

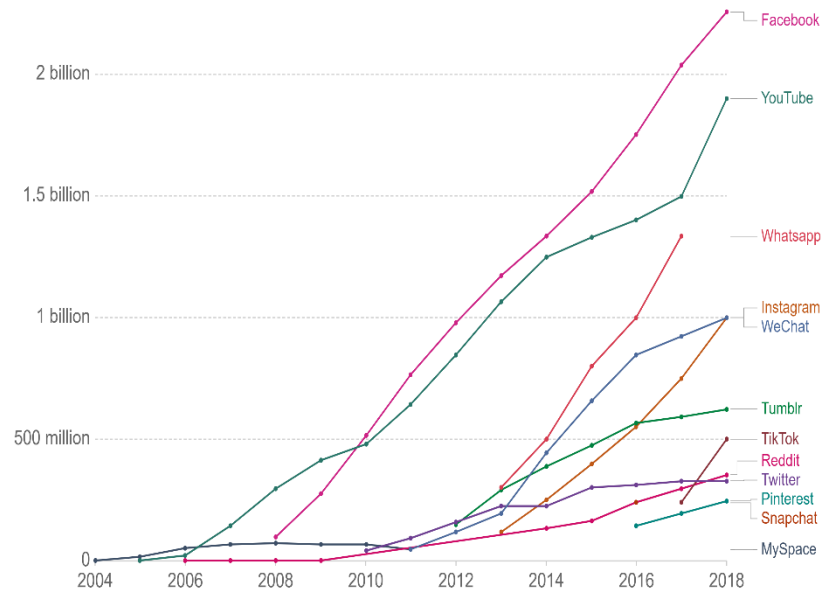


Network Growth

Rise of Online Social Networks

Number of people using social media platforms, 2004 to 2018

Estimates correspond to monthly active users (MAUs). Facebook, for example, measures MAUs as users that have logged in during the past 30 days. See source for more details.

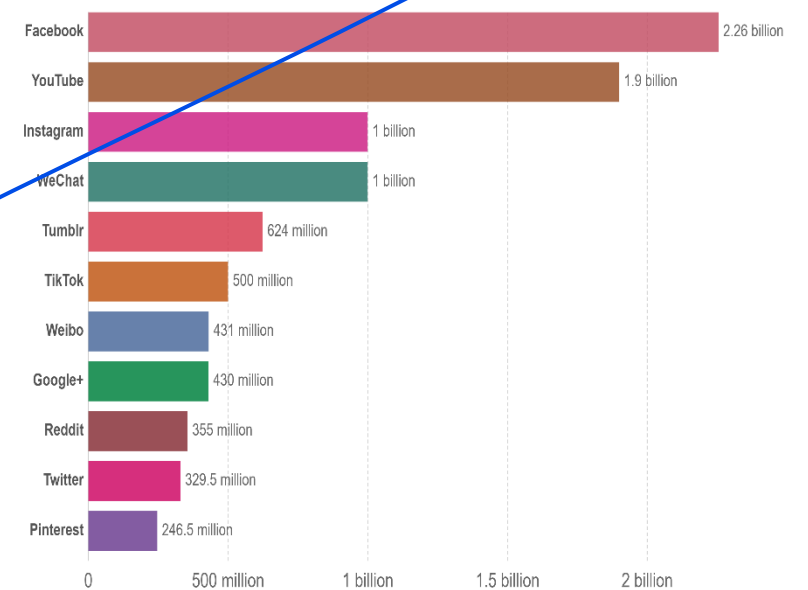


Source: Statista and TNW (2019)

- Online social networks growing rapidly year-by-year
- Their sizes are huge
- In 2014, Facebook possessed 1.39 billion active users and 400 billion friendship links
- March 2015, Twitter has 288 million active users and 60 billion followers

Number of people using social media platforms, 2018

Estimates correspond to monthly active users (MAUs). Facebook, for example, measures MAUs as users that have logged in during the past 30 days. See source for more details.



Source: Statista and TNW (2019)

CC BY

Such huge volume of these online networks restrict the active research with real social networks

<https://ourworldindata.org/rise-of-social-media>

Synthetic Networks

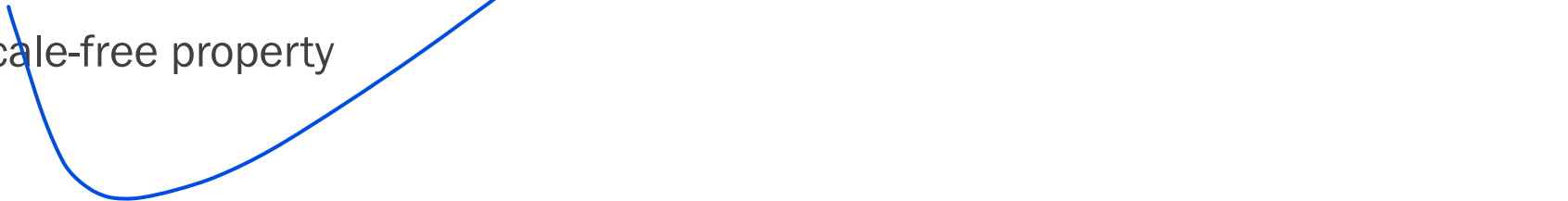
- ❑ Generated using theoretical network models
- ❑ Often possesses strong underlying mathematical foundation
- ❑ Often can simulate important real-world network characteristics
- ❑ Help getting insights of the real-life networks
- ❑ Allow experimentation through simulation when real networks are unavailable
- ❑ can establish network insights on concrete theoretical foundations

Properties of Real-world Networks

☐ High average local clustering coefficient

☐ Small-world property

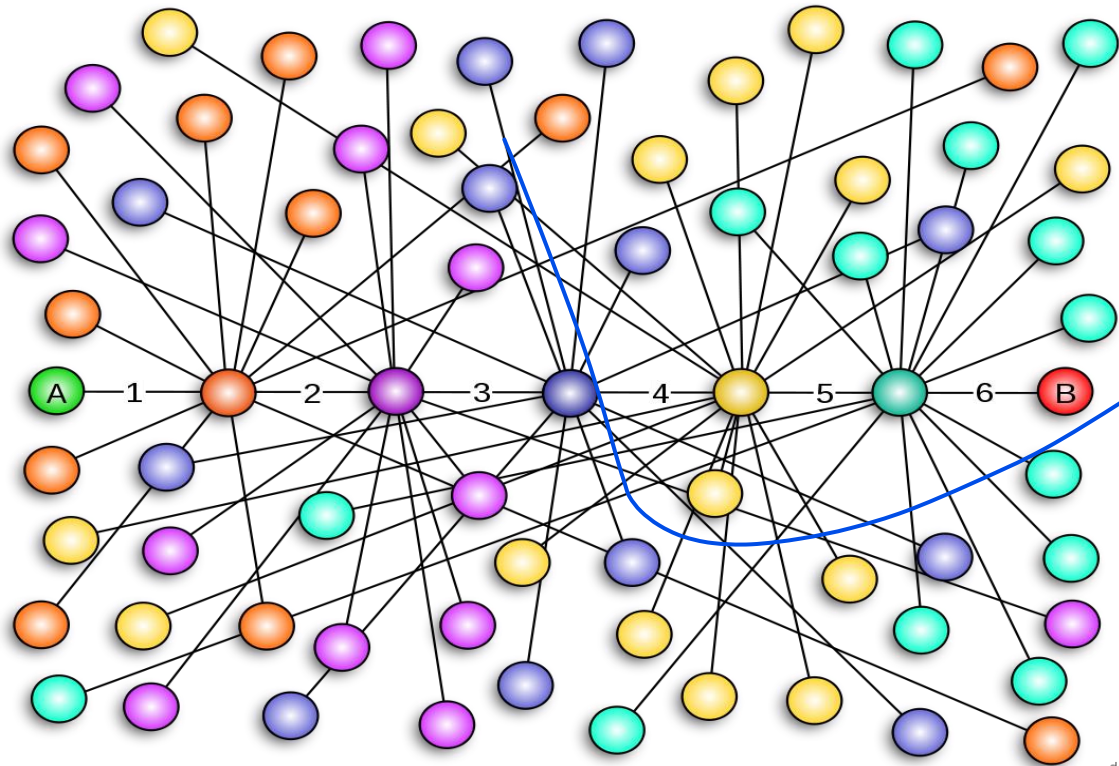
☐ Scale-free property



High Average Local Clustering Coefficient

- ❑ Neighbors of a node tend to be highly connected with each other at individual level
- ❑ In Facebook social graph, for users with 100 friends have average local clustering coefficient of 0.14
- ❑ For a graph of 150 million nodes, the number is unexpectedly high

Small-world Property



□ An outcome of Stanley Milgram's small-world experiment (1967) to measure the probability of two random persons being known to each other

The "six degrees of separation" model

https://en.wikipedia.org/wiki/Small-world_experiment

dw 2010

Six Degrees of Separation: Milgram's Small-World Experiment



A possible path of a letter

https://en.wikipedia.org/wiki/Small-world_experiment

- ☐ Source cities: Omaha, Nebraska, and Wichita, Kansas
- ☐ Destination city: Boston, Massachusetts
- ☐ Information packets sent initially to random persons in Omaha/Wichita
- ☐ Recipient was to forward the letter
 - ☐ directly to target, if she knew the target personally
 - ☐ Else to some friend/relative who more likely to know the target
- ☐ 64 out of 296 letters eventually reach the target contact
- ☐ Number of intermediates among the traversal path chains:

Maximum: 11

Minimum: 1

Median: 5.2 (≈6)

Small-world of Social Media!!!

- ❑ Average chain of contacts in Microsoft Messenger was 6.6 people [[Leskovec and Horvitz 2007](#)]
- ❑ Average distance between two random Facebook users in 2011 was 4.74 with 3.74 intermediaries [[Four Degree of Separation](#) by [Backstrom et al. \(2012\)](#)]
- ❑ Average distance between two random Facebook users in 2016 was 4.57 with 3.57 intermediaries [Repeat Experiment by Facebook in 2016]

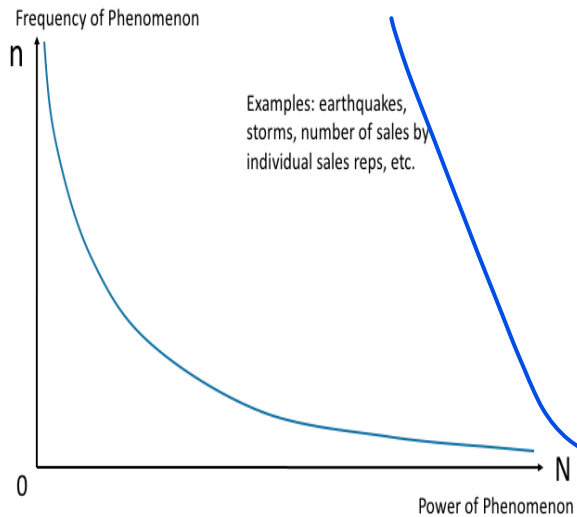
Small-world Property for Social Networks

- ❑ Can be viewed in light of the **average path length between two randomly chosen nodes in a network**
- ❑ A network G is said to follow the **small-world property** if the average path length of the network is logarithmically proportional to the network size.

$$\textit{Average Path Length} \propto \log(\textit{Network Size})$$

Scale-free Property: Power Law

Basic Power Law



© 2016 Alcera Consulting Inc.

- ❑ **Power Law**: a relative change in the value of one variable leads to the proportional change in the value of other variable

- ❑ Independent of the initial values of both the variable

- ❑ Mathematically,

$$y \propto x^{-b}, b \in \mathbb{R}$$

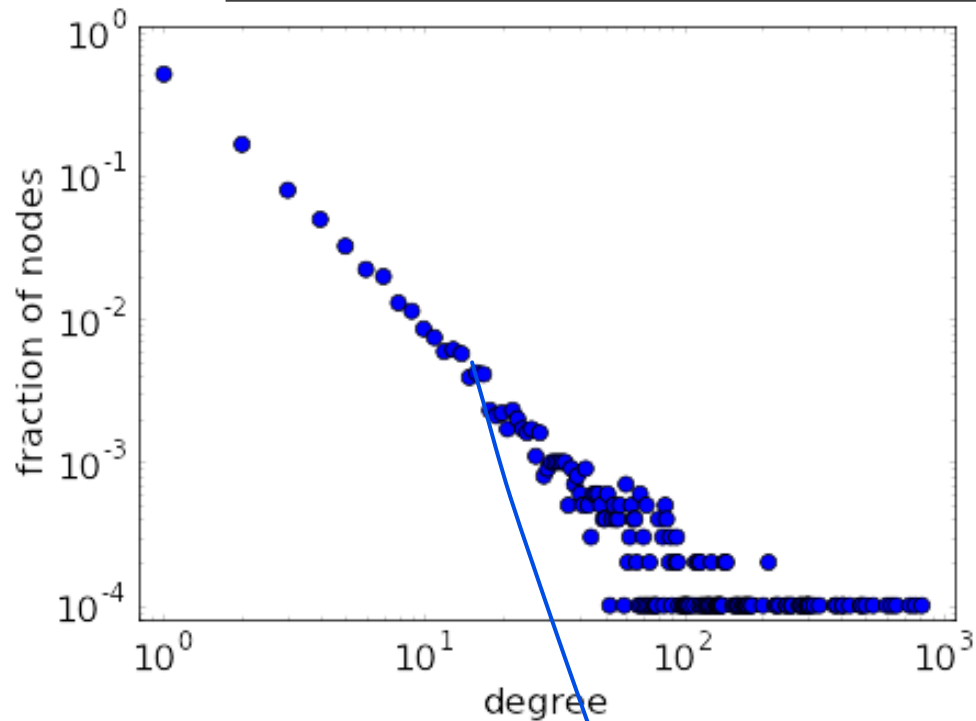
- ❑ Functions that follows power-law are **scale-invariant**

- ❑ **Pareto Principle** (or the **80/20 rule**) in Economics: 80% of the outcomes are results of 20% of the causes

- ❑ Power law principle is also coined as **Pareto Distribution**

<https://exploitingchange.com/2016/09/14/another-powerful-idea-the-power-law/>

Scale-free Property: Real Networks

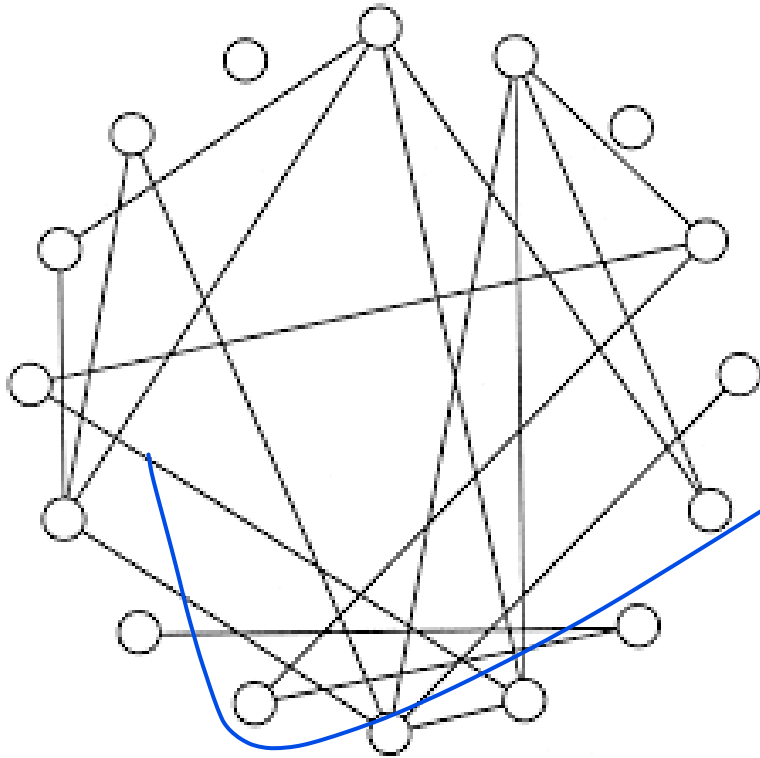


□ Networks whose degree distribution follows power-law are known as **scale-free** networks

□ Most real-world networks are found to be scale-free

Degree distribution of a scale-free network
https://mathinsight.org/scale_free_network

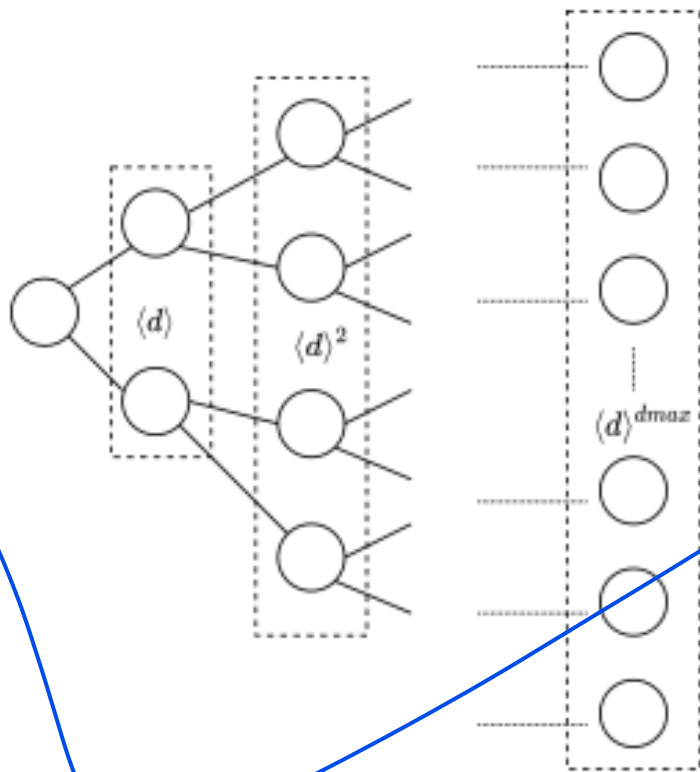
Synthetic Networks: Random Network Model



An instance of $G(16, \frac{1}{7})$ network

- ❑ Also popularly known as [Erdős and Rényi model](#) (or ER model)
- ❑ A number of variants of the model
- ❑ Popular variants:
 - ❑ $G(N, K)$ model [[Erdős and Rényi 1959](#)]: From the set of all networks of N nodes and K edges. a network is chosen uniformly at random
 - ❑ $G(N, p)$ model [[Gilbert 1959](#)]: Network has N nodes, and any random pair of nodes has a probability p of being adjacent independently with any other pair of nodes in the network
- ❑ Both the variants behave identically in the limiting case
- ❑ $G(N, p)$ model considered as the standard random network model

Erdős-Rényi Network: Average Path Length



- When l_{max} represent the maximum path length of $G(N, p)$,

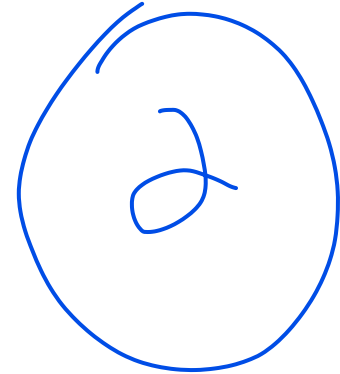
$$1 + \langle d \rangle + \langle d \rangle^2 + \langle d \rangle^3 + \dots + \langle d \rangle^{l_{max}} = N$$

- When $\langle d \rangle \gg 1$, the above yields,

$$l_{max} \approx \frac{\log N}{\log \langle d \rangle}$$

- Further approximation yields,

$$\langle l \rangle \propto \log N$$



Theorem: Erdős-Rényi Networks follow small-world property.

A depiction of random network in tree format

Erdős-Rényi Network: Clustering Coefficient

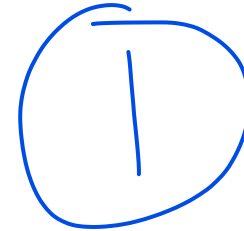
□ In $G(N, p)$

- Number of possible edges between neighbours of a node: $\binom{\langle d \rangle}{2}$
- Expected number of edges between these nodes: $p \times \binom{\langle d \rangle}{2}$

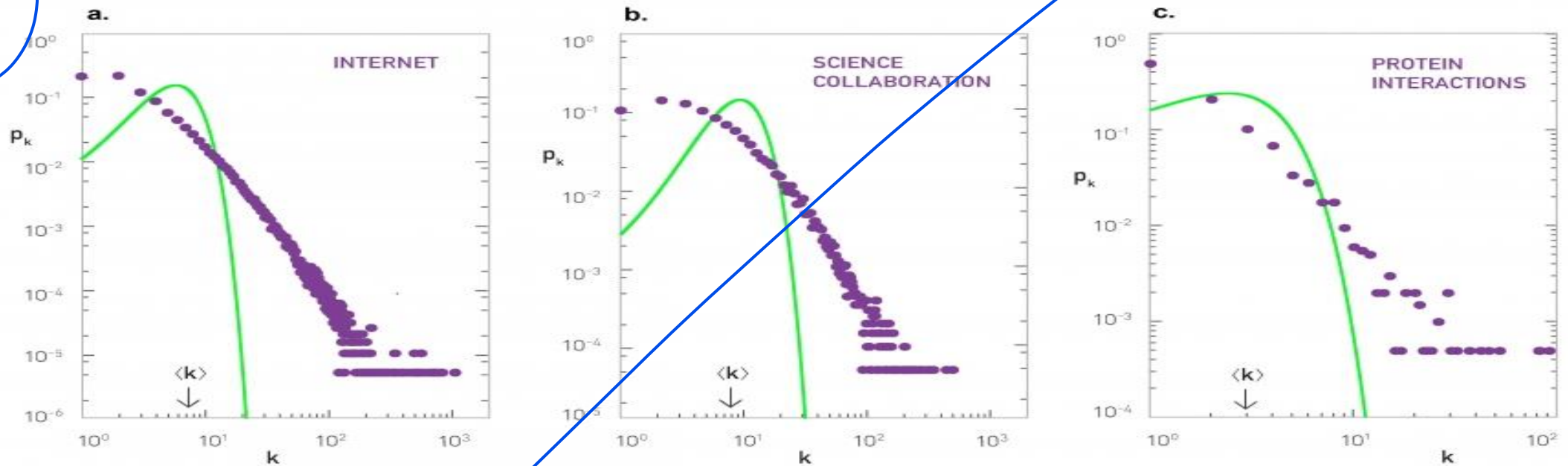
□ The above yields, the local clustering coefficient for a node $v_i \in G$

$$C_i = p \approx \frac{\langle d \rangle}{N}$$

Theorem: The local clustering coefficient for any node in an Erdős-Rényi Network is inversely proportional to the size of the network



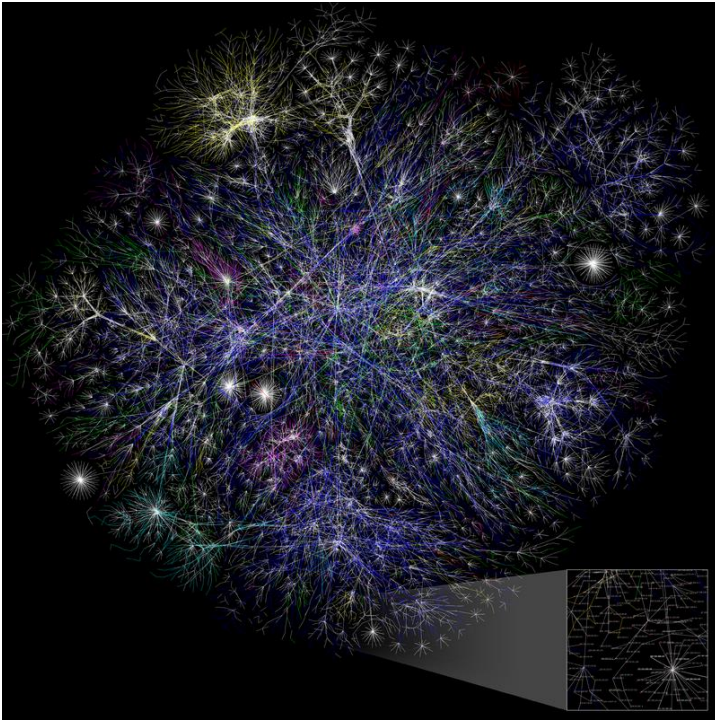
Erdős-Rényi Networks vs. Real-life Networks: Degree Distribution



Real-life networks are often scale-free; however, Erdős-Rényi Networks are not

<http://networksciencebook.com/chapter/3#not-poisson>

Erdős-Rényi Networks vs. Real-life Networks: Presence of Outliers



Partial map of the Internet based on the January 15, 2005. Hubs are highlighted

[https://en.wikipedia.org/wiki/Hub_\(network_science\)](https://en.wikipedia.org/wiki/Hub_(network_science))

- ❑ In an Erdős-Rényi Network, probability of having a node with a high degree is extremely low
- ❑ probability of a node with 2000 neighbours is 10^{-27} !
- ❑ In real-world networks, such nodes exist (Hubs)
- ❑ Celebrities in social networks

Erdős-Rényi Networks vs. Real-life Networks: Small-world Property

□ Erdős-Rényi Networks follow small-world property as the maximum path length in Erdős-Rényi

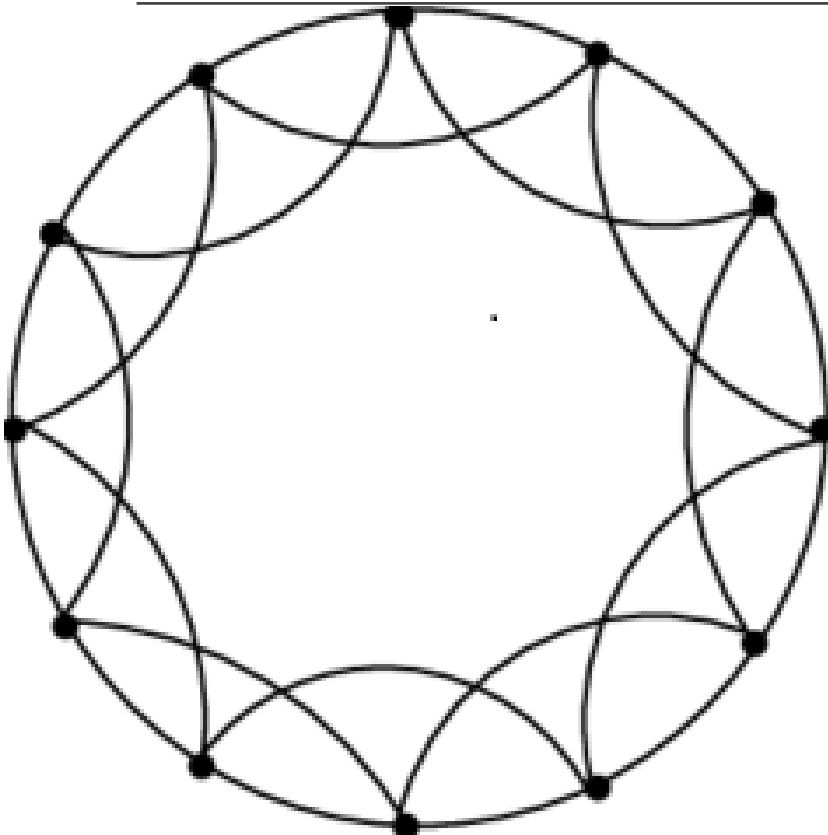
Networks $\approx \frac{\log N}{\log \langle d \rangle}$



Erdős-Rényi Networks vs. Real-life Networks: Clustering Coefficient

- ❑ local clustering coefficients in Erdős-Rényi Networks decreases with the increase in network size
- ❑ Nodes having high local clustering coefficient exist in enormously large real-life networks
- ❑ Echo chambers in large social networks like Facebook

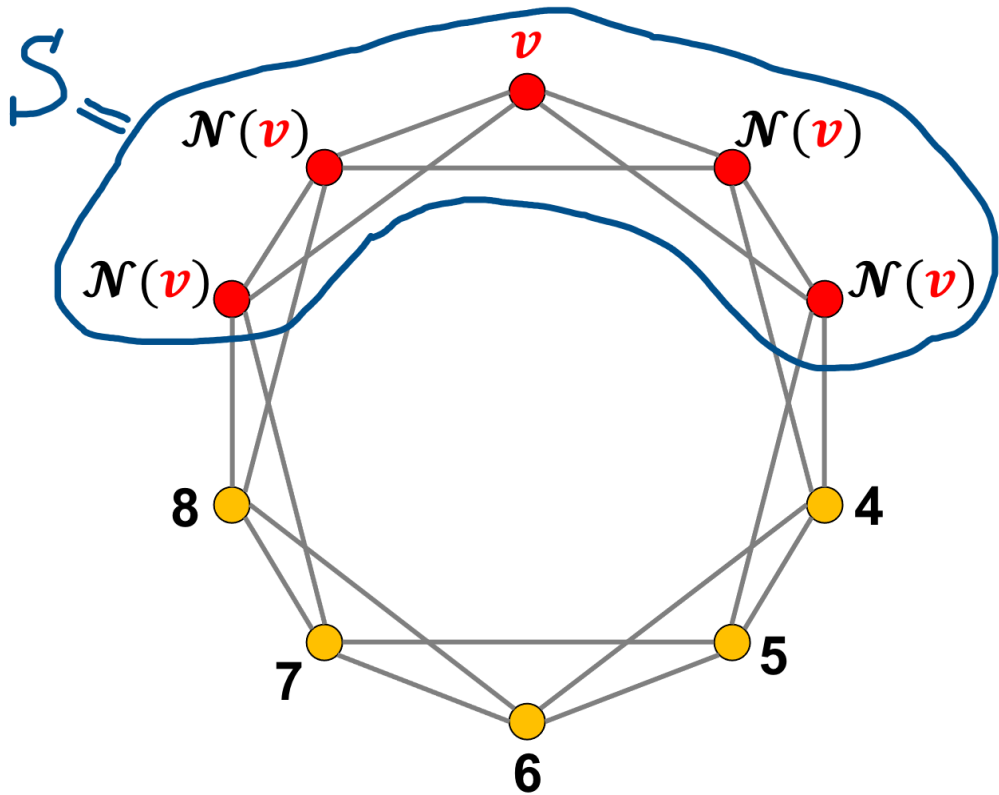
Regular Ring Lattice Network Model



4-Regular Ring Lattice Network

- ❑ A ring lattice network consists of N nodes labeled $0, 1, 2, \dots, N - 1$ arranged in circular order
- ❑ Every node in the network is connected to exactly k other nodes, immediate $\frac{k}{2}$ rightmost nodes and $\frac{k}{2}$ leftmost nodes relative to the position of the node in the network

Regular Ring Lattice Network Model: Local Clustering Coefficient

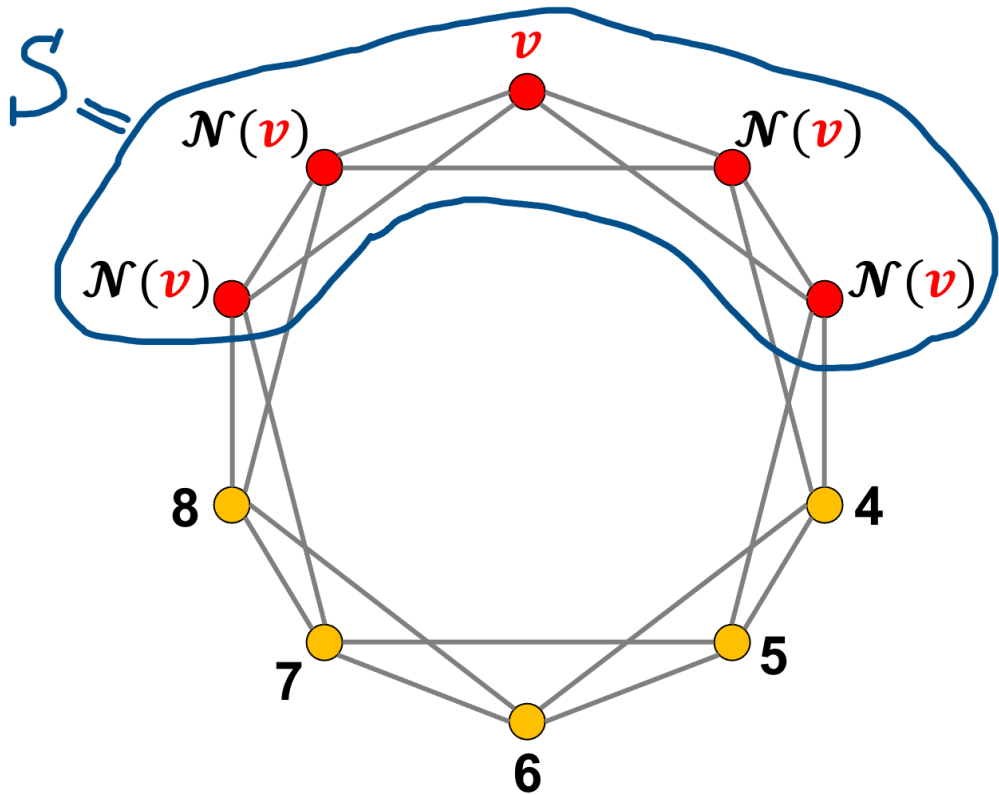


4-Regular Ring Lattice Network

Theorem: Local clustering coefficient of any node v in a k – Regular Ring Lattice Network is $\frac{3(k-2)}{4(k-1)}$

- When k is large, then the above expression boils down to $\frac{3}{4}$
- Regular Ring Lattice networks has high local clustering coefficients

Regular Ring Lattice Network Model: Small-world Property

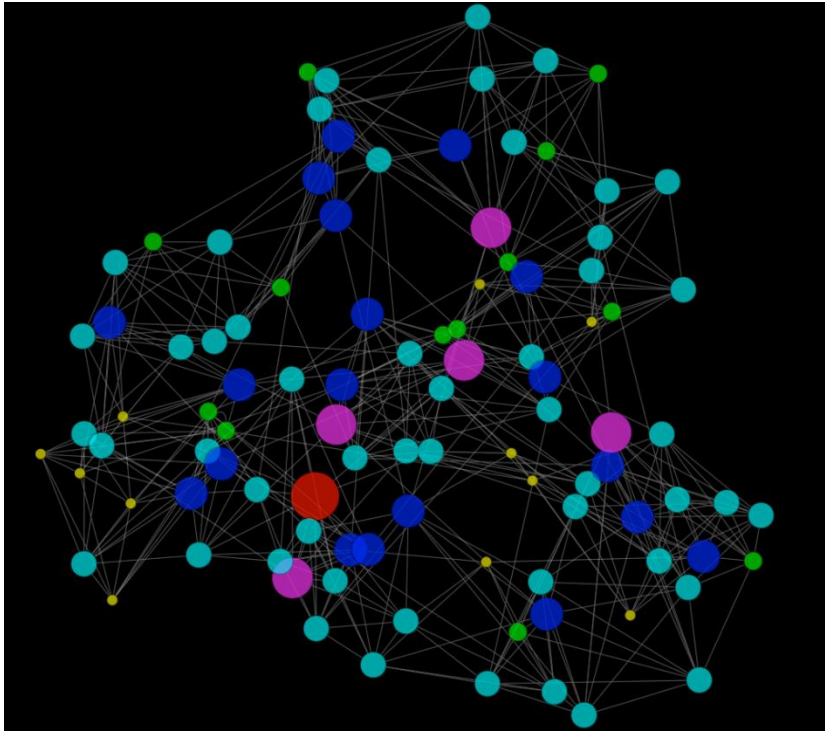


- ❑ All the connection in ring lattice are to nearby nodes
- ❑ Two nodes that are in the opposite sides of the ring will be connected by long chains

Theorem: Regular Ring Lattice networks does not follow small-world property

4-Regular Ring Lattice Network

Watts-Strogatz Network Model



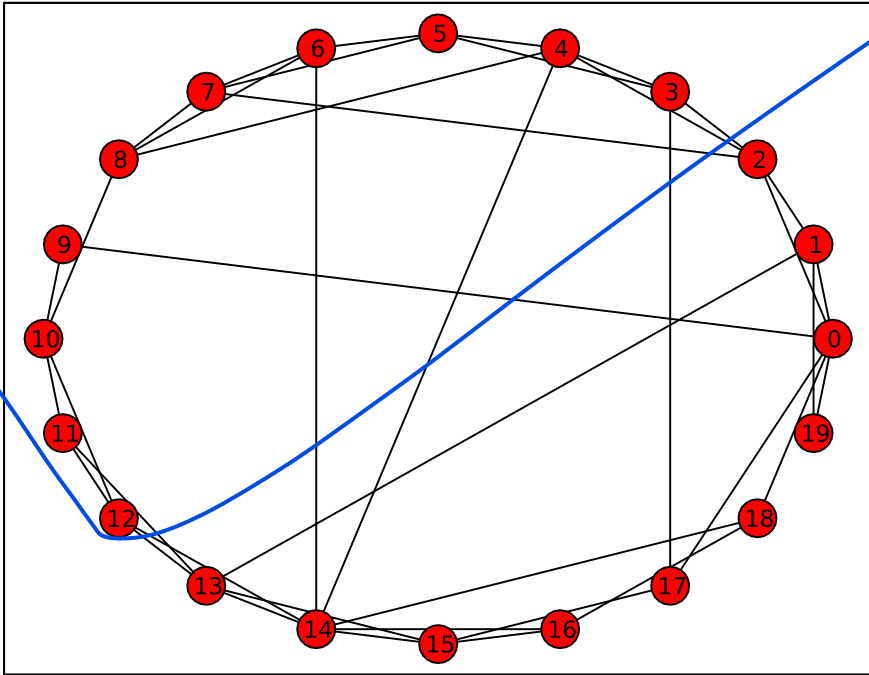
- ☐ Erdős-Rényi Networks
 - ☐ possesses the small-world property,
 - ☐ have small local clustering coefficients for all the nodes when the network size is large
- ☐ Regular Ring Lattice networks
 - ☐ possess high local clustering coefficients for all the node
 - ☐ does not follow small-world property
- ☐ Watts-Strogatz Model form networks that has a high local clustering coefficient and possesses the small-world property
- ☐ Proposed by Watts and Strogatz in 1998

Watts-Strogatz network with 100 nodes formed by igraph and visualized by Cytoscape 2.5

https://en.wikipedia.org/wiki/Watts%E2%80%93Strogatz_model

Watts-Strogatz Network Model: Network Formation

Watts-Strogatz model $N=20$, $K=4$, $\beta=0.2$



<https://www.hindawi.com/journals/mpe/2014/693743/fig2/>

- 1) Start with a k -regular lattice network of size N
- 2) List the nodes of the lattice as $1, 2, \dots, N$
- 3) Choose i^{th} node from the list, $i = 1, 2, \dots, N$
- 4) Select edges that link i^{th} node to some j^{th} node ($j > i$)
- 5) With a fixed rewiring probability β , rewire the other end of these edges
- 6) Avoid formation self-loops and link-duplication
- 7) Repeat steps 3 through 6 until all the nodes are scanned

Watts-Strogatz Network Model: Properties

❑ What if $\beta \rightarrow 0$?

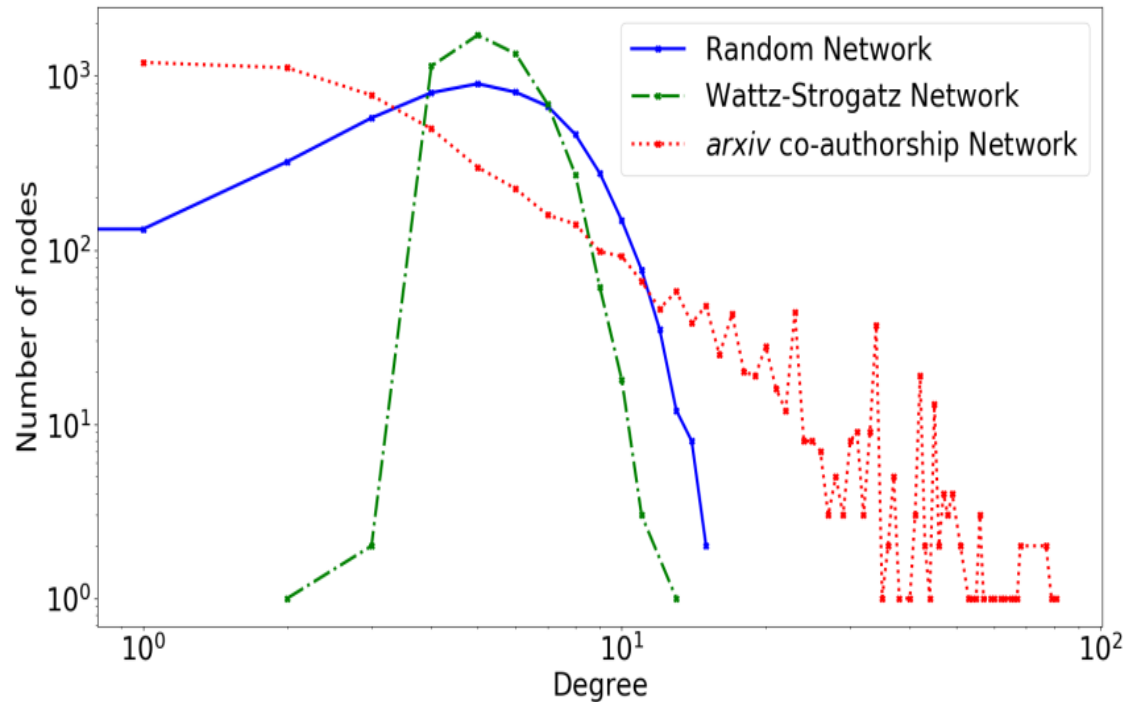
❑ What is the effect on small world property and clustering coefficient ?

❑ What if $\beta \rightarrow 1$?

❑ What is the effect on small world property and clustering coefficient ?

❑ Can we have outliers in Watts-Strogatz Network ? What about real networks ?

Watts-Strogatz Networks vs. Real-world Networks



Comparisons of Degree Distribution

- ❑ Watts-Strogatz networks have only few outliers, whereas, real-life networks have significantly high number of outliers
- ❑ Watts-Strogatz networks follow small-world property if $\beta \rightarrow 1$, however, real-world networks always follow small-world
- ❑ Degree distribution in Watts-Strogatz does not follow power law, whereas, real networks often does that
- ❑ Both Watts-Strogatz networks and real-world networks have high clustering coefficient

Preferential Attachment Model

What is the common property of the following networks ?

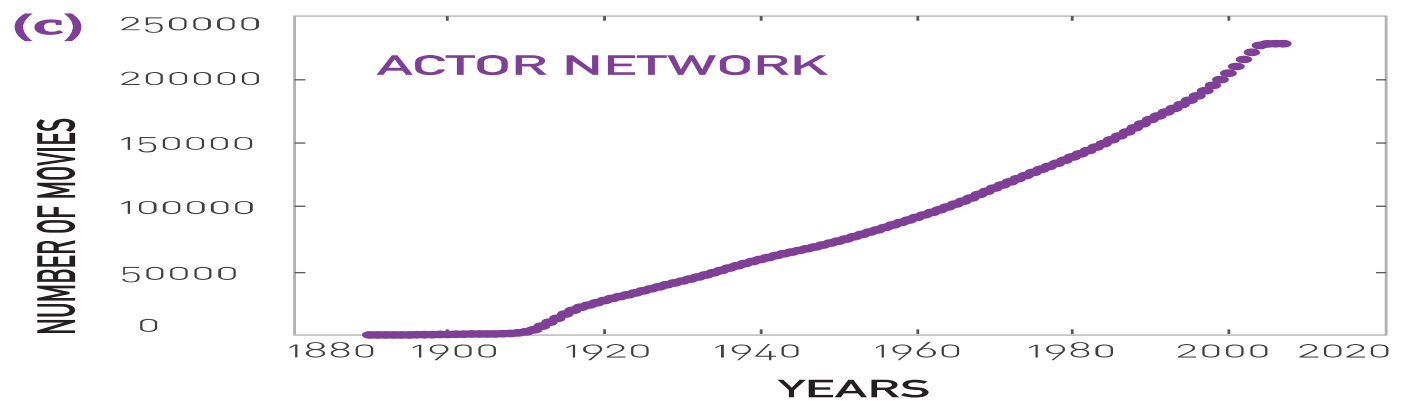
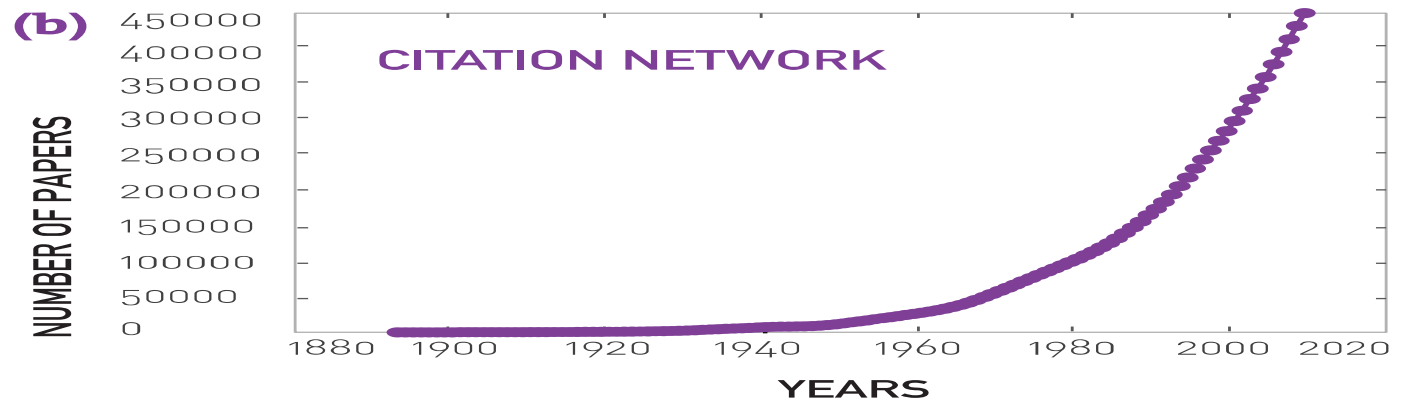
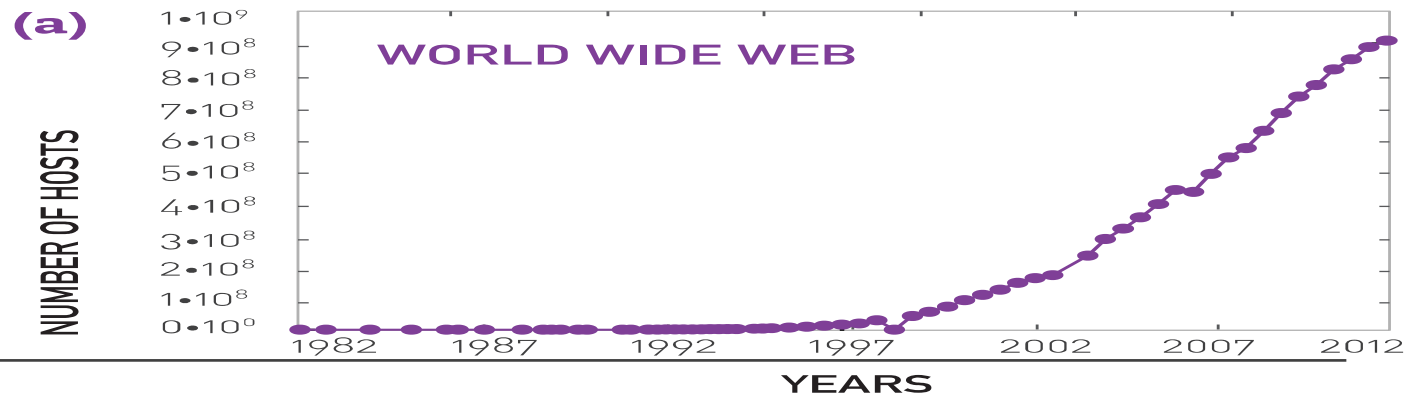
- Citation Network
- WWW Network
- Social Network
- Collaboration Network

They are dynamic, evolve over time, expanding by adding new nodes and edges...

Growth:

Erdos-Renyi ?

Watts-Strogatz ?



Preferential Attachment

New nodes prefer to connect to the more connected nodes

- Do we observe this property in real networks ? Can you give example ?
- Higher degree nodes are preferred over less degree nodes !!!
 - Rich gets richer

Barabasi-Albert Model

either all the nodes are having the same degree or there are few nodes which have a degree greater than all or some

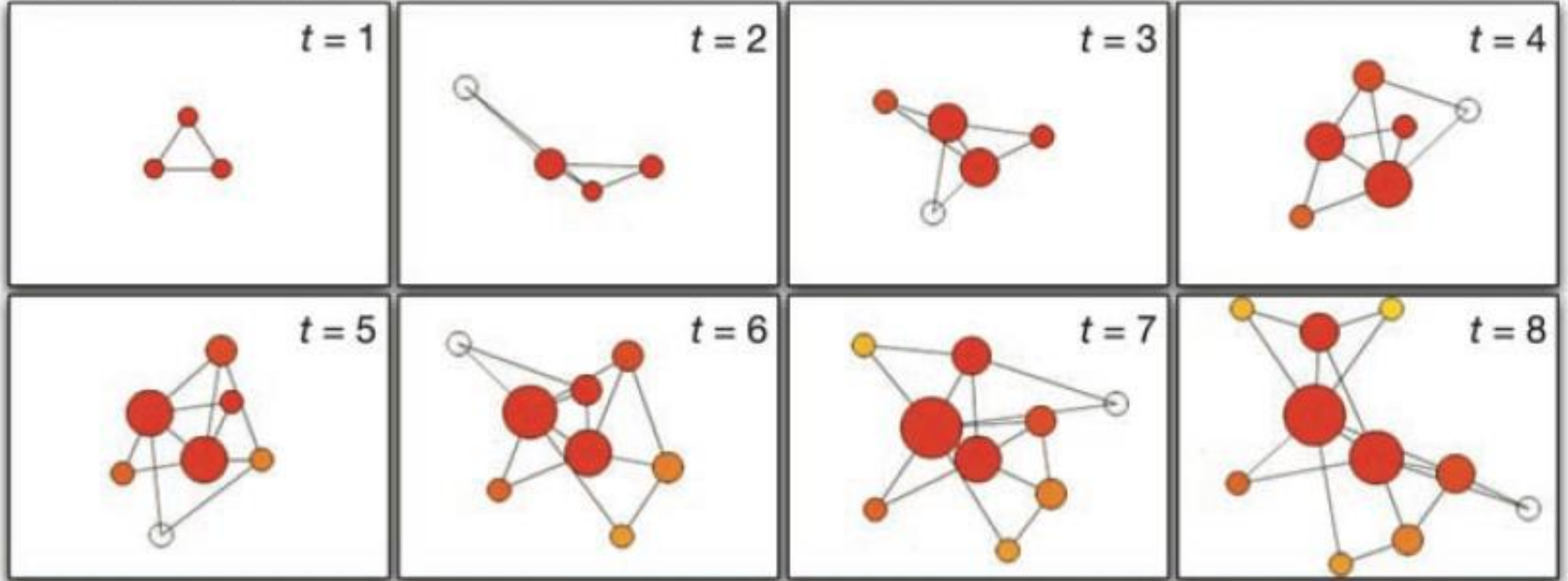
- **Dynamic Growth:** start at $t=0$ with n_0 nodes and $m_0 \geq n_0$ edges
 - **1) Growth**
 - At each time step add a new node with m edges ($m \leq n_0$), connecting to m nodes already in network $k_i(i) = m$
 - **2) Preferential Attachment**
 - The probability of linking to existing node i is proportional to the node degree k_i

$$P(k_i) = \frac{k_i}{\sum_i k_i}$$

- after t timesteps: $t + n_0$ nodes, $mt + m_0$ edges

Barabasi-Albert Model

Scale-Free Model



Barabasi-Albert Model

Let see if the degree of a node $k_i(t)$ at time t i.e. $k_i(t)$ is dependent on initial time.

$$k_i(t + \delta t) = k_i(t) + m\Pi(k_i)\delta t$$

$$\frac{dk_i(t)}{dt} = m\Pi(k_i) = m\frac{k_i}{\sum_i k_i} = \frac{mk_i}{2mt} = \frac{k_i(t)}{2t}$$

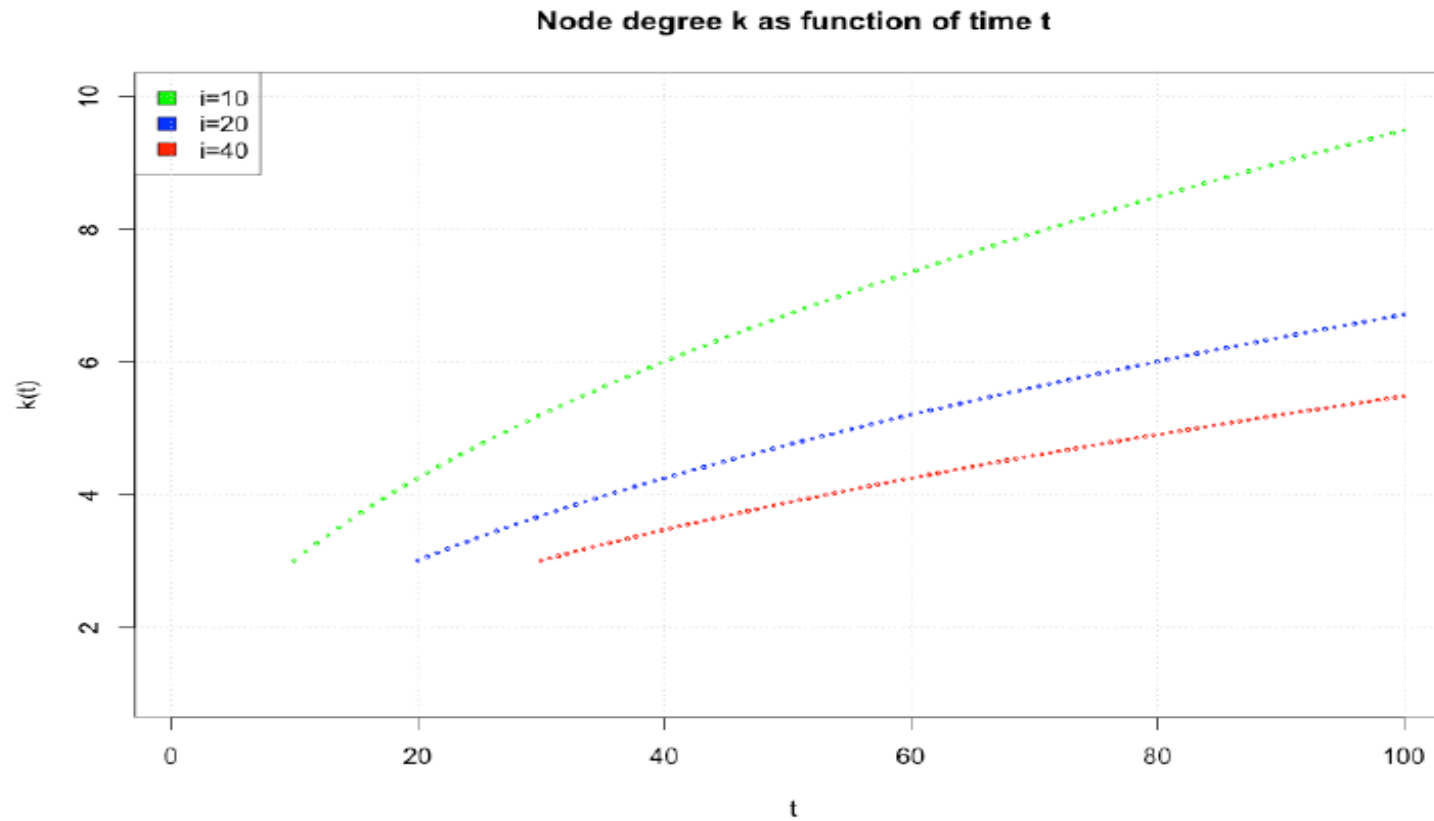
node i is added at time t_i : $k_i(t_i) = m$

$$\int_m^{k_i(t)} \frac{dk_i}{k_i} = \int_{t_i}^t \frac{dt}{2t}$$

Solution:

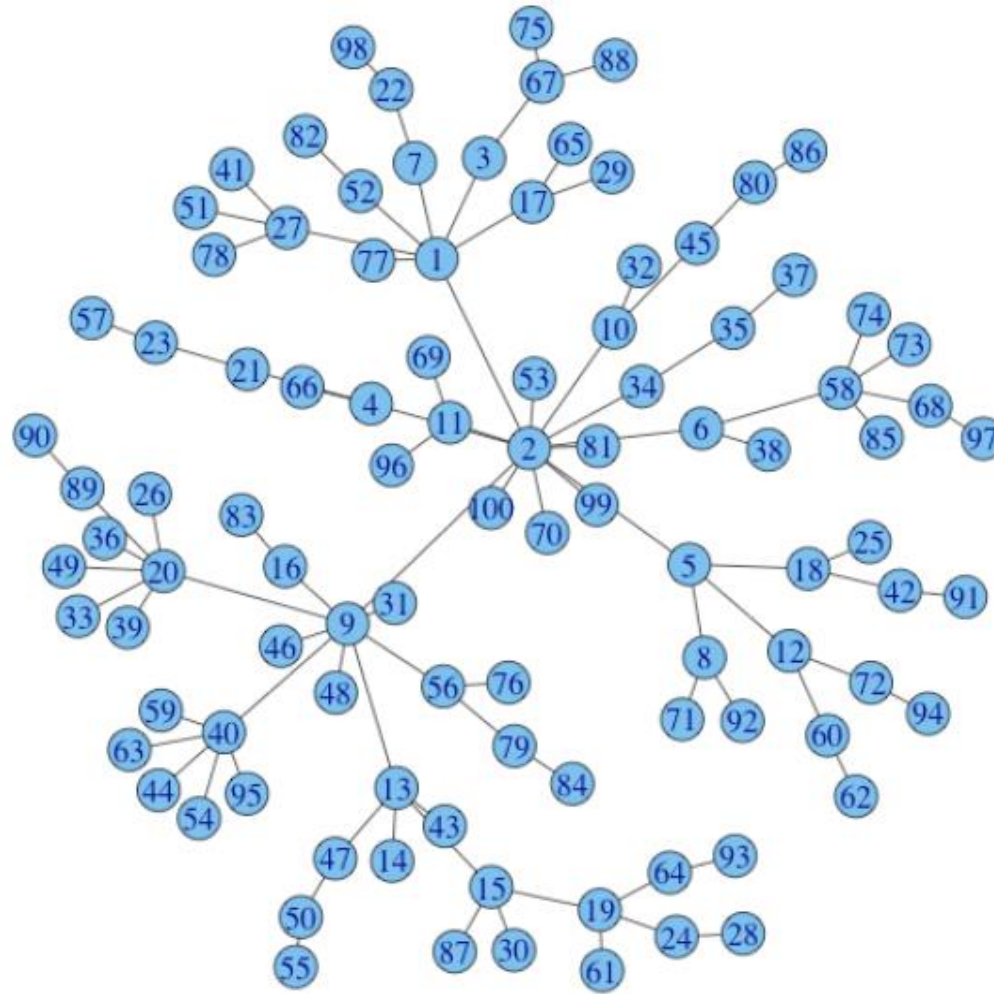
$$k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2}$$

Barabasi-Albert Model



$$k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2} ; \quad \frac{dk_i(t)}{dt} = \frac{m}{2} \frac{1}{\sqrt{tt_i}}$$

Barabasi-Albert Model



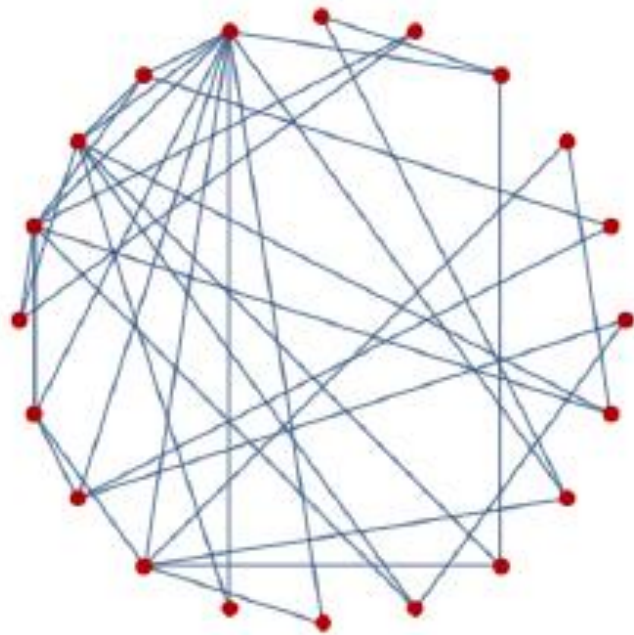
Barabasi-Albert Network Model

- Generates **scale-free** networks as per the following degree distribution (the derivation is skipped)

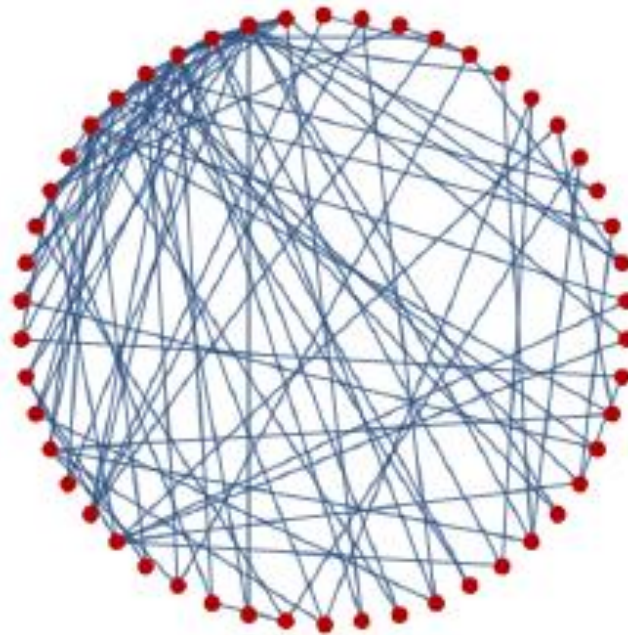
power law vadi property....

$$P(k) = \frac{2m^2}{k^3}$$

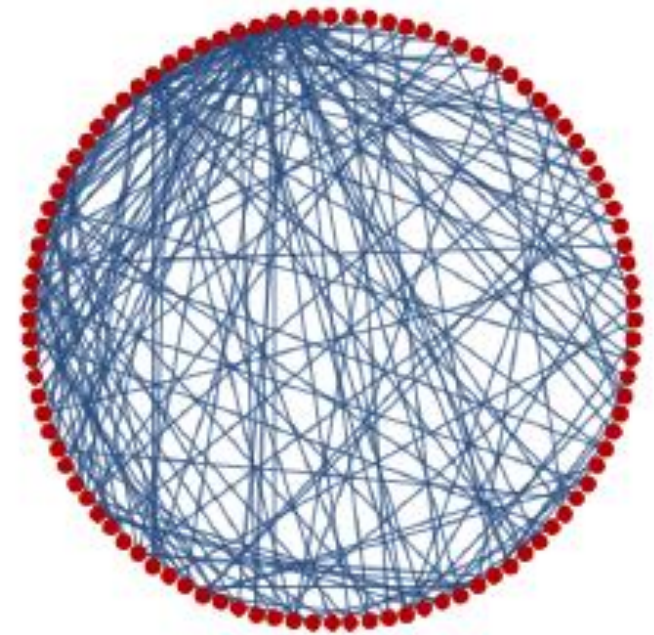
Barabasi-Albert Model: Network Growth



$|G| = 20$
Average Path Length =
2.16



$|G| = 50$
Average Path Length =
2.69



$|G| = 100$
Average Path Length =
3.02