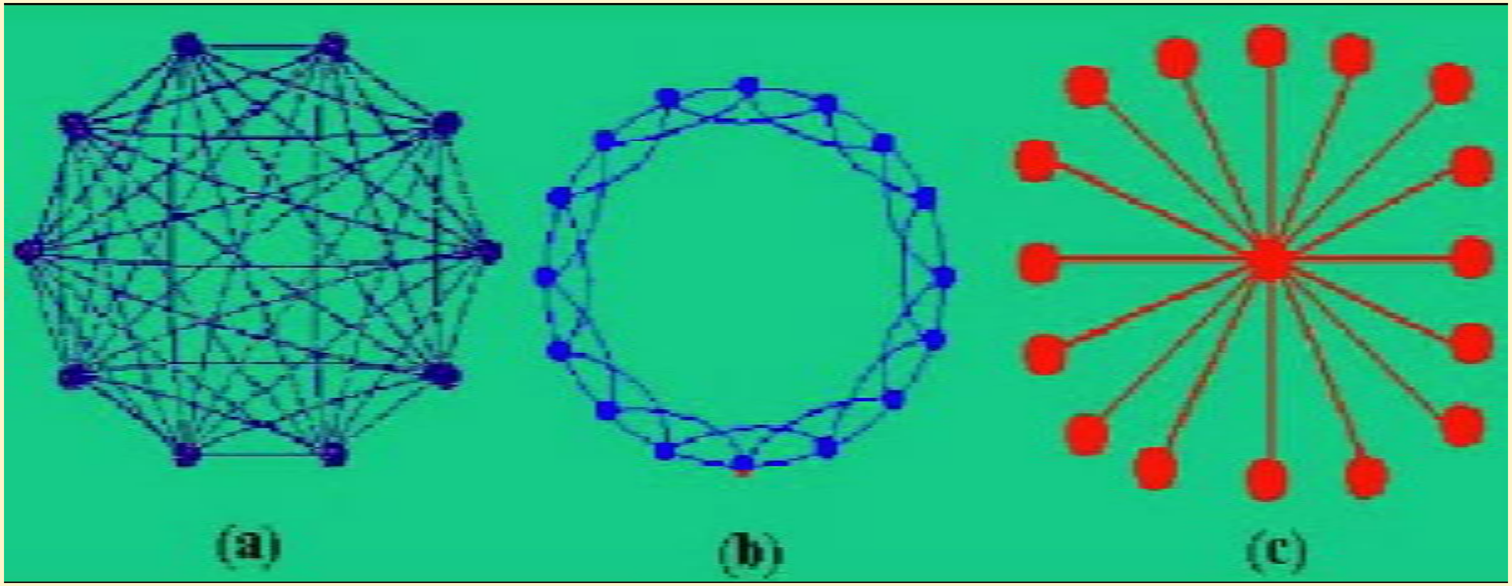
- Average Degree:

$$\langle k \rangle = (3 + 3 + 2 + 1 + 1) / 5 = 2$$

- Degree Distribution:

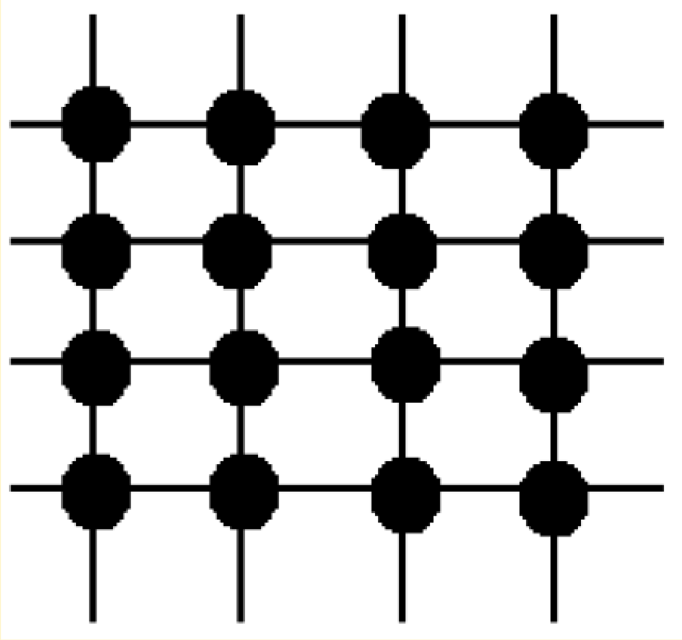


Regular Networks

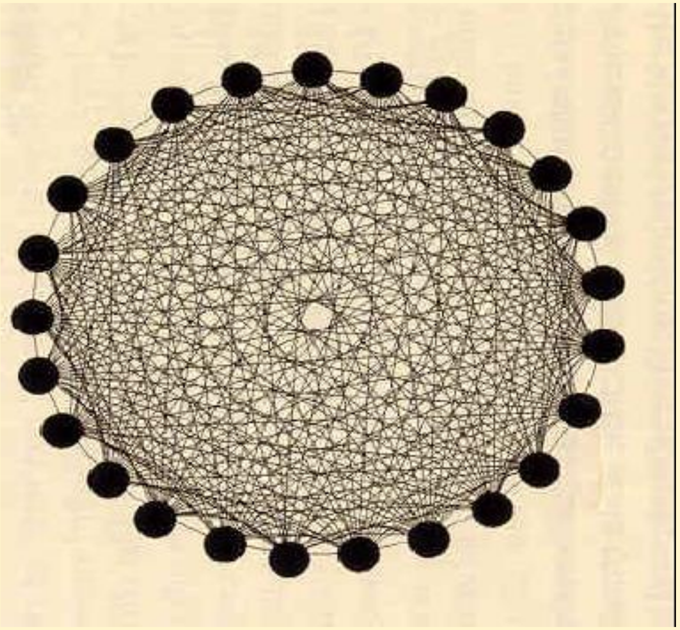


- (a) Fully-connected network
- (b) Ring-shaped coupled network
- (c) Star-shaped coupled network

Regular Networks (continued)



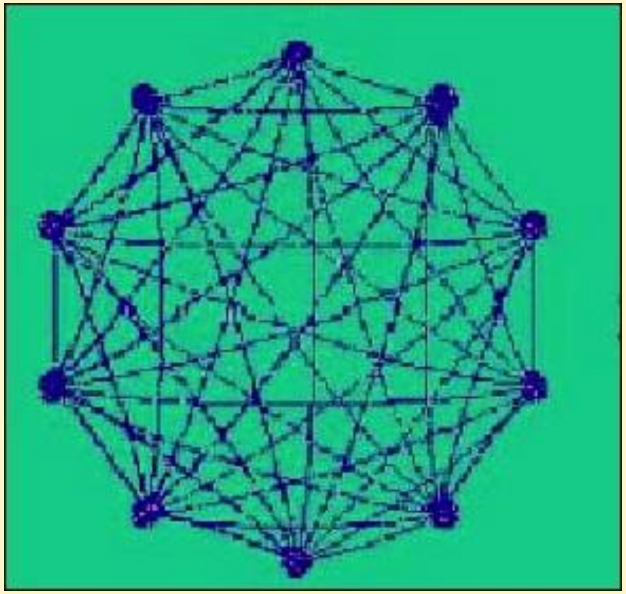
…



(a) Lattice … (z) Layers

Basic Properties of Regular Networks:

(a) Fully-connected networks



- Average path length:

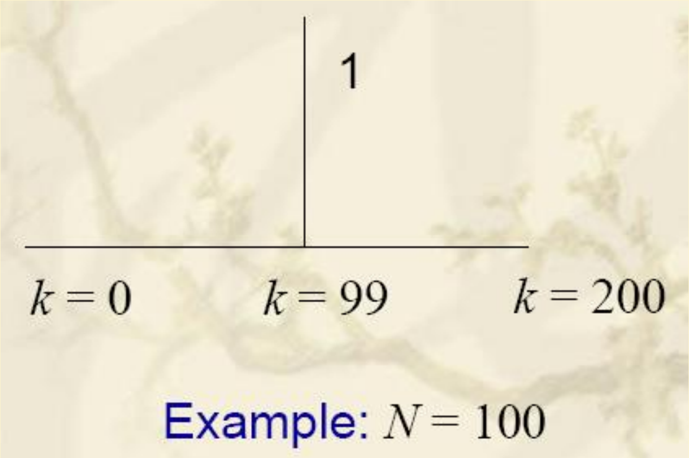
$$L_{full} = 1$$

- Network clustering coefficient:

$$C_{full} = 1$$

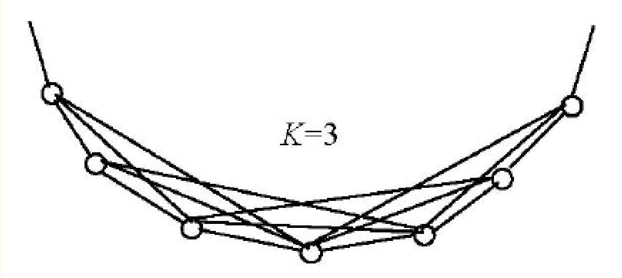
- Degree distribution:

delta



Basic Properties of Regular Networks:

(b) Ring-shaped coupled networks



- Average path length: (not proved)

  M – number of edges,

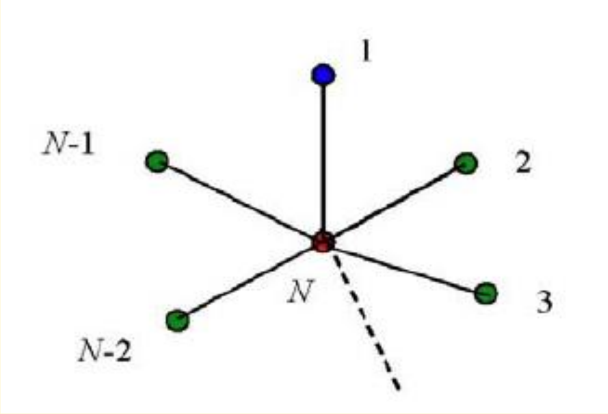
  K – number of neighboring connections

$$L_{ring} \sim \frac{M(M+1)-2(K-1)(M-K+1)}{2M} \to \infty \quad (M \to \infty)$$

- Clustering coefficient: (not proved)

$$C_{ring} = \frac{3(K-1)}{2(2K-1)} \to \frac{3}{4} \qquad (K \to \infty)$$

(c) Star-shaped coupled networks



- Average path length:

$$L_{star} = 2 - \frac{2}{N} \to 2 \qquad (N \to \infty)$$

verifying:

$$\frac{(N-1)\times 1 + [(N-2)+(N-3)+\cdots+2+1]\times 2}{(N-1)+[(N-2)+(N-3)+\cdots+2+1]} = 2 - \frac{2}{N}$$

- Clustering coefficient:

$$C_{star} = 0$$

Random-Graph Networks

- Average degree:

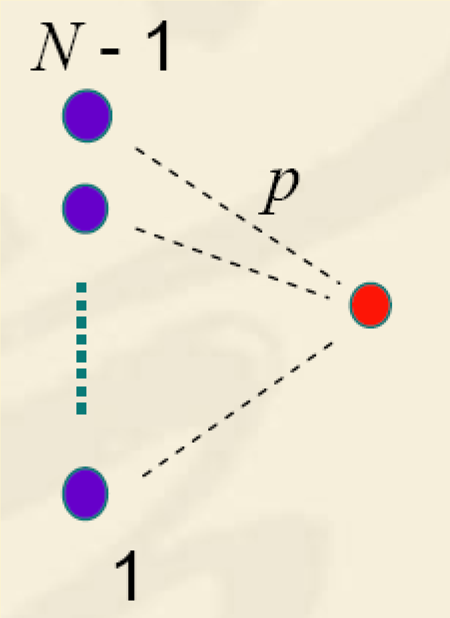
$$\langle k \rangle = p(N-1) \approx pN$$

- Clustering coefficient:

$$C = \langle k \rangle / N = p$$

- Degree distribution: Poisson

$$P(k) = \frac{\mu^k}{k!} e^{-\mu}$$



Degree Distributions

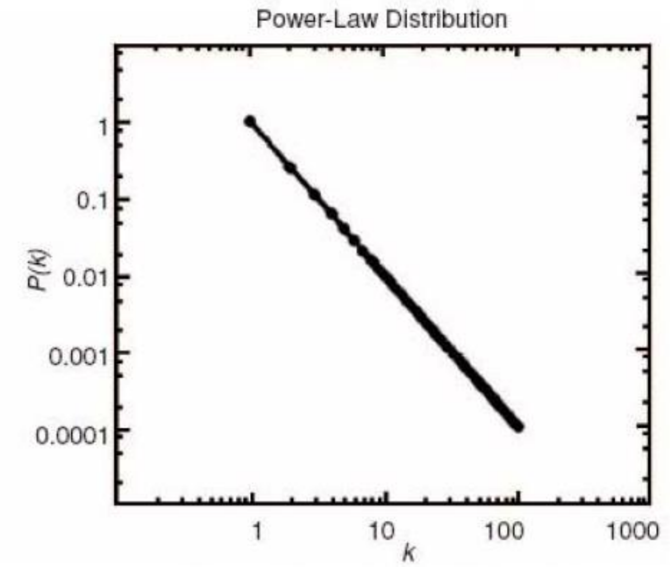
- Completely regular lattice:

  P(k) ~ Delta distribution

- Most networks:

  $P(k) \sim k^{-\gamma}$ (Power Law)

  - scale-free feature



(b) Power-law distribution (log-log)

- Completely random networks:

  P(k) ~ Poisson distribution

(Regular) Delta --- $P(k) \sim k^{-\gamma}$ --- Poisson (Random)

Theorem 1-2 (*Scale-free property of power-law distributions*) Consider a probability distribution function $f(x)$. If, for any given constant a, there is a constant b such that the following "scale-free" property holds:

$$f(ax) = bf(x) \tag{1-8}$$

then, with the assumption of $f(1)f'(1) \neq 0$, the function $f(x)$ is uniquely determined by

$$f(x) = f(1)\,x^{-\gamma}, \quad \gamma = -f'(1)/f(1) \tag{1-9}$$

Proof. Let $x=1$ in (1-8). Then, one has $f(a) = bf(1)$, so $b = f(a)/f(1)$; therefore,

$$f(ax) = \frac{f(a)f(x)}{f(1)}$$

Since this equality holds for arbitrary a and x, one may also consider a as a variable and take a derivative of the equality with respect to a, obtaining

$$\frac{df(ax)}{d(ax)}\frac{d(ax)}{da} = \frac{f(x)}{f(1)}\frac{df(a)}{da}$$

In particular, letting $a=1$ gives

$$x\frac{df(x)}{d(x)} = \frac{f'(1)}{f(1)}f(x)$$

which has a unique solution

$$\ln f(x) = \frac{f'(1)}{f(1)}\ln x + \ln f(1)$$

This is equivalent to (1-9), completing the proof of the theorem. □

Contact me at:

Email: hjzhang@hit.edu.cn
Tel: 26033086
Office: Rm.1601, Bld. L