# Social Networks & Power Laws

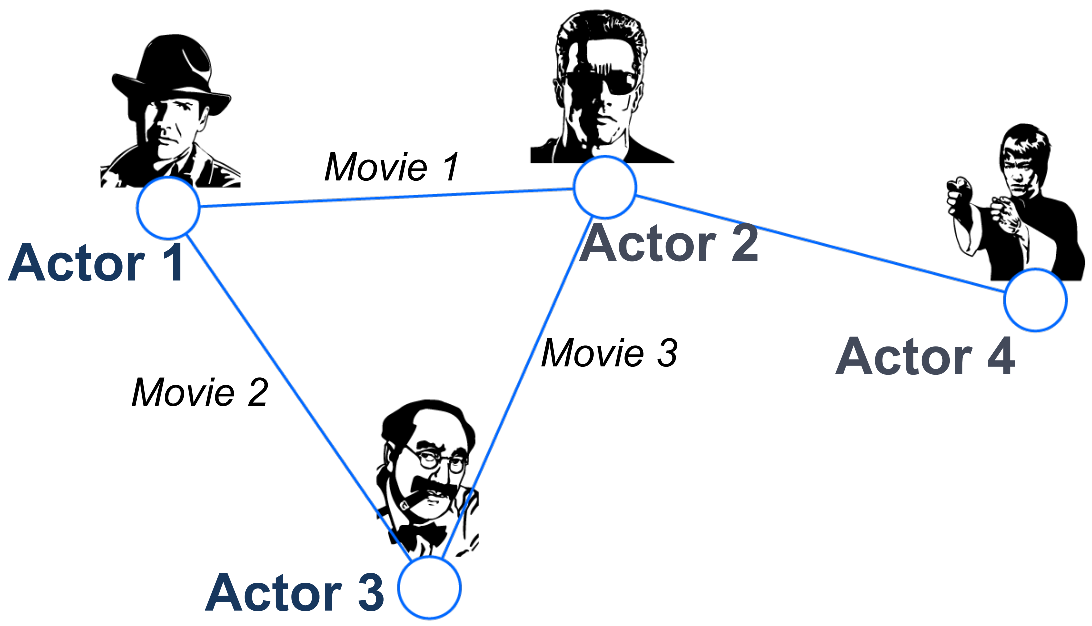
## Social Networks

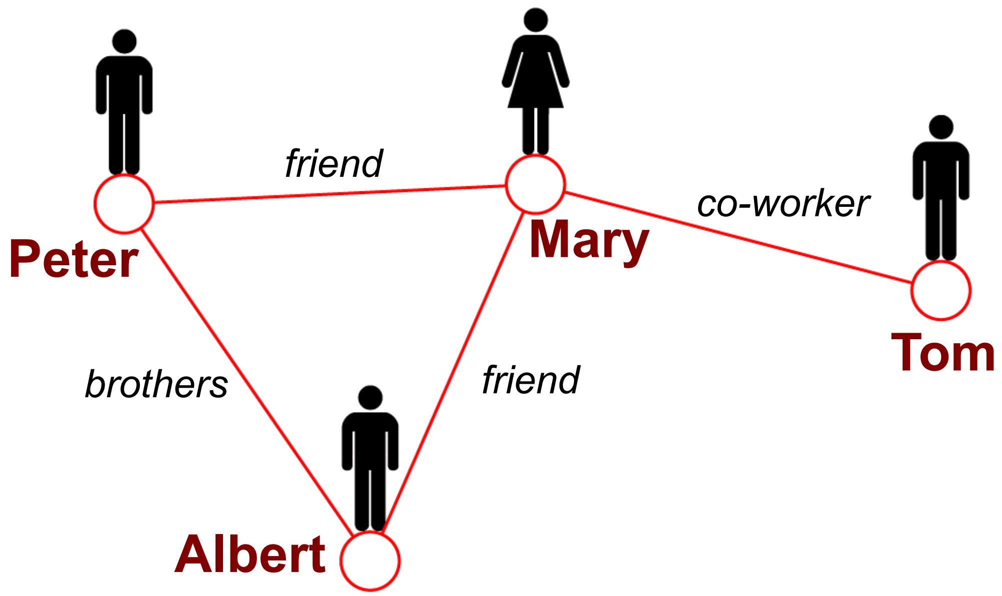
# What is a social network?

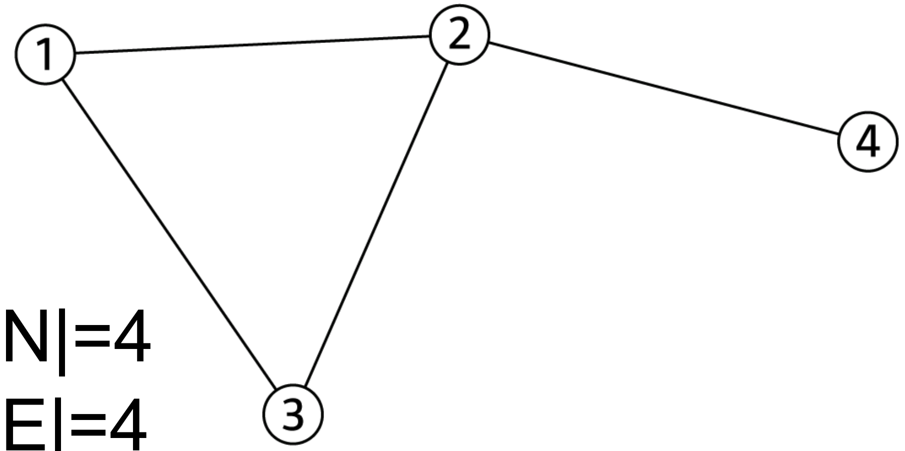
A graph metaphor for studying the relationships/interactions among a group of people

| | |
|---|---|
| People: vertices/nodes | $V$ |
| Relationship: edges | $E$ |
| System: network, graph | $G(V,E)$ |

# 3 Sample Graphs



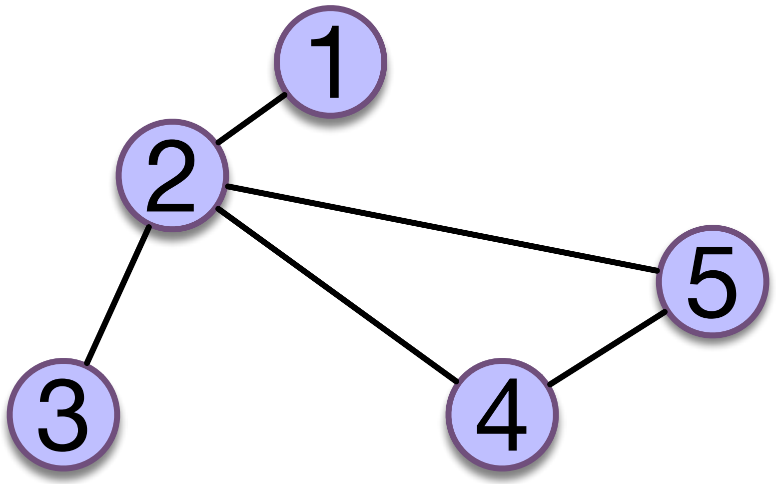




|N|=4
|E|=4

Figures from Jure Leskovec, CS224W

# Networks can be undirected or directed

## Undirected

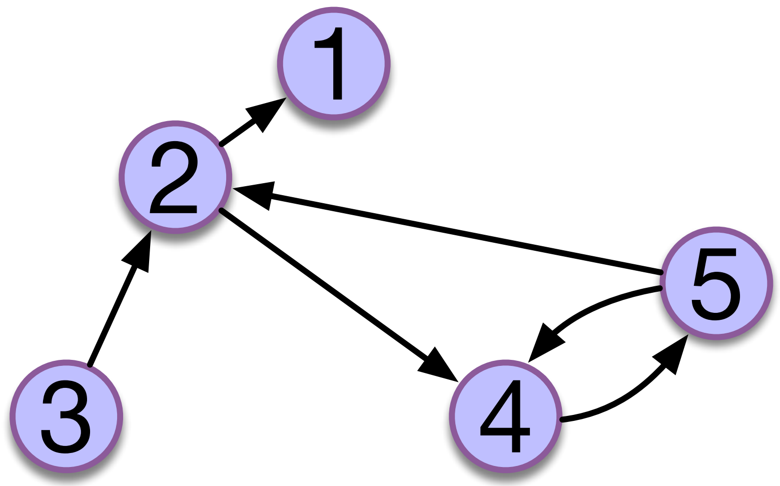



Friendship on Facebook

Actor Collaborations

Twitter conversations

## Directed




Citations

Following on Twitter or IG

# Some useful properties of nodes

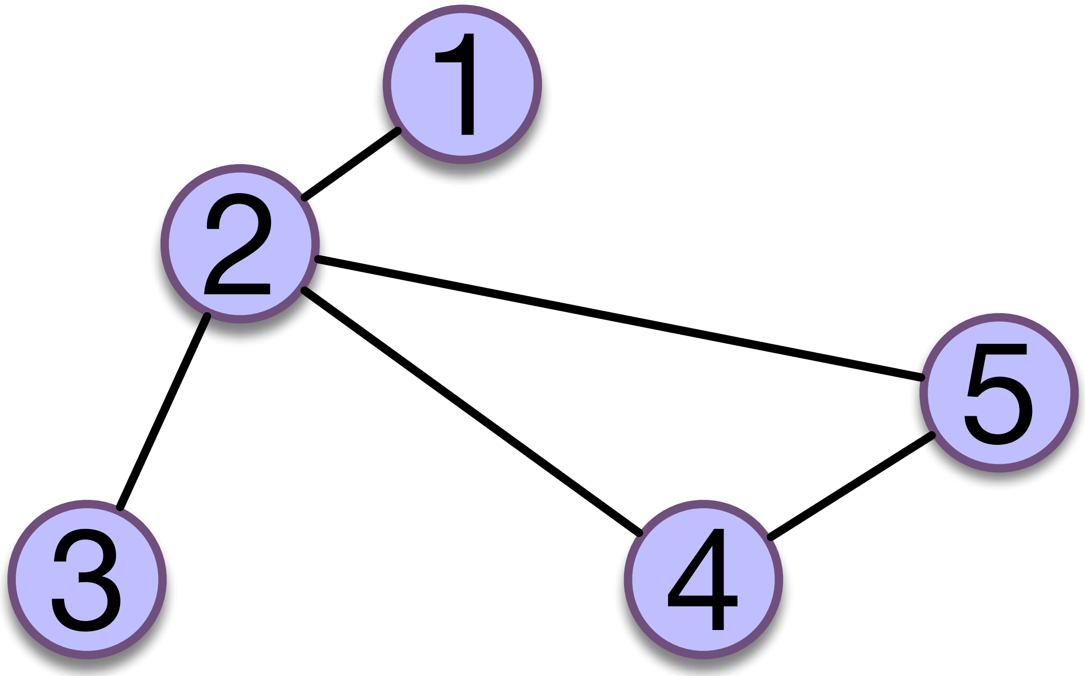
Centrality Measures

- Degree (Undirected, in-degree, out-degree)
- Betweenness Centrality

Clustering Coefficient

# Measuring Centrality: Degree

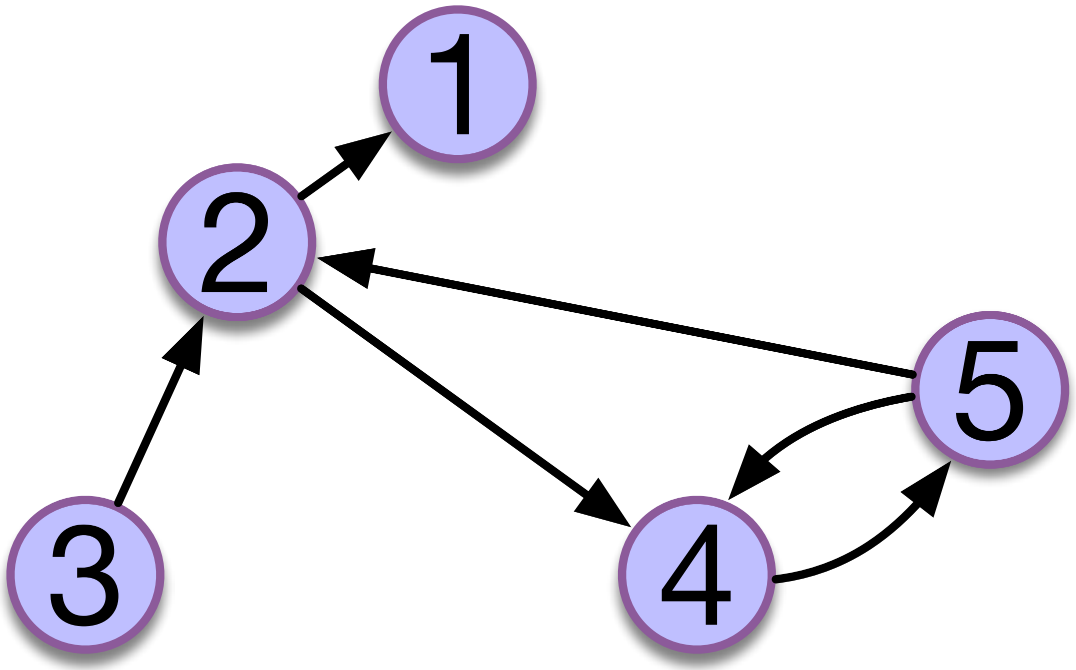
Node degree $k_i$: the number of edges touching node $i$




$k_2=4$ $k_1=k_3=1$ $k_4=k_5=2$

# Measuring Centrality: in-degree

Node in-degree $k^{in}_i$: the number of edges into node $i$



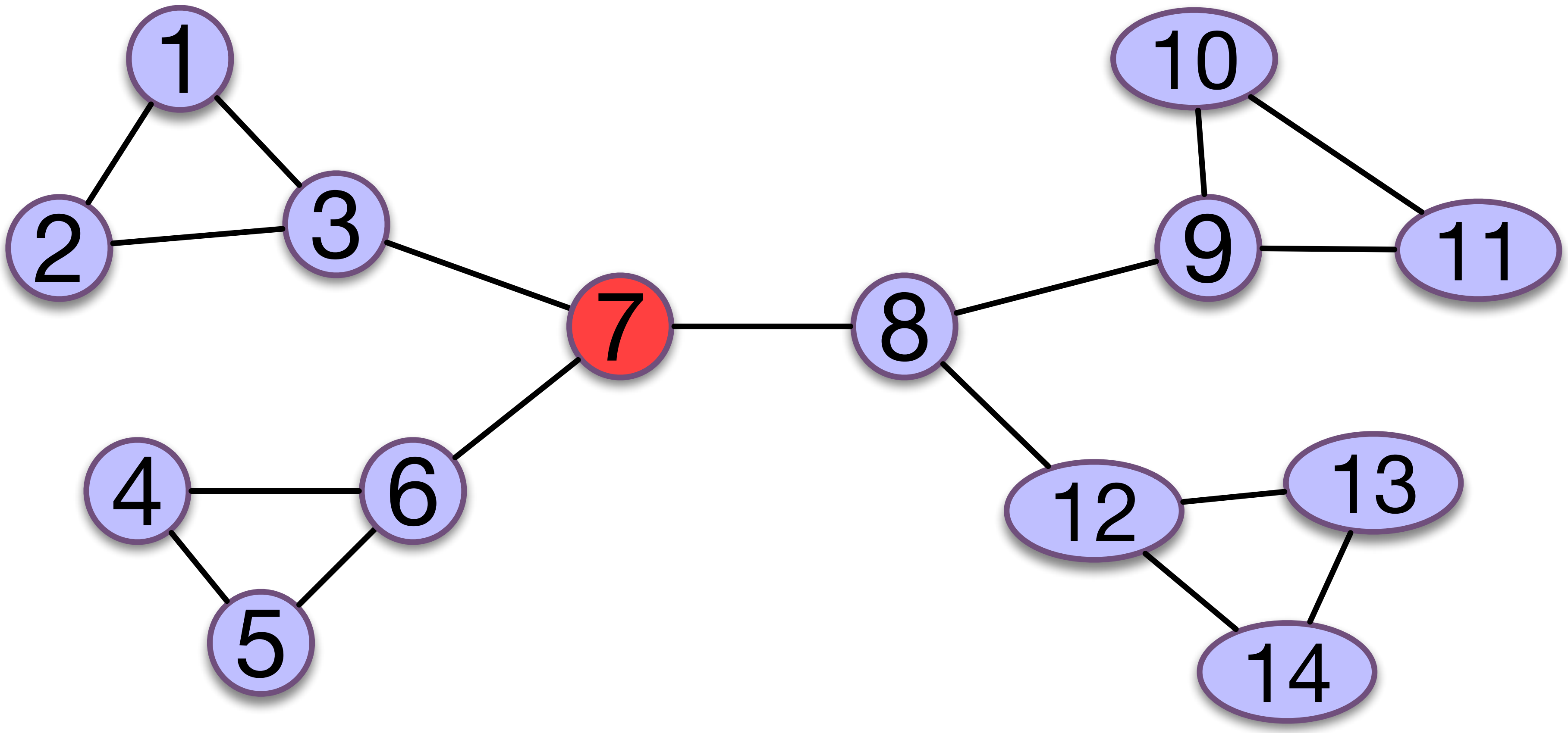

$k^{in}_2 = k^{in}_4 = 2$ $k^{in}_1 = k^{in}_5 = 1$ $k^{in}_3 = 0$

# Betweenness Centrality

A node with high **betweenness**

- lots of shortest paths between nodes must pass through it
- is a bottleneck for information flow in the network

**Betweenness** of node 7 should be high



# Betweenness Centrality

The **betweenness** of a node A (or an edge A-B)=

$$\frac{\text{number of shortest paths that go through A (or A-B)}}{\text{total number of shortest paths that exist between all pairs of nodes}}$$

# Betweenness

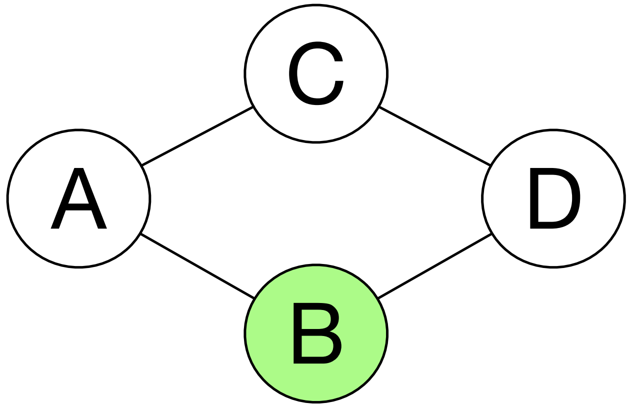
$$b_A = \frac{\text{number of shortest paths that go through A}}{\text{total number of shortest paths between all pairs of nodes}}$$

More formally: $$b_j = \sum_{i=1}^{n}\sum_{k=1}^{n}\frac{s_{ik}(j)}{s_{ik}}, i \neq j \neq k$$

Where: $s_{ik}$ = the # of shortest paths between i and k

$s_{ik}(j)$ = the # of shortest paths between i and k that go through j

Betweenness of B?



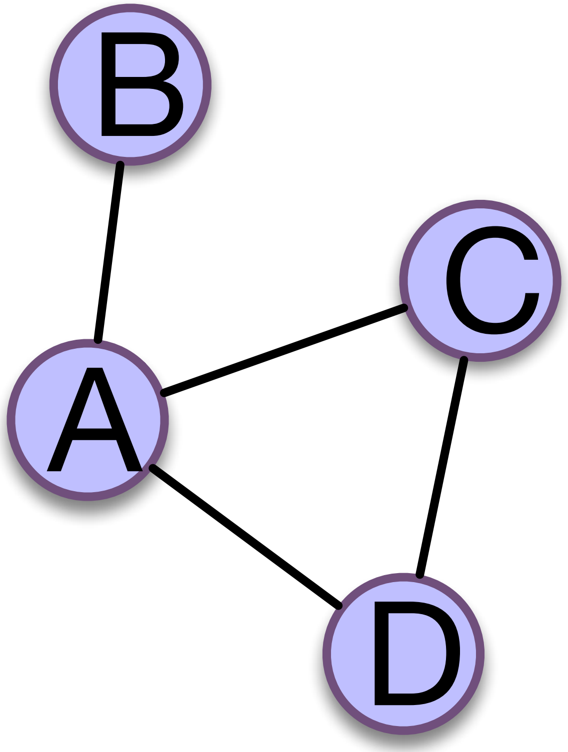

| | # sh. paths through B | # shortest paths |
|---|---|---|
| A,C | 0 | 1 |
| A,D | 1 | 2 |
| C,D | 0 | 1 |

Centrality $b_B$ = 1/4

# Clustering coefficient

$C_A$ for a node A: what % of my friends know each other?

- A has 3 friends (B,C,D), 3 possible links (B-C, B-D, C-D)
- Only C-D exists
- $C_A = 1/3$
- How about $C_C$?
- $C_C=1/1$
- Ranges from 0 (no friends know each other) to 1 (all do)




$$C = \frac{\text{\# of edges between A's friends}}{n_A * (n_A - 1)/2}$$

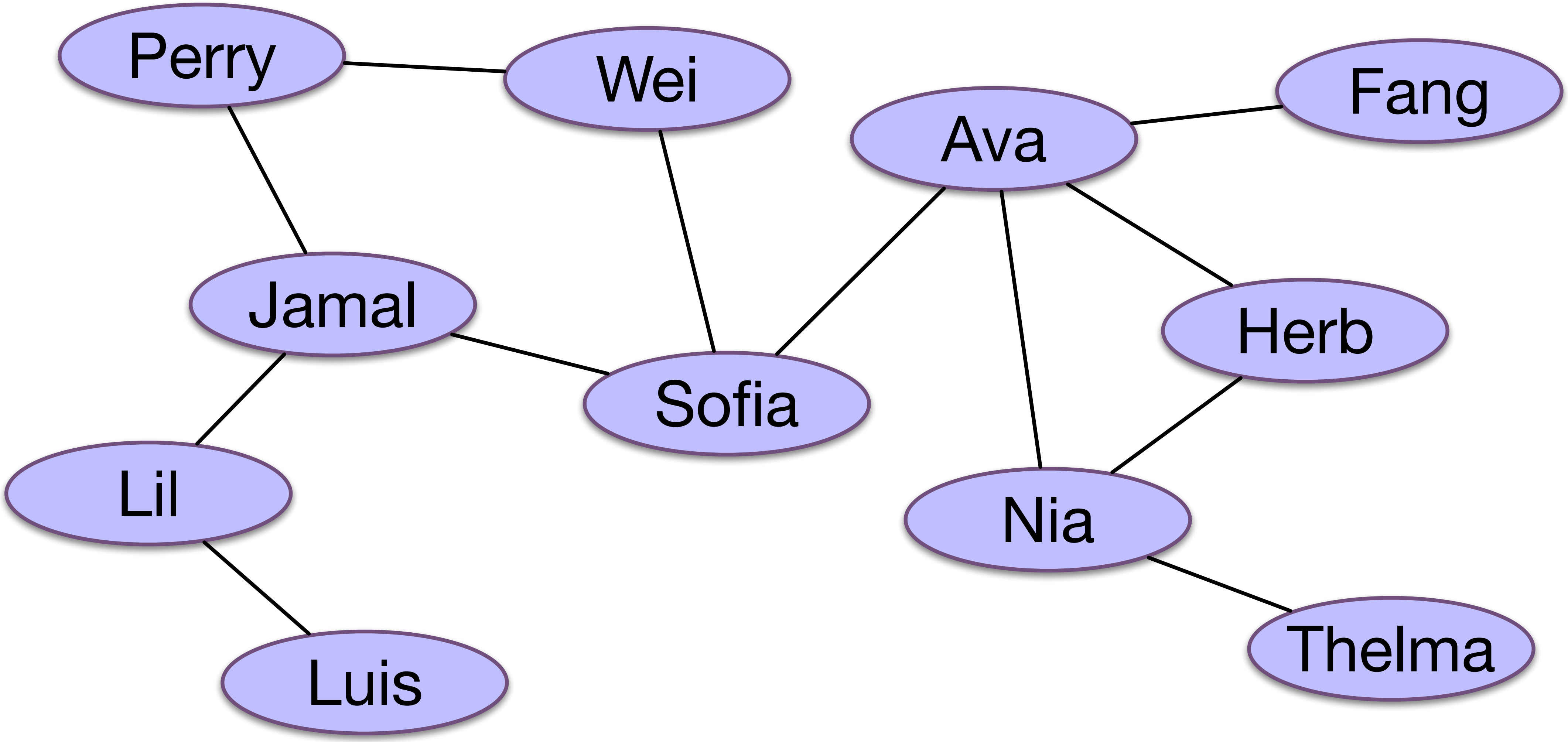
# Social Networks & Power Laws

## Social Networks

# Social Networks & Power Laws

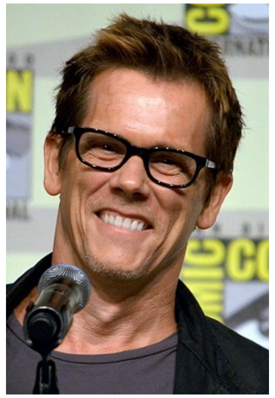
## Small Worlds

# "6 Degrees of Separation"?



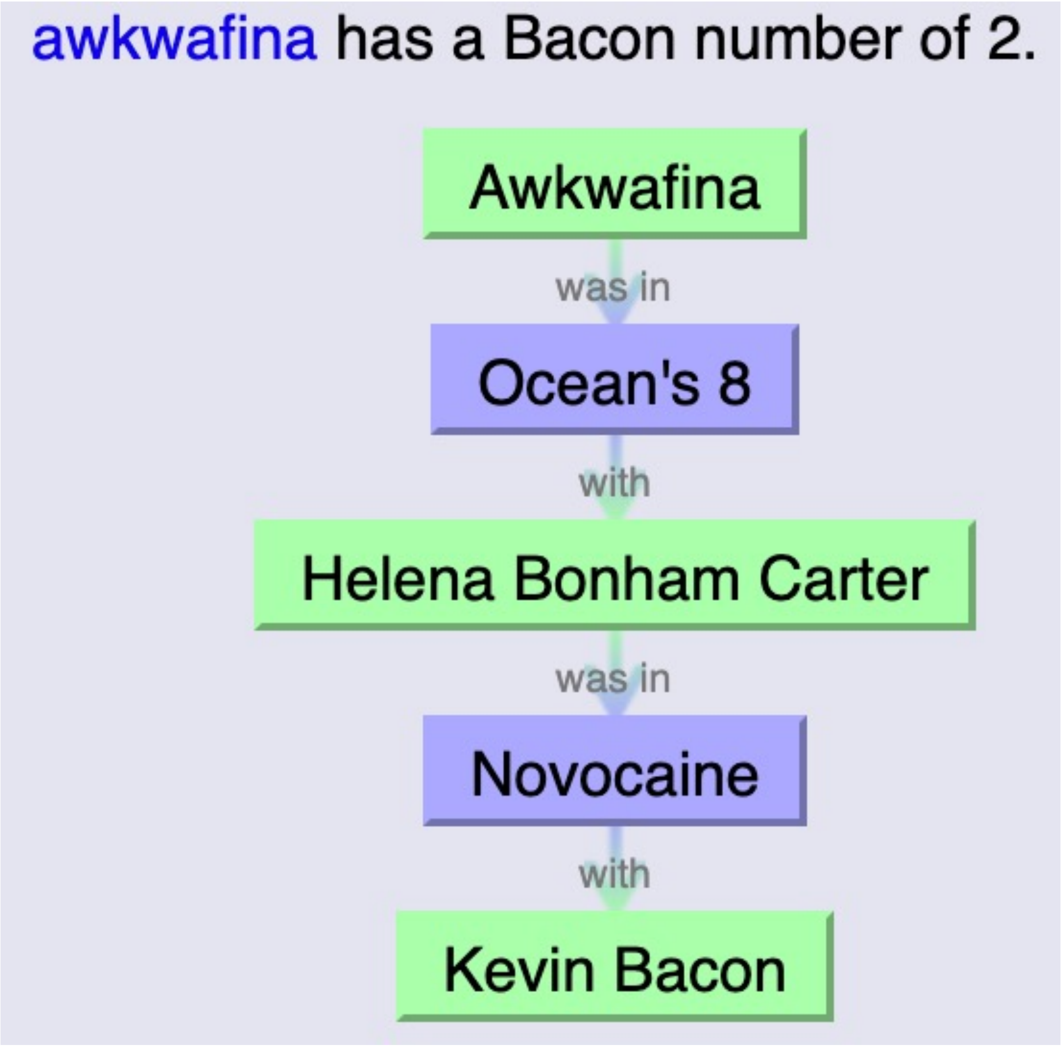

Small world: people are connected by small # of links

# The Kevin Bacon number

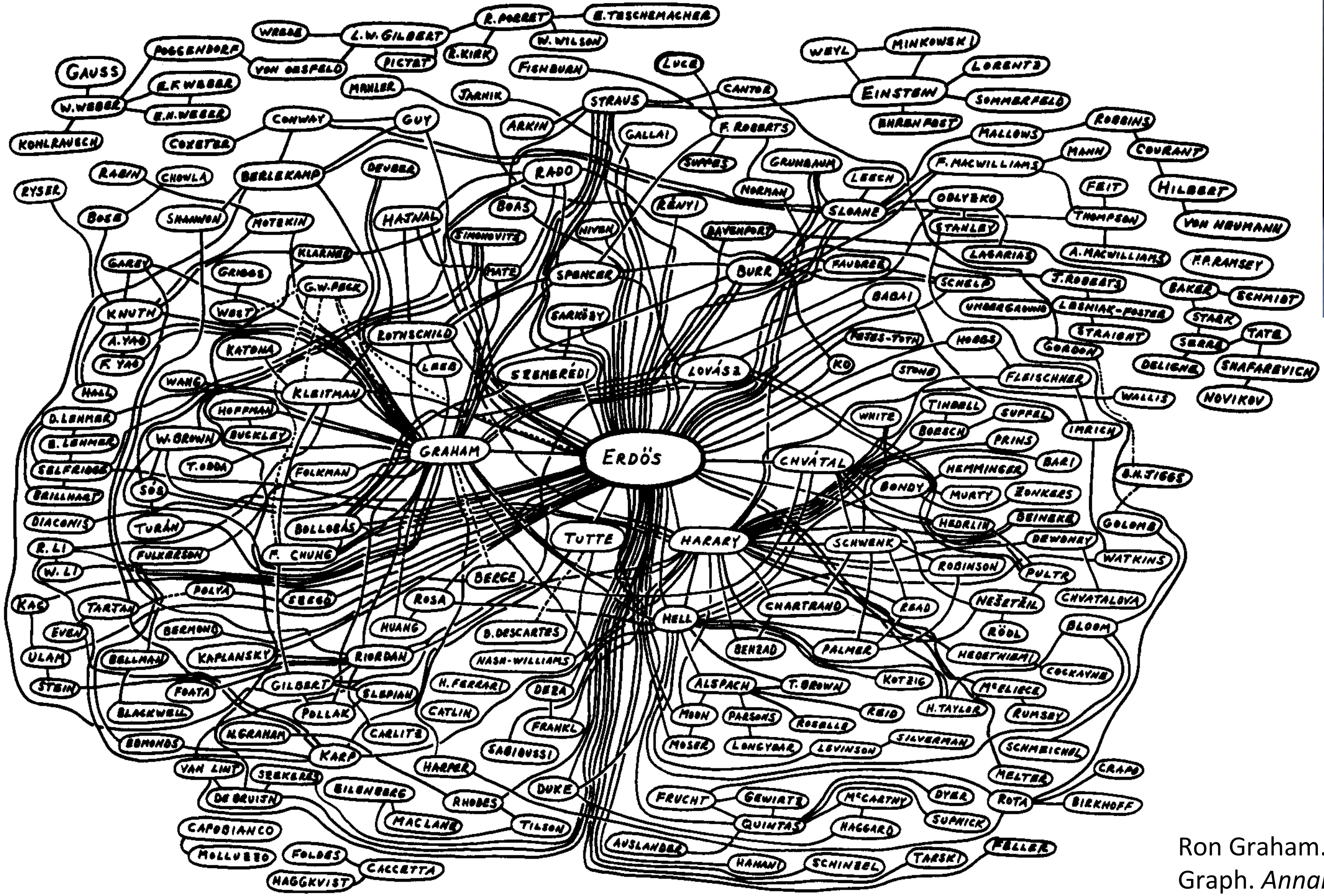


Genevieve, Wikimedia Commons

- Create a network of movie actors
- Connect 2 actors if they're in the same movie
- Count the number of steps to Kevin Bacon
- Almost all actors have Bacon numbers <= 4



# Erdös numbers are small too

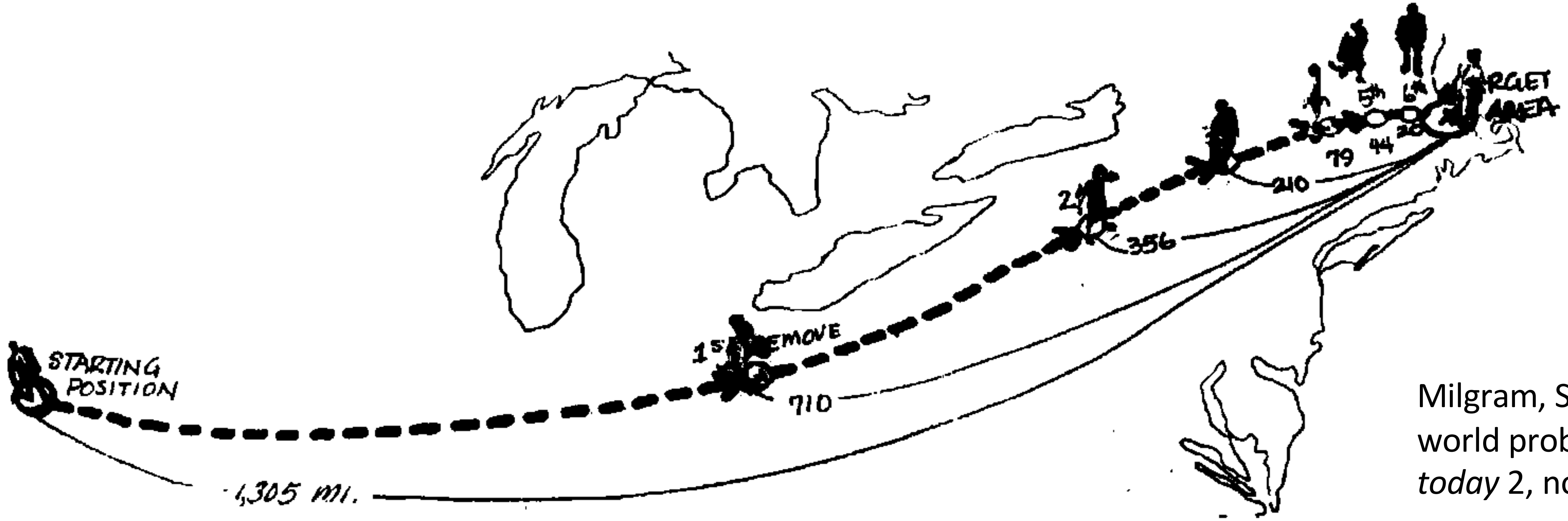



[Topsy Kretts](#)
[Wikimedia Commons](#)

Ron Graham. 1979. On Properties of a Well-Known Graph. *Annals N.Y. Acad. Sci.* **328** (1979), 166-172.

# The Milgram Small World Experiment

## How to find the typical shortest path between any two people?




Milgram, Stanley. "The small world problem." *Psychology today* 2, no. 1 (1967): 60-67.

- Travers and Milgram (1969): Asked 300 people in Omaha NE and Boston
    - Get a letter to a stock-broker in Boston by passing it through friends
- How many steps did it take?

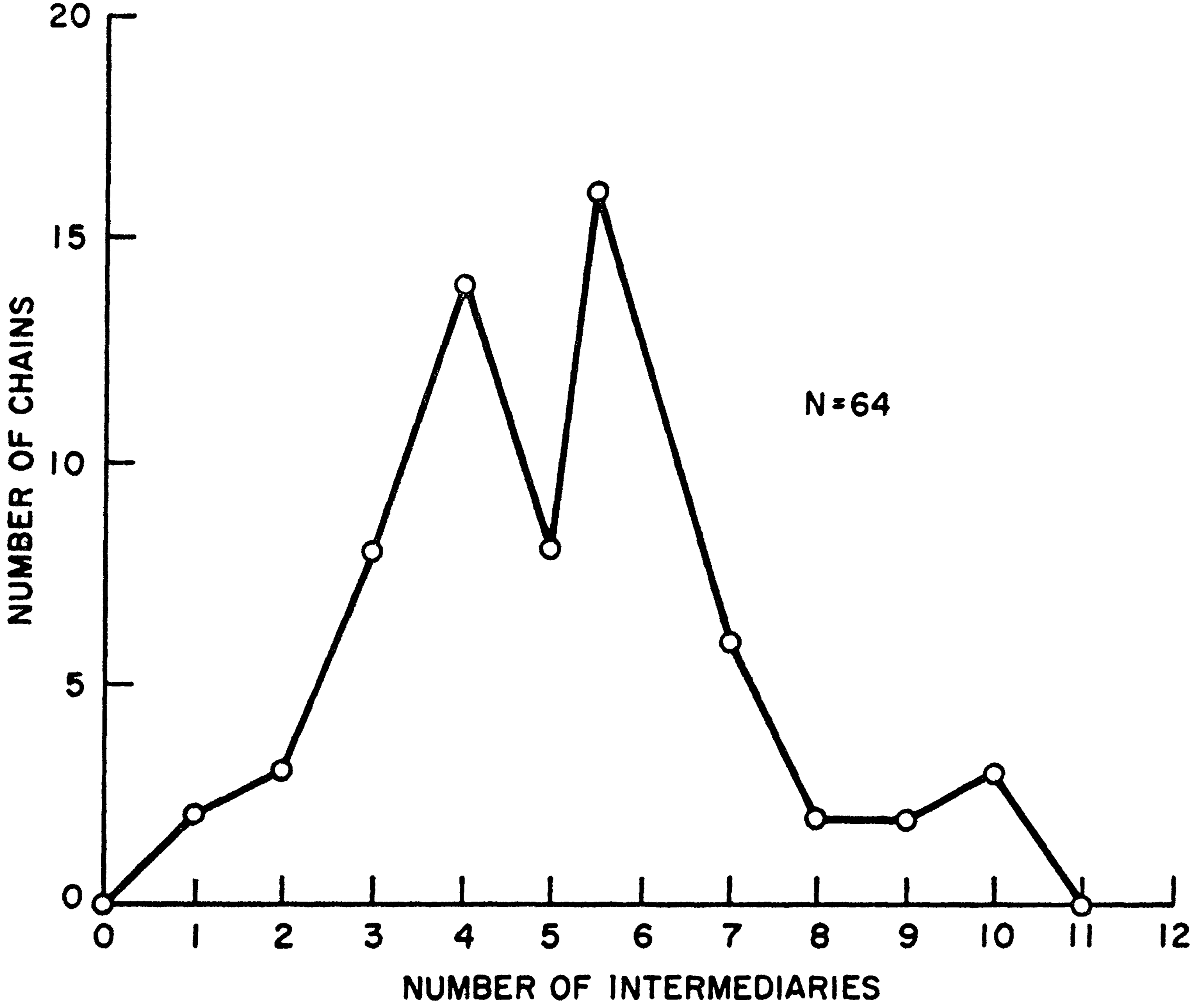
Travers, J., & Milgram, S. (1969). An exploratory study of the small world problem. *Sociometry*, *32*, 425-43.

# The Milgram small world experiments

It took 4.4-5.7 intermediaries on average

**"Six degrees of separation"**

Most letters passed through 2 "hub" people





Travers, J., & Milgram, S. (1969). An exploratory study of the small world problem. *Sociometry*, *32*, 425-43.

# Problems with Milgram experiment

Not enough data

- Only 300 observations to start
- Of those, only 64 letters made it through

Almost certainly an overestimate

- People may not find actual shortest path

# Online small world: Facebook friends

721 million users
69 billion friendship links


Paul Butler, Facebook

99.6% of all pairs of users connected by paths of 5 degrees (6 hops)

92% are connected by only four degrees (5 hops).

Average: 3.74 intermediaries/degrees (4.74 hops)

Four degrees of separation

Backstrom, Lars, Paolo Boldi, Marco Rosa, Johan Ugander, and Sebastiano Vigna. "Four degrees of separation." In *Proceedings of the 4th Annual ACM Web Science Conference*, pp. 33-42. 2012.

# Fun facts: Origins of the “Six degrees” hypothesis

Hungarian writer Karinthy’s 1929 play “Chains” (Láncszemek)

- https://djjr-courses.wdfiles.com/local--files/soc180%3Akarinthy-chain-links/Karinthy-Chain-Links_1929.pdf

# Why is the world small? What are the implications?

Duncan Watts: Small world is interesting because:

1. The network is **large** - *O*(Billions)
2. The network is **sparse** - people are connected to a small fraction of the total network
3. The network is **decentralized** -- no single (or small #) of stars
4. The network is highly **clustered** -- most friendship circles are overlapping

Slide from James Moody

# Watts and Strogatz (1998) Model

A small world graph: very clustered (high C) but small path-length (low L)

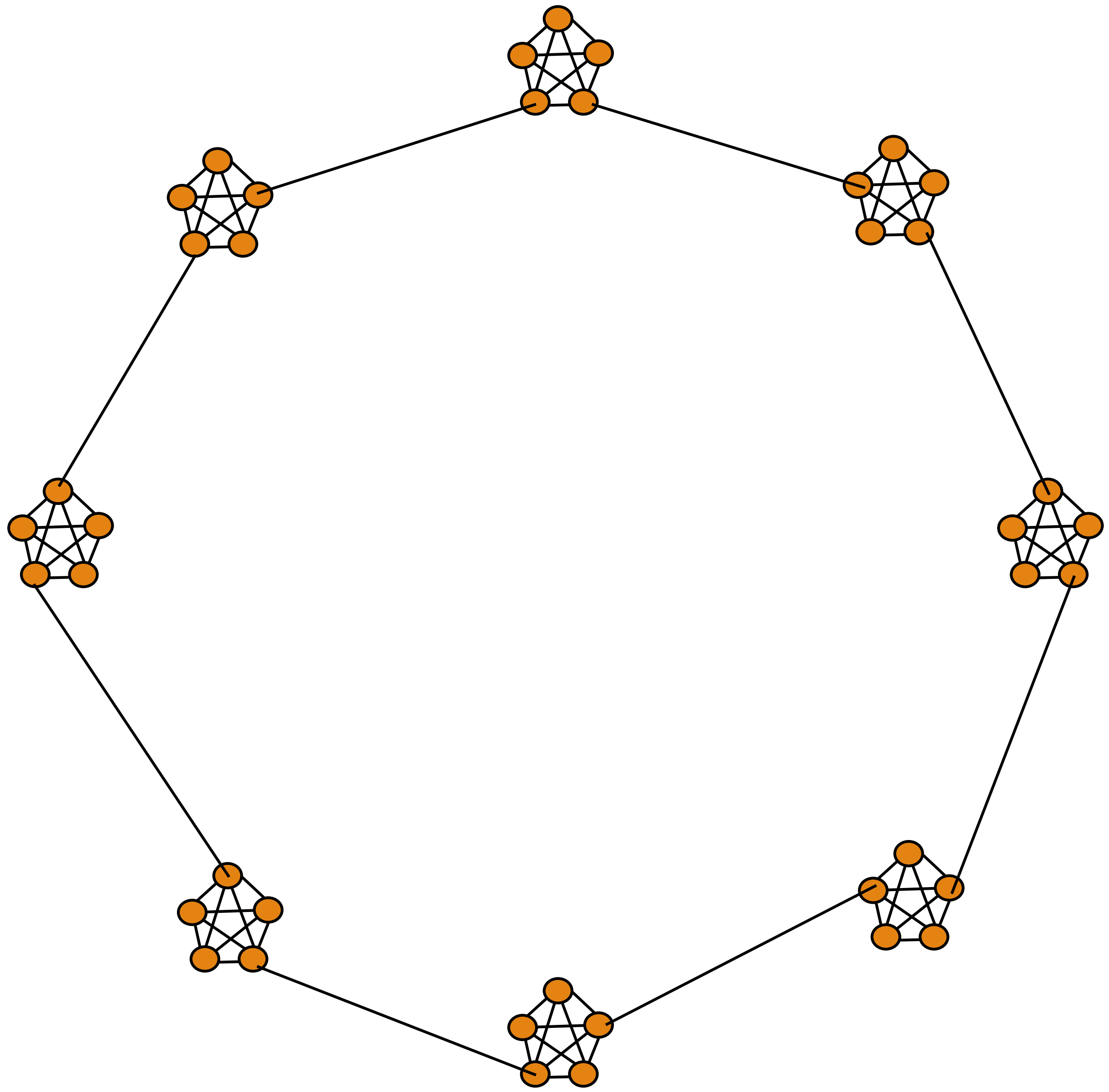
1) **Clustering coefficient C**
   - How many of my friends know each other
   - Averaged over all nodes: the average local density
2) **Characteristic path length *L***
   - The average length of the shortest paths connecting any two nodes.

*(Note: this is not quite the same as the **diameter** of the graph, which is the **maximum** shortest path connecting any two nodes)*

Watts, D., Strogatz, S. Collective dynamics of 'small-world' networks. *Nature* **393,** 440–442 (1998).

# Watts and Strogatz (1998) "Caveman Network"

- Everyone in a cave knows each other
- A few people make connections
- Are C and L high or low?
- C high, L high

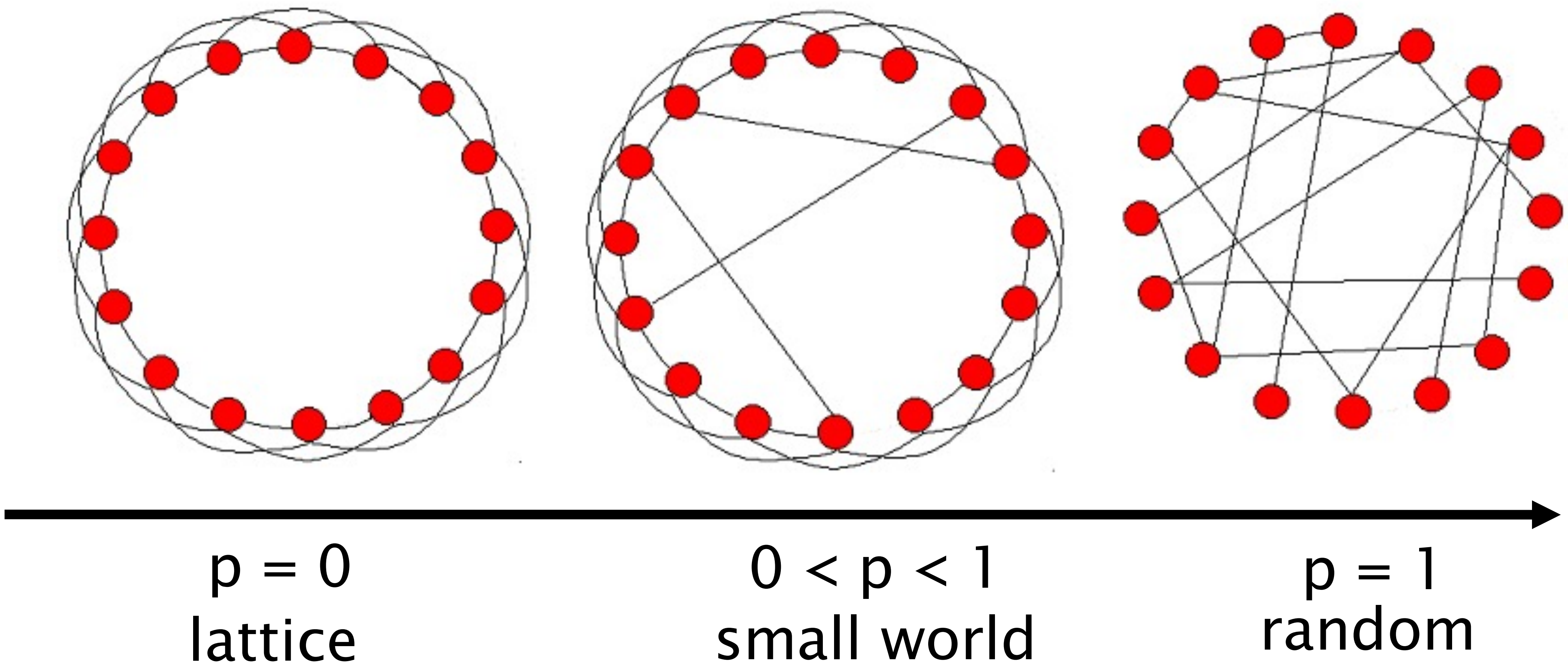


Slide from Lada Adamic

# Watts and Strogatz 1998 model

Start with a ring, where every node is connected to the next z nodes ( a regular lattice)

With probability p, rewire every edge (or, add a shortcut) to a random, uniformly chosen destination.




Slide from
Lada Adamic

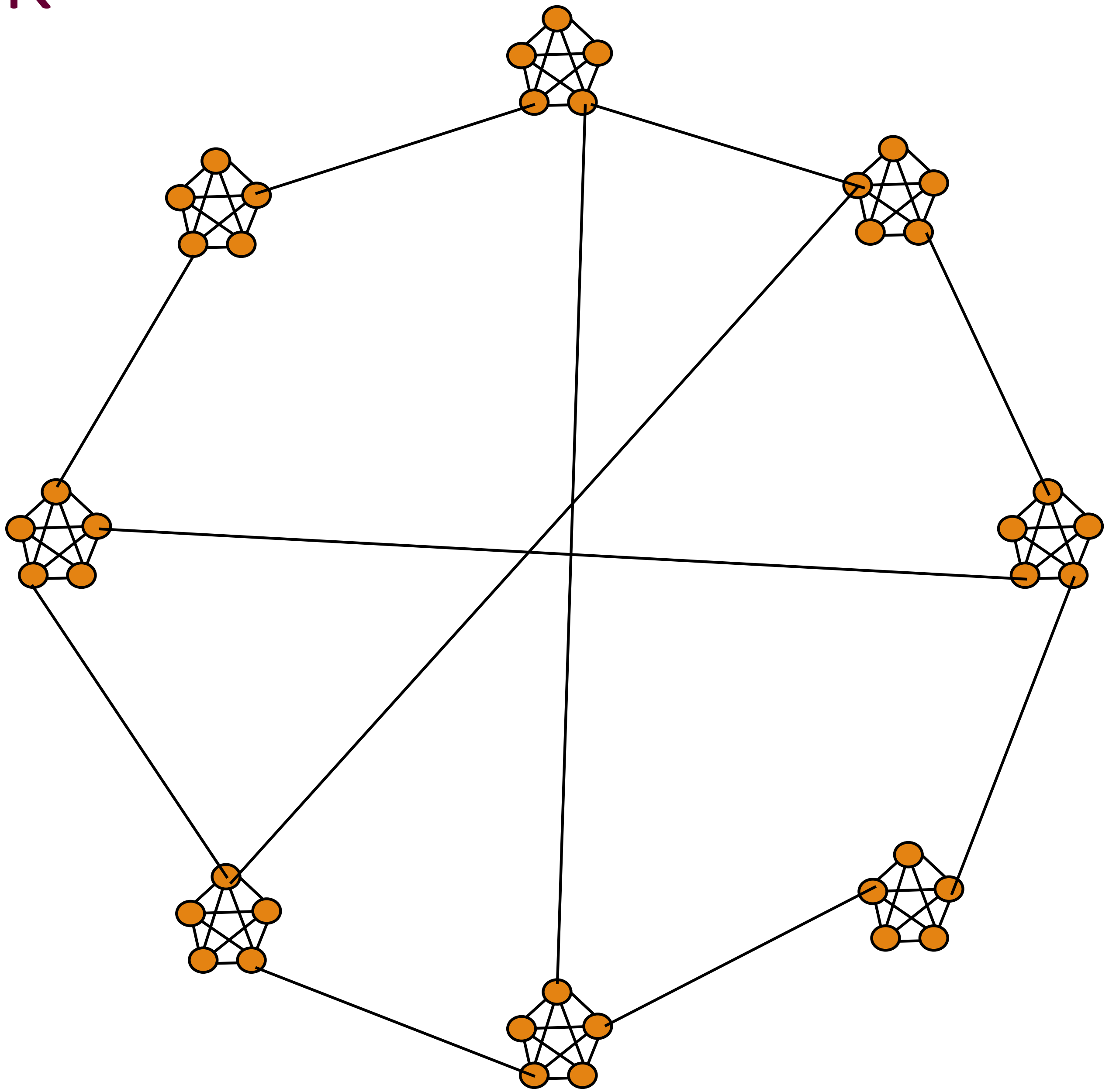
# Why does this work? Key is fraction of shortcuts in the network

In a highly clustered, ordered network, a single random connection will create a shortcut that lowers *L* dramatically

Without big changes in C

Small world: very clustered (high C) but small path-length (low L)

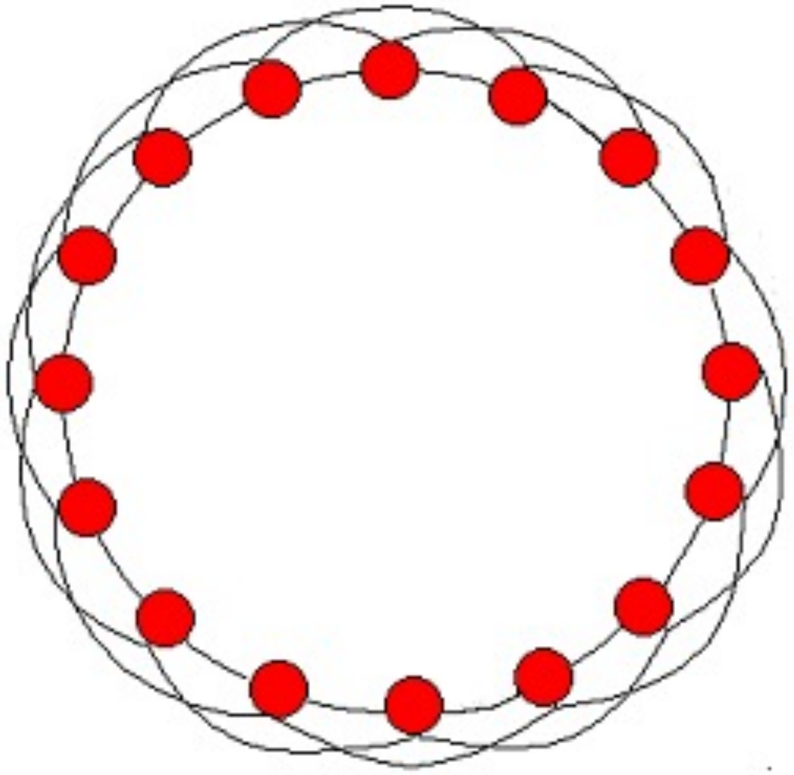
Small world properties can be created by a small number of shortcuts



Slide from Lada Adamic

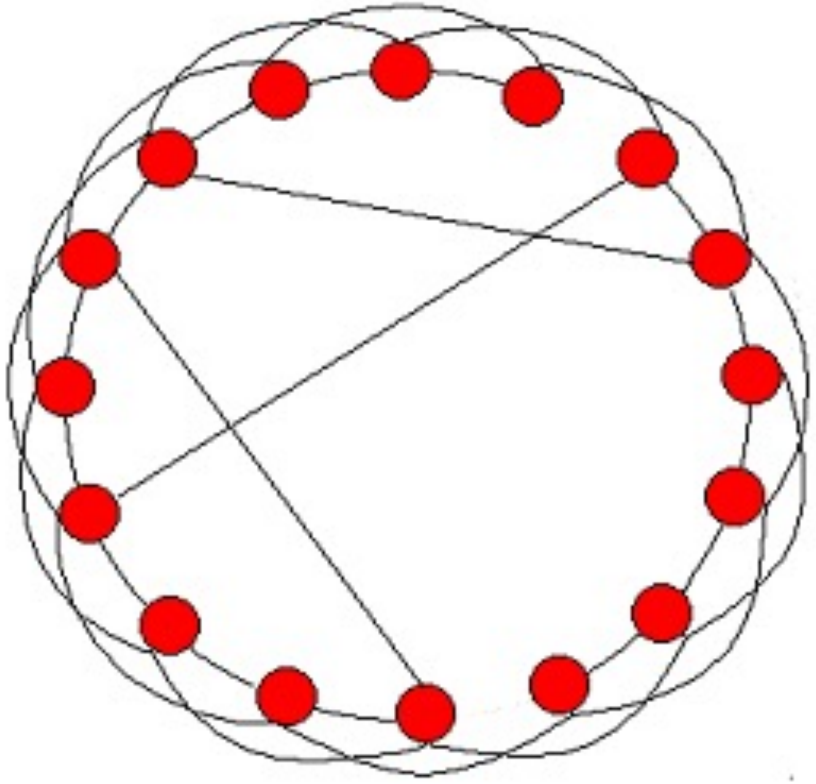
# Clustering and Path length

Lattice



High C
High L

Small world



High C
Low L

Figures from
Lada Adamic

# Small World: Summary

Can a network with high clustering be a small world?

- Yes! You don't need more than a few random links

The Watts Strogatz Model:

- Provides insight on the interplay between clustering and the small-world
- Captures the structure of many realistic networks
- Accounts for the high clustering of real networks

Slide from Jure Leskovec

# Social Networks & Power Laws

## Small Worlds

# Social Networks & Power Laws

## Weak Links

# Weak links

In the 1960s, Mark Granovetter  studied how people find jobs.

He found out that most job referrals were through personal contacts.

But more by acquaintances and not close friends.

Why didn't jobs come from close friends?

Mark Granovetter. The strength of weak ties. American Journal of Sociology, 78:1360– 1380, 1973.

# Triadic Closure

"If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future." (Anatol Rapoport 1953)

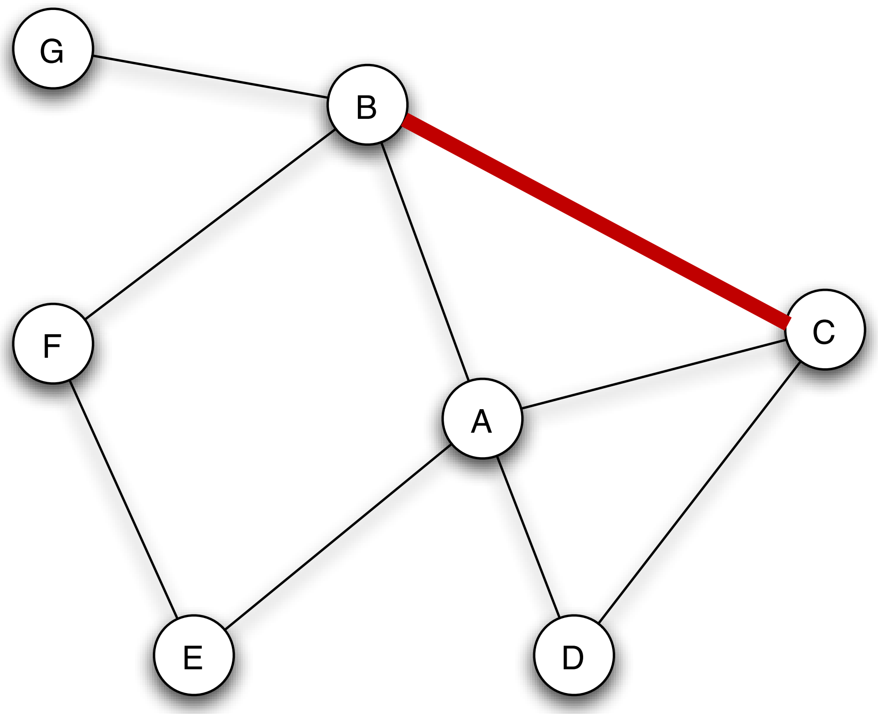


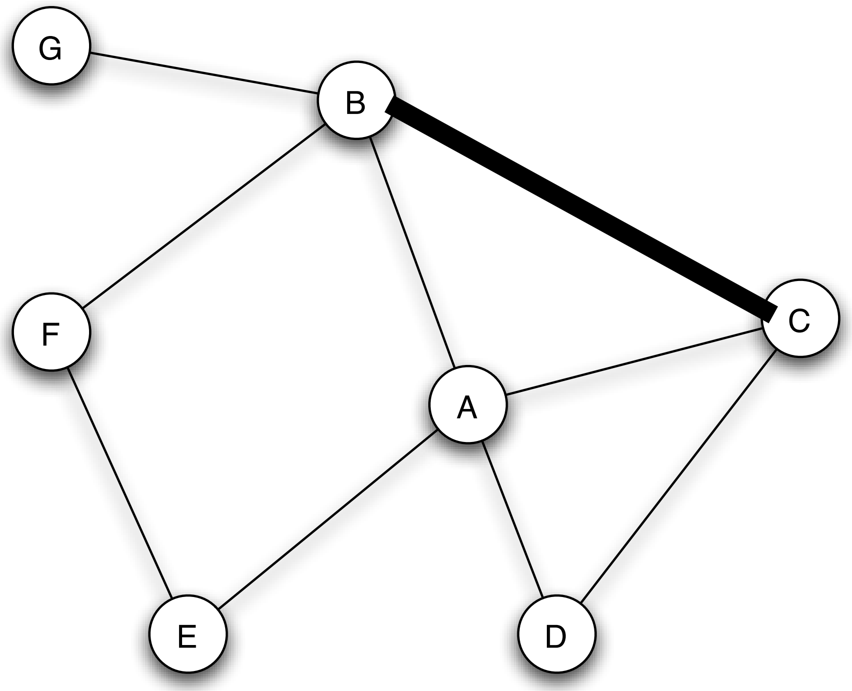

Figure from D. Easley and J.Kleinberg. 2010. Networks, Crowds, and Markets. Cambridge

# Reminder: clustering coefficient C

$C_A$ of a node A is the percentage of friends of A who are friends themselves

$C_A$ before new edge = 1/6

(of B-C, B-D, B-E, **C-D**, C-E, D-E)




$C_A$ after new edge? = 2/6

Triadic closure leads to higher clustering coefficients

Figure from D. Easley and J.Kleinberg. 2010. Networks, Crowds, and Markets. Cambridge

# Why Triadic Closure?

1. We meet our friends through other friends
    - B and C have **opportunity** to meet through A
2. B and C’s mutual friendship with A gives them a reason to **trust** A
3. A has incentive to bring B and C together to avoid **stress**:
    - if A is friends with two people who don’t like each other it causes stress

# An aside about the importance of triangles

Bearman, Peter S., and James Moody. "Suicide and friendships among American adolescents." *American Journal of Public Health* 94, no. 1 (2004): 89-95.

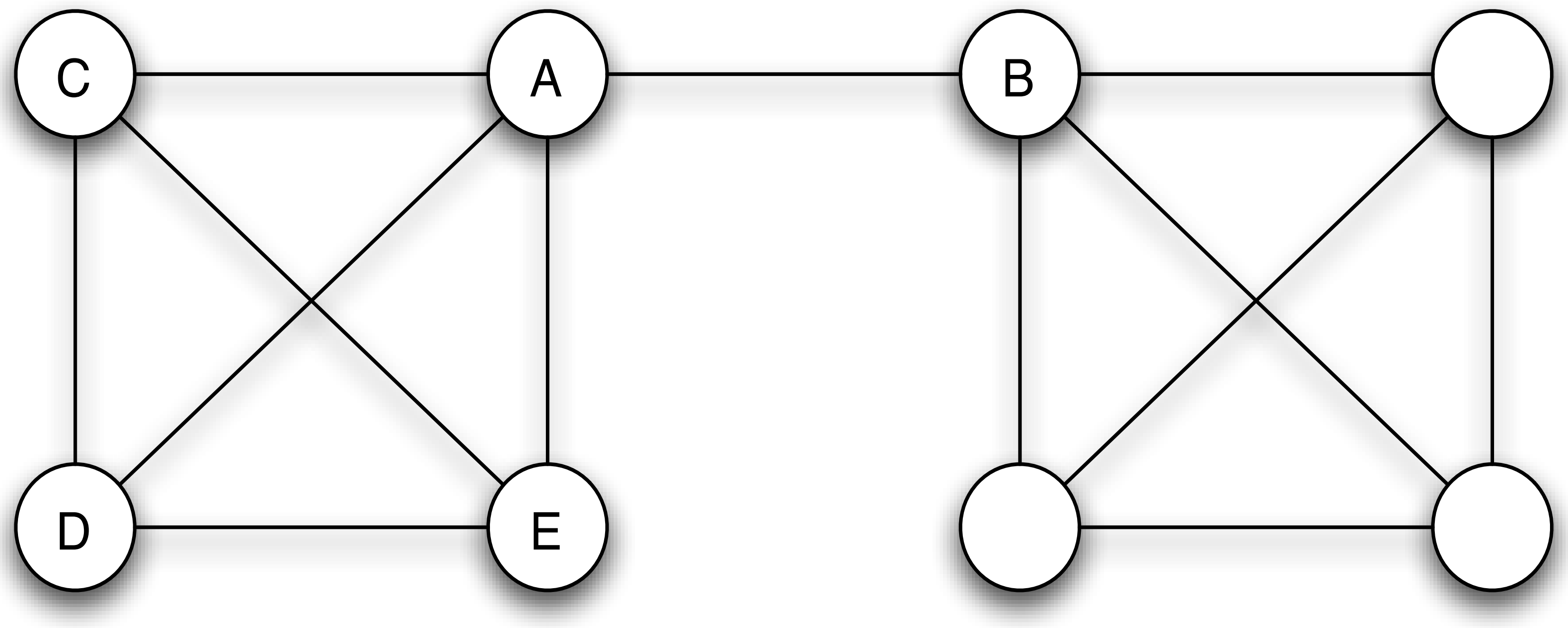
Famous, replicated Bearman and Moody result

Adolescents with low clustering coefficients in their friend network more likely to consider suicide.

Lack of triangles correlates with lack of social groups

# Bridges

A bridge is an edge whose removal places A and B in different components

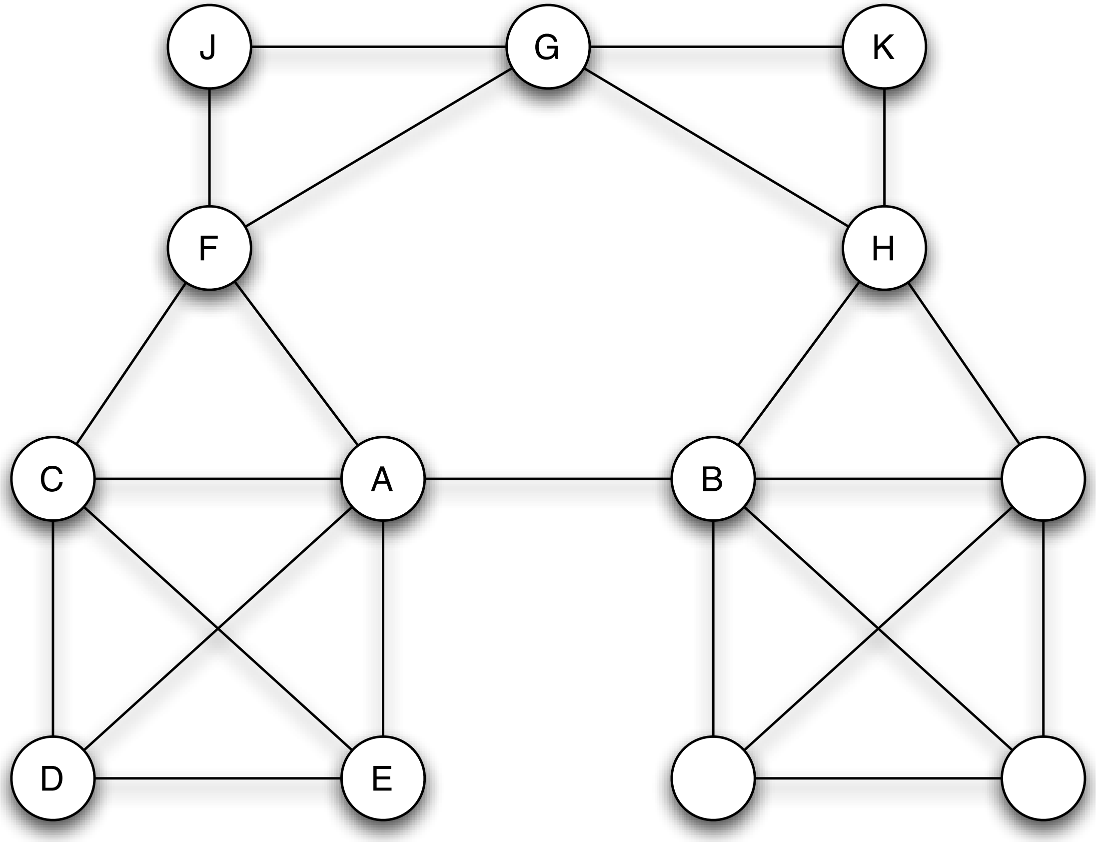



If A is going to get new information (like a job) that she doesn't already know about, it might come from B

Figure from D. Easley and J.Kleinberg. 2010. Networks, Crowds, and Markets. Cambridge

# Local Bridge

A local bridge is an edge X-Y where X,Y have no friends in common
(a local bridge does not form the side of any triangle)




If A is going to get new information (like a job) that she doesn't already know about, it might come from B

Figure from D. Easley and J.Kleinberg. 2010. Networks, Crowds, and Markets. Cambridge

# Strong and Weak Ties

Some ties are stronger

- closer friendships
- more time spent together

Simplifying assumption

- ties are either strong (s) or weak (w)

# Strong ties and triadic closure

The new B-C edge more likely to form if A-B and A-C are **strong** ties

More extreme: if A has strong ties to B and to C, there **must** be an edge B-C

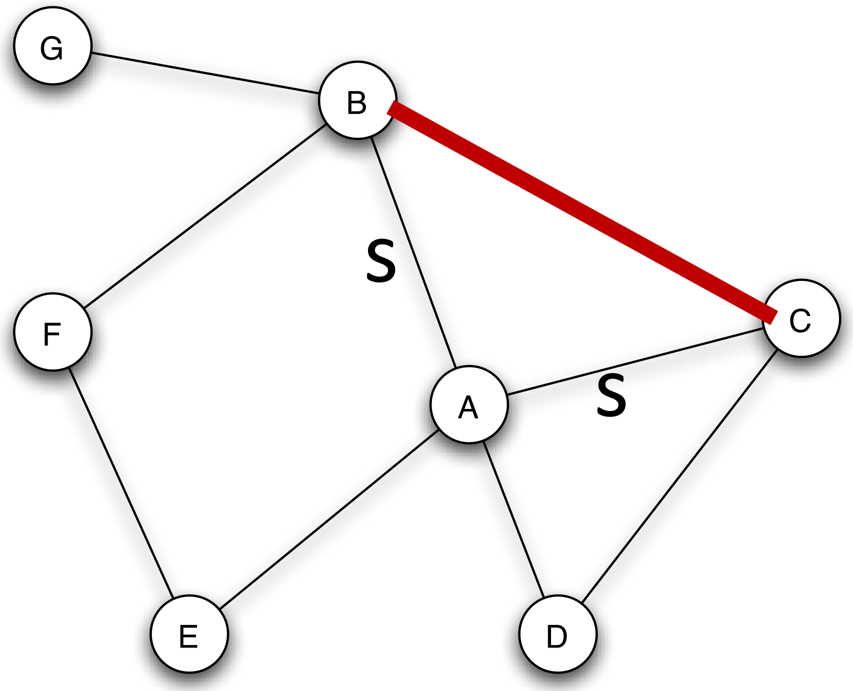



Figure from D. Easley and J.Kleinberg. 2010. Networks, Crowds, and Markets. Cambridge

# Strong triadic closure

If a node 1 has two strong ties to nodes 2 and 3, there is an edge between 2 and 3

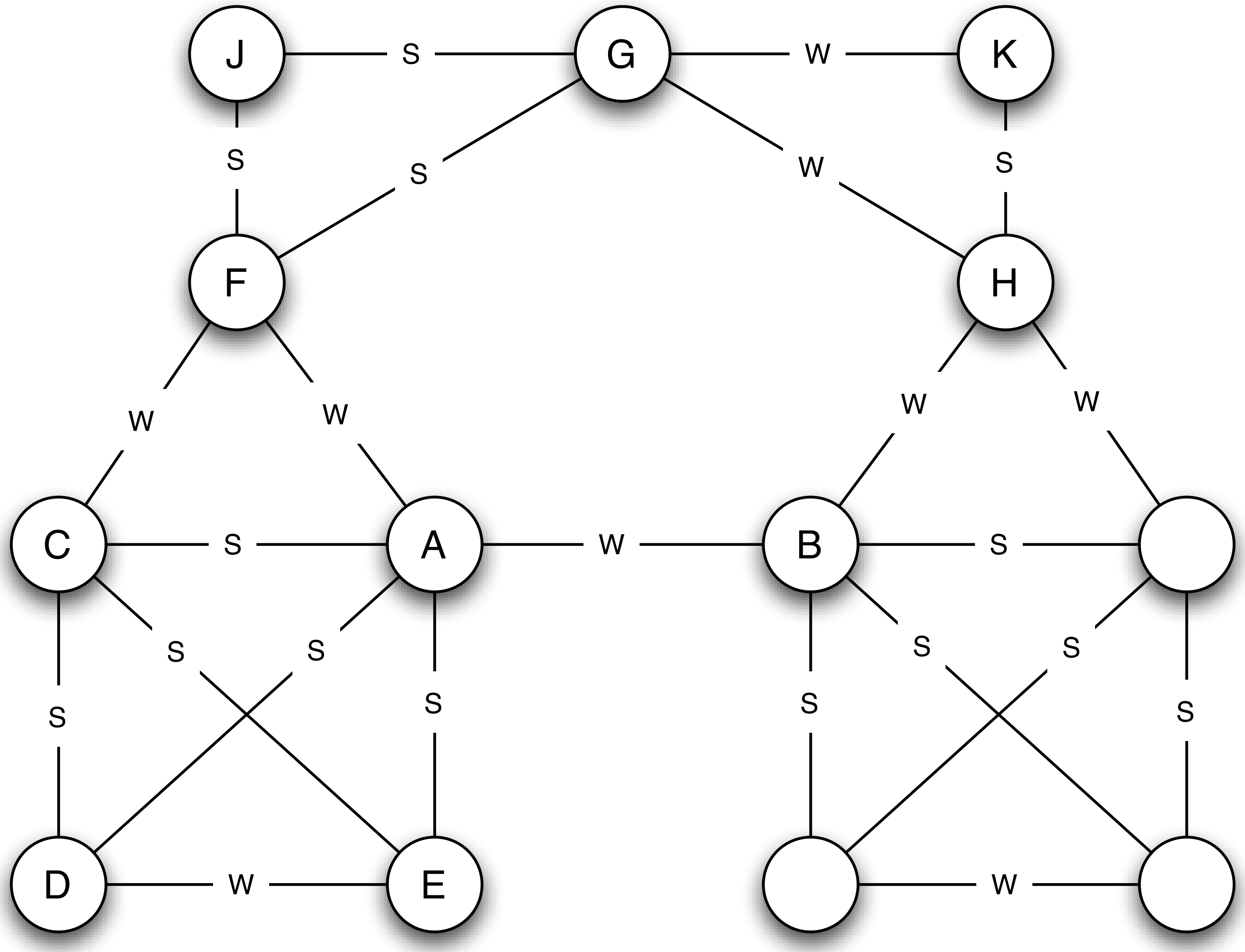



Figure from D. Easley and J.Kleinberg. 2010. Networks, Crowds, and Markets. Cambridge

# Closure and bridges

If a node A in a network satisfies the Strong Triadic Closure Property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie.

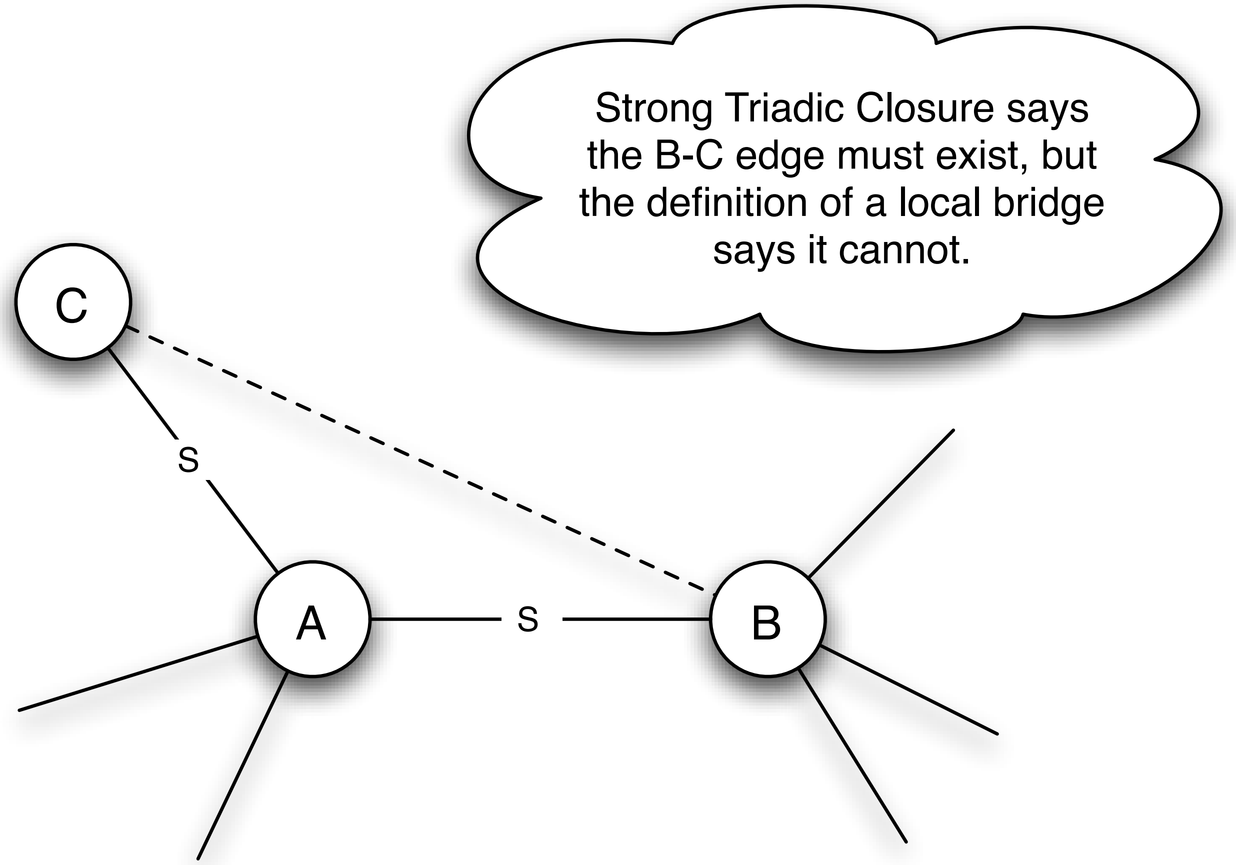



Figure from D. Easley and J.Kleinberg. 2010. Networks, Crowds, and Markets. Cambridge

# Closure and bridges

So local bridges are likely to be weak ties

Explaining why jobs came from weak ties

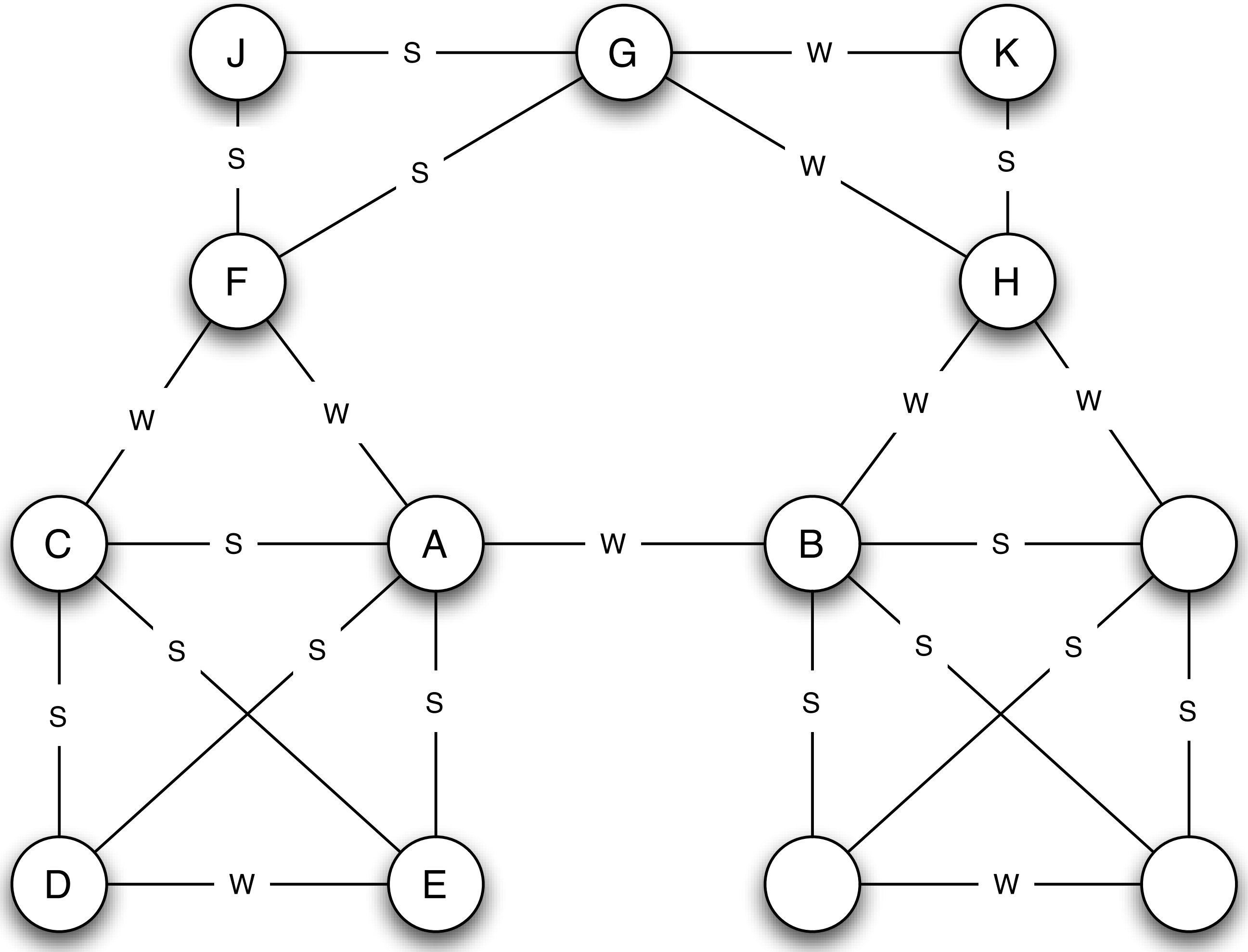



Figure from D. Easley and J.Kleinberg. 2010. Networks, Crowds, and Markets. Cambridge

# The Strength of Weak Ties

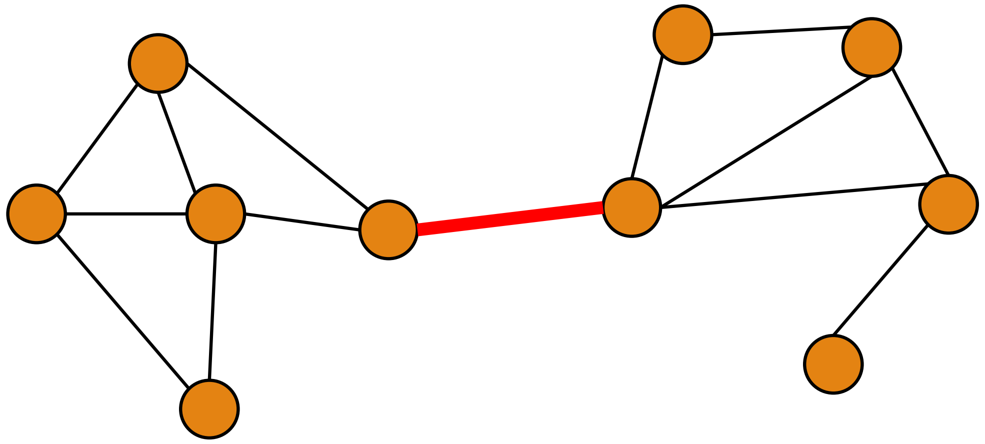
Weak ties can occur between cohesive groups: low transitivity

- old college friend
- former colleague from work

Weak ties have different *information* than close contacts

Granovetter found weak ties were short paths

- Compatible with Watts/Strogatz small world model



SLIDE ADAPTED FROM JAMES MOODY

# Summary

Triangles (triadic closure) lead to higher clustering coefficients

- Your friends will tend to befriend each other

Local bridges will often be weak ties

Information comes over weak ties

Social Networks & Power Laws

# Weak Links

Social Networks & Power Laws

# Power Laws

# Degree of nodes

Many nodes on the internet have low degree

- One or two connections

A few (hubs) have very high degree

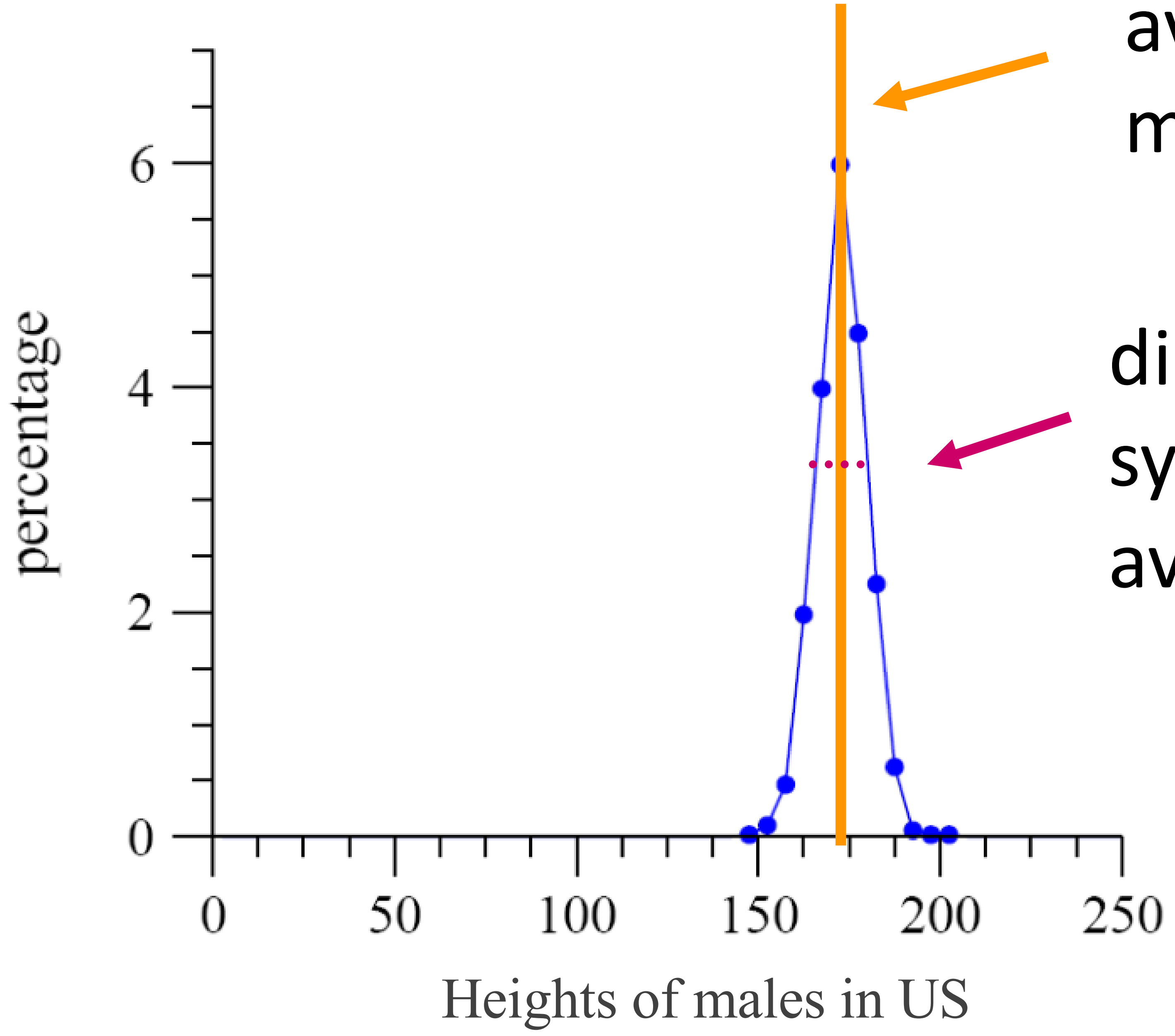
$P(k)$, # of nodes with degree k, follows a power law:

$$P(k) \propto k^{-\alpha}$$

Where α for the internet is about 2.1

- So the fraction of web pages with k in-links is proportional to $1/k^2$

# Gaussian (normal) distribution of human heights

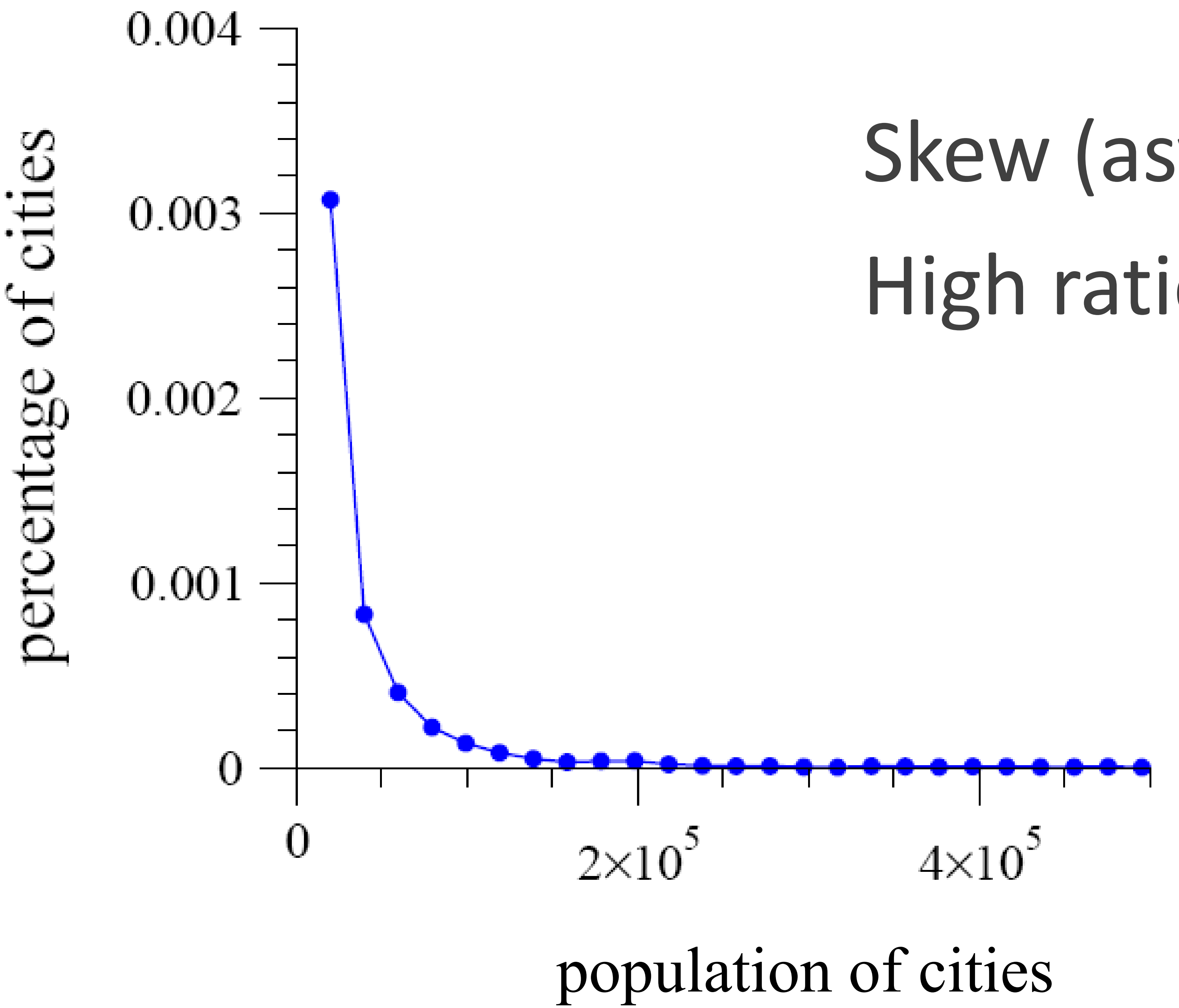



average value close to most typical

distribution close to symmetric around average value

Figure from Newman, Mark. "Power laws, Pareto distributions and Zipf's law." *Contemporary physics* 46, no. 5 (2005): 323-351. Slide from Lada Adamic

# Power-law distribution




Skew (asymmetry)

High ratio of max to min

Figure from Newman, Mark. "Power laws, Pareto distributions and Zipf's law." *Contemporary physics* 46, no. 5 (2005): 323-351. Slide from Lada Adamic

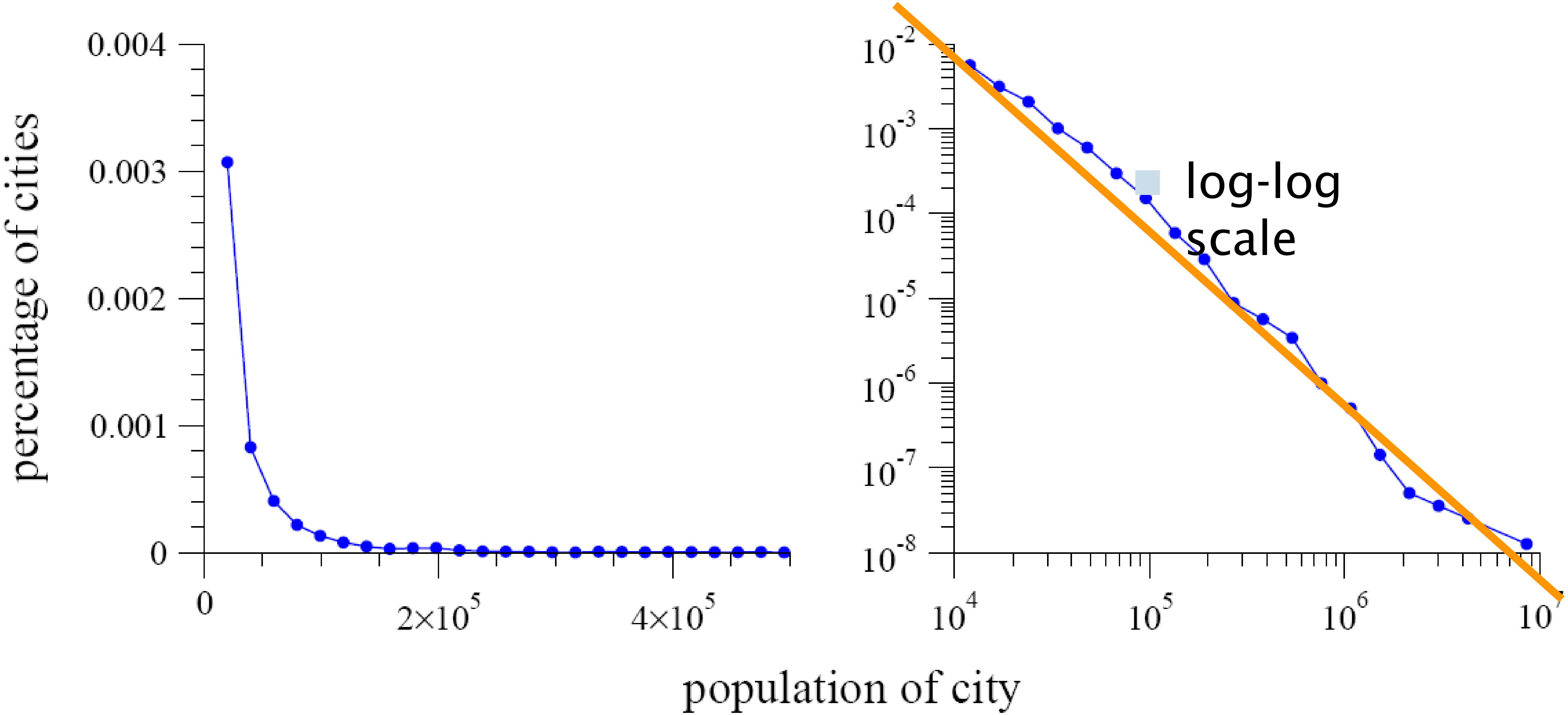
# Power-law distributions

Right skew
- normal distribution is centered on mean
- power-law or Zipf distribution is not

High ratio of max to min
- Populations of cities are power-law distributed
- Contrast: human heights are not (max and min not that different)

Slide from Lada Adamic

# Power-law distribution




- straight line on a log-log plot

Figure from Newman, Mark. "Power laws, Pareto distributions and Zipf's law." *Contemporary physics* 46, no. 5 (2005): 323-351. Slide from Lada Adamic

# Power law distribution are straight lines on log-log plots

$p(x)$: probability of observing an item with value x

Power law:

power law exponent α

$$p(x) = \frac{c}{x^{\alpha}} = cx^{-\alpha}$$

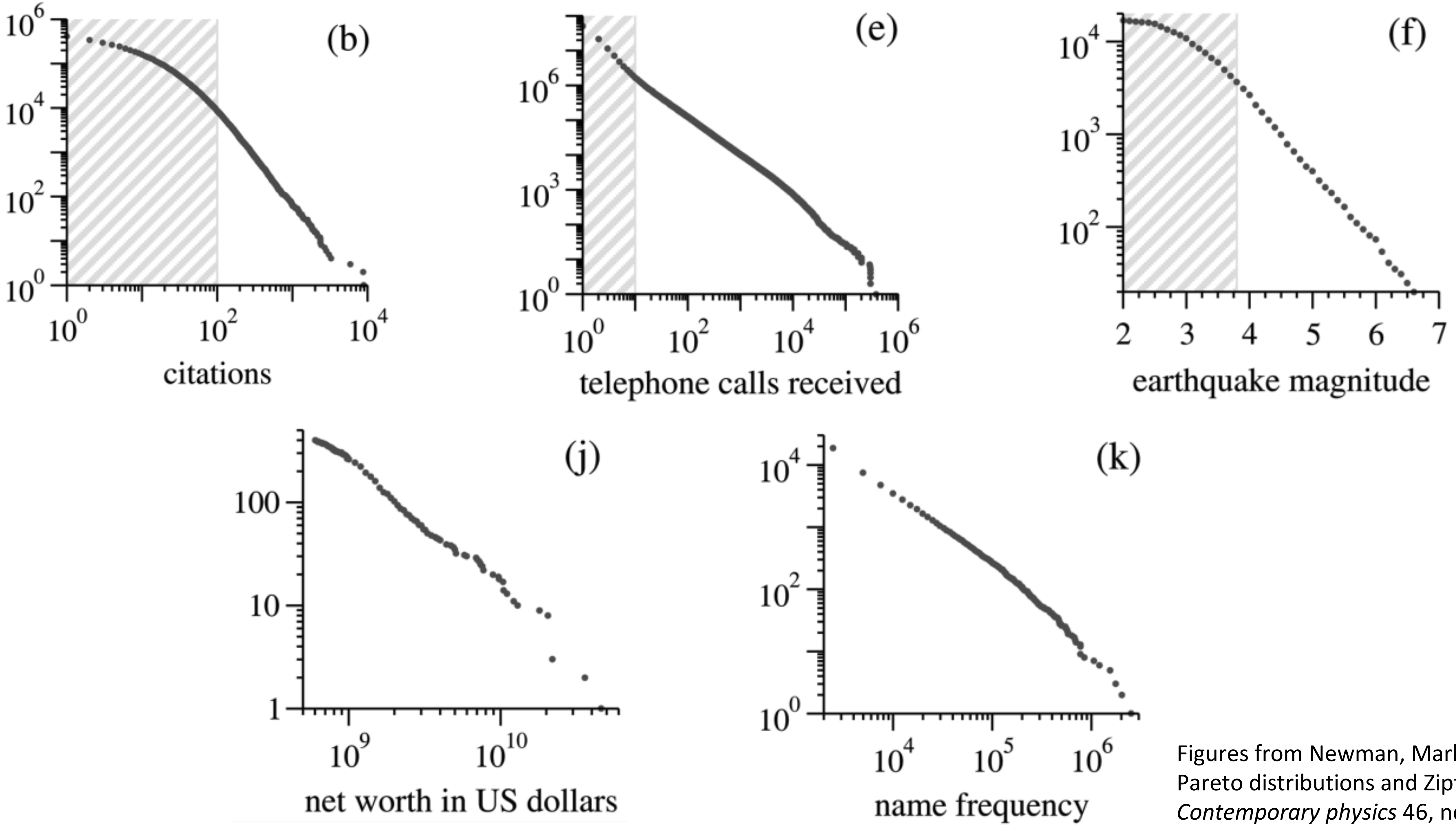
Take the log of both sides

$$\log(p(x)) = \log(c) - \alpha\log(x)$$

A line with $-\alpha$ as slope

Slide from Lada Adamic

# Power laws are everywhere!




Figures from Newman, Mark. "Power laws, Pareto distributions and Zipf's law." *Contemporary physics* 46, no. 5 (2005): 323-35

# Power laws are "scale free"

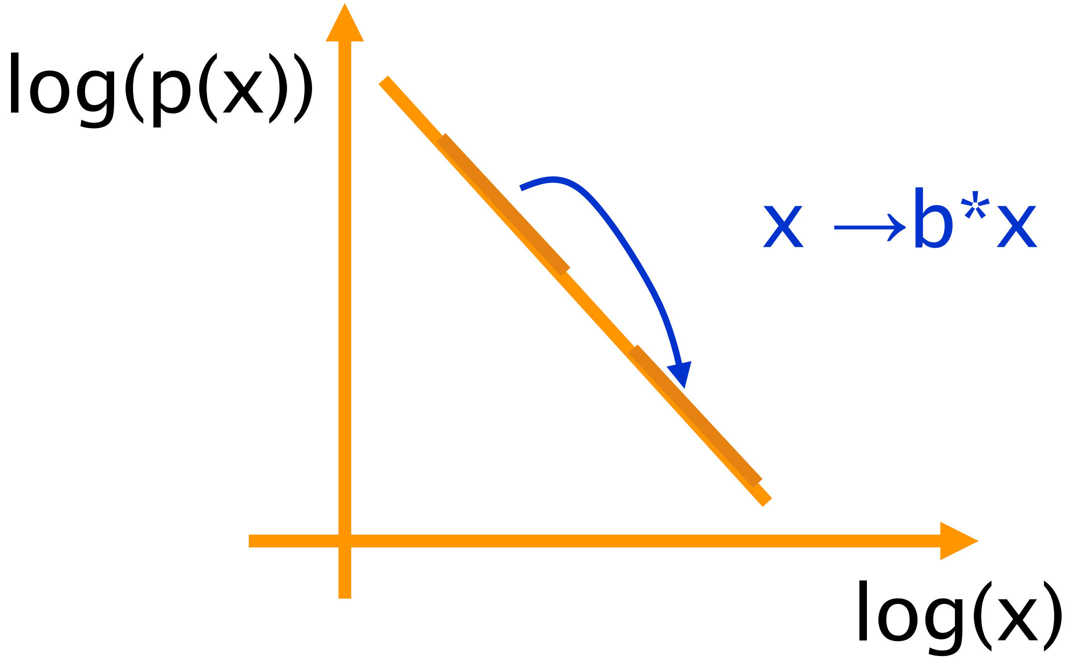
A power law looks the same no mater what scale we look at it on (2 to 50 or 200 to 5000)

Only true of a power-law distribution!

$p(bx) = g(b)\ p(x)$

Shape of the distribution unchanged except for a multiplicative constant

$p(bx) = (bx)^{-\alpha} = b^{-\alpha}\ x^{-\alpha}$




Slide from Lada Adamic

# Language: Zipf's law is a power-law

How frequent is the 3rd or 100th most common word?

Few very frequent words ("the", "of"), lots of rare ones ("expressive", "Jurafsky")

**Zipf's law (George K. Zipf):** the frequency of the $r$'th most frequent word is inversely proportional to its rank:

$$\text{freq}(r'\text{th most frequent word}) \propto \frac{1}{r^{\beta}}$$

with $\beta$ close to unity.

The most frequent word is twice as frequent as the 2nd, 3 times as frequent as the 3rd, etc.

# Pareto's law and power-laws

## Pareto

- The Italian economist Vilfredo Pareto was interested in the distribution of income.
- Pareto's law is expressed in terms of the cumulative distribution (the probability that a person earns X or more).

$$P[X > x] \sim x^{-k}$$

Slide from Lada Adamic

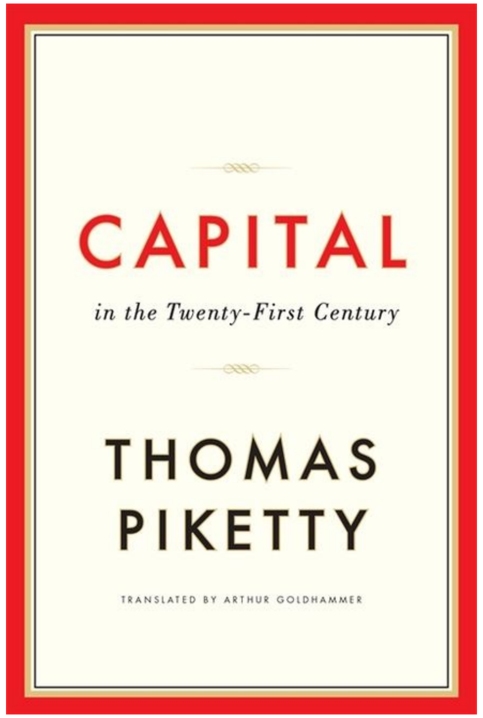
# Income

The fraction i of the income going to the richest P of the population:

$$i = (100/P)^{\eta-1}$$

- if $\eta = 0.5$

  top 1 percent gets $100^{-0.5} = .10$
- In 2015, $\eta = 0.6$ top 1 percent gets $100^{-0.4} = .16$
- (higher $\eta$ = more inequality)




Thomas Piketty's book, #1 on NY Times best seller list in 2014, studies how $\eta$ relates to rise of wealth inequality

Charles Jones. "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality" *Journal of Economic Perspectives*, Winter 2015.

# Where do power laws come from?

Many different processes can lead to power laws

There is no one unique mechanism

Slide from Lada Adamic

# Preferential attachment

Price (1965) model for citation networks

- **new citations to a paper are proportional to the number it already has**
- each new paper is generated with m citations
- new papers cite previous papers with probability proportional to their in-degree (citations)

Slide from Lada Adamic

# Preferential attachment is a "Rich get Richer" Model

Explanation for various power law effects

1. **Citations:** randomly draw citations from the reference section of papers
2. Assume **cities** are formed at different times, and that, once formed, a city grows in proportion to its current size simply as a result of people having children
3. **Words**: people are more likely to use a word that is frequent (perhaps it comes to mind more easily or faster)

# Power laws: Summary

Many processes are distributed as power laws

- Word frequencies, citations, web hits

Power law distributions have interesting properties

- scale free, skew, high max/min ratios

Various mechanisms explain their prevalence

- rich-get-richer, etc.

Explain lots of phenomena we have been dealing with

- the use of stop words lists (a small fraction of word types cover most tokens in running text)

# Social Networks & Power Laws

## Power Laws