



# CSCI4190

# Introduction to Social Networks



# Course Instructor

Laiwan Chan 陳麗雲 Lauren

- Room 1016, Ho Sin Hang Engineering Building
- Tel : 3943 8434
- Email : lwchan AT cse.cuhk.edu.hk
- <http://www.cse.cuhk.edu.hk/~lwchan>



- Tutors :

Liu Jie (jliu AT cse.cuhk.edu.hk)

Hou Yifan (yfhou AT cse.cuhk.edu.hk)



# Lecture Hours

- Lectures :
  - Tue 12:30 – 2:15pm, LSB LT5
  - Thu 1:30 – 2:15pm, LSB LT5
- Tutorial :
  - Wed 5:30 – 6:15 pm, LSB LT2



# Course Outline

- **Social Network and Graph Theory** (Chapters 1,2) (1 wk)
- **Network Structure** (3 wk)
  - Strong and weak ties (Chapter 3)
  - Networks in their surrounding contexts (Chapter 4)
  - Positive and negative relationships (Chapter 5)
- **Network Dynamics : population models** (3 wk)
  - Information Cascades (Chapter 16)
  - Network Effects (Chapter 17)
  - Power laws and rich-get-richer phenomena (Chapter 18)
- **Network Dynamics : structural models** (3 wk)
  - Cascading behavior in networks (Chapter 19)
  - The small world phenomenon (Chapter 20)
  - Epidemics (Chapter 21)
- **Information networks and the WWW** (3 wk)
  - Link analysis and web search (Chapter 14)
  - Sponsored search markets (Chapter 15)



# Course objectives

- To introduce the concepts and principles of **social network analysis**
- To model the **network structure** using graphs and introduce their characteristics
- To model the **interaction and aggregate behavior** of people who connect through social networks



# This course is NOT

- for students who spend a few hours a day on social network.
- for students who only know how to click



- A blue icon of a tablet computer.



# This course is

- for students who want to learn how we model the behaviors of individuals and how these behaviors influence others in the network.



# Text Book

- Networks, Crowds, and Markets: Reasoning About a Highly Connected World

by David Easley and Jon Kleinberg,  
Cambridge University Press, 2010

<https://www.cs.cornell.edu/home/kleinber/networks-book/>



# Course Information

- blackboard : <https://blackboard.cuhk.edu.hk>

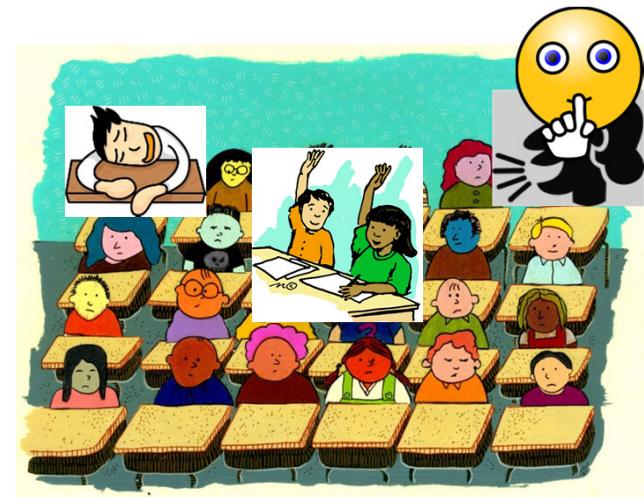
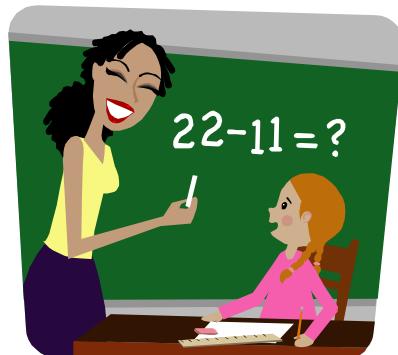


# Assessment

- 60 % Final Exam and with a minimum passing score of 40/100
- 20 % Assignments
- 20 % Course Project



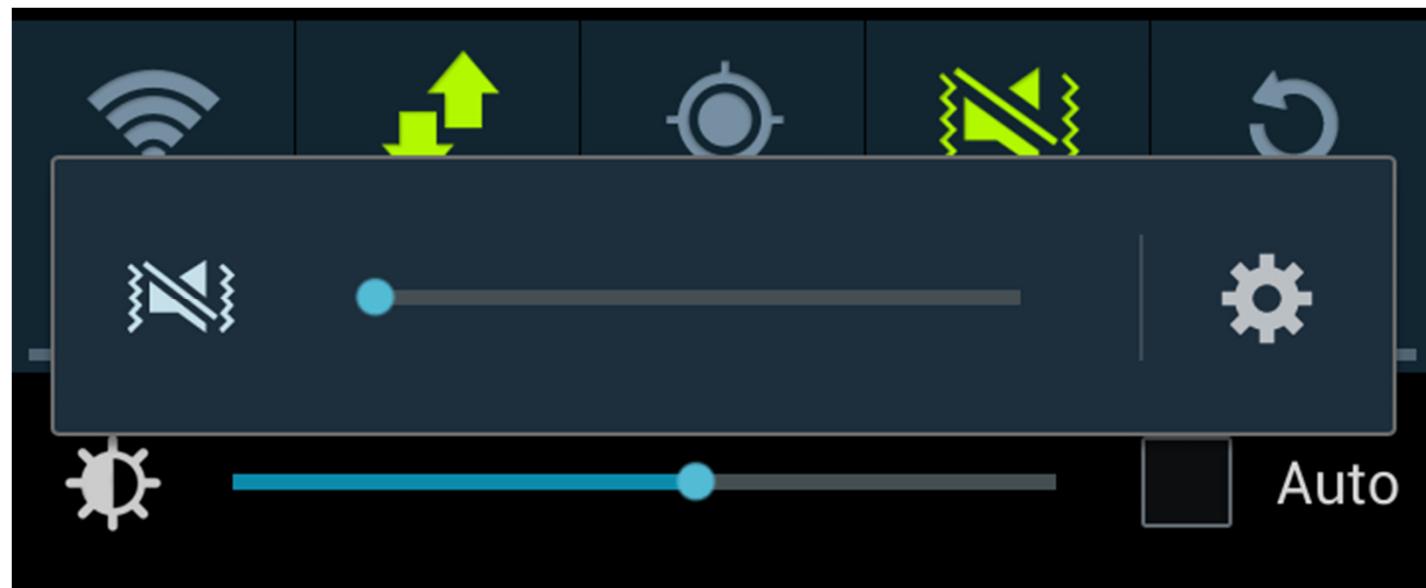
# CLASSROOM MANAGEMENT





**PLEASE TURN OFF YOUR  
PHONES/BEEPERS**







# PLEASE DO NOT RECORD





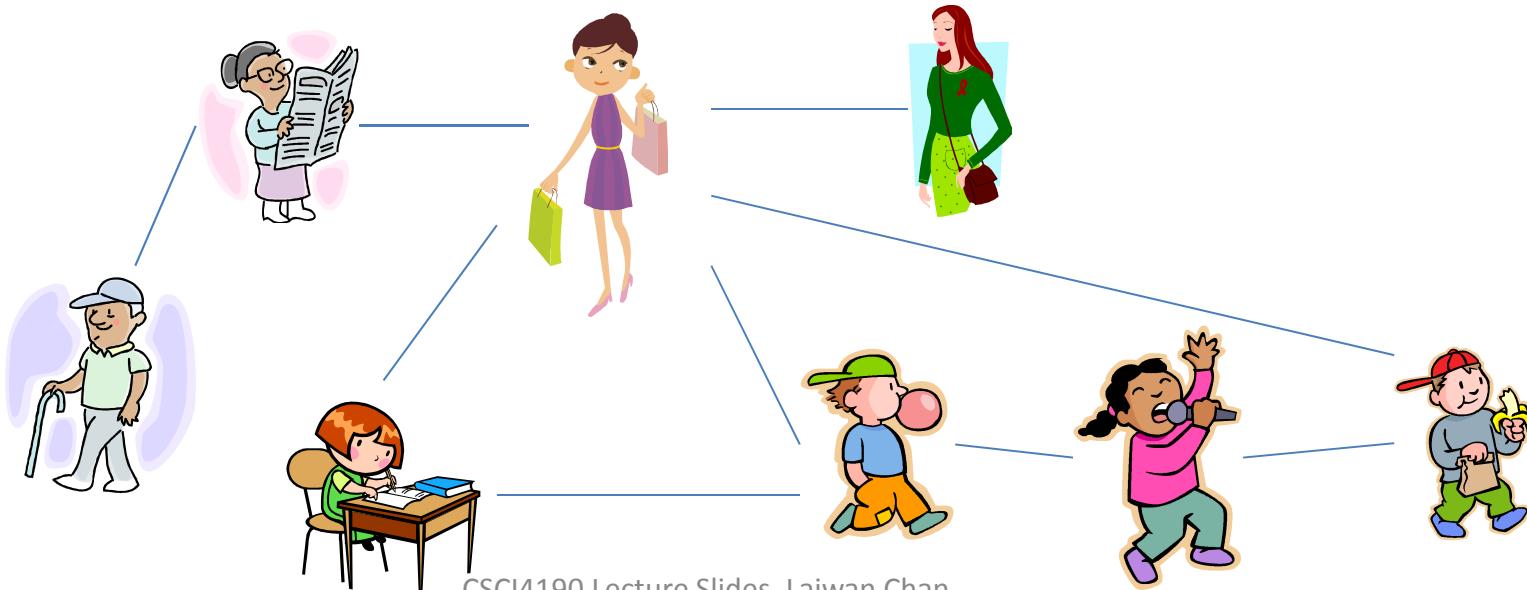
# Honesty in Academic Works

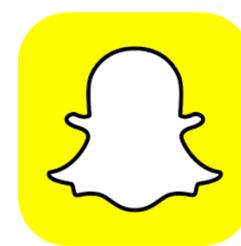
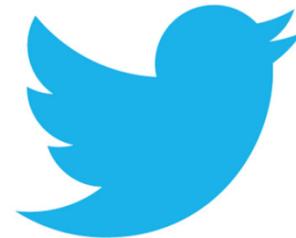
- Attention is drawn to University policy and regulations on honesty in academic work, and to the disciplinary guidelines and procedures applicable to breaches of such policy and regulations. Details may be found at <http://www.cuhk.edu.hk/policy/academichonesty/> .
- With each assignment, students will be required to submit a signed declaration that they are aware of these policies, regulations, guidelines and procedures. For group projects, all students of the same group should be asked to sign on the declaration.
- For assignments in the form of a computer-generated document that is principally text-based and submitted via VeriGuide, the statement, in the form of a receipt, will be issued by the system upon students' uploading of the soft copy of the assignment. Assignments without the receipt will not be graded by teachers. Only the final version of the assignment should be submitted via VeriGuide.



# Social Network

- A social network is a social structure made up of a set of social actors (such as individuals or organizations) and a set of the dyadic ties between these actors.





Telegram



Instagram





|              |             |              |
|--------------|-------------|--------------|
| Bebo         | Blogger     | Vimeo        |
| Brightkite   | Delicious   | Windows      |
| Design Float | Digg        | Yahoo!       |
| Dopplr       | Email       | Yelp         |
| Facebook     | Flickr      | Tumblr       |
| Friendfeed   | Gamespot    | Viddler      |
| Google       | Google Talk | Virb         |
| Last.fm      | Linkedin    | Wordpress    |
| Mixx         | Mobile Me   | Yahoo! Buzz  |
| Myspace      | Netvibes    | YouTube      |
| Newsvine     | Openid      | Stumble Upon |
| Picasa       | Plurk       | Technorati   |
| Posterous    | Reddit      |              |
| RSS          | Skype       |              |





- **Social network** : a social structure made up of social actors (such as individuals or organizations) and a set of dyadic ties between these actors.
- **Social network analysis (SNA)** : analysis of social network via network theory, consisting of nodes (actors) and ties (relationship)
- **Social network service** : a platform to build social networks or social relations among people who share interests, activities, backgrounds or real-life connections.
- **Social media** : computer-mediated tools that allow people to create, share or exchange information, ideas and pictures/videos in virtual communities and networks.



# Social media

- A group of internet-based applications that build on the ideological and technological foundations of Web2.0 and that allow the creation and exchange of user-generated content.
  - collaborative projects (for example, [Wikipedia](#))
  - Blogs (for example, [WordPress](#))
  - content communities (for example, [YouTube](#) )
  - social networking sites (for example, [Facebook](#))
  - virtual game-worlds (e.g., [World of Warcraft](#))
  - virtual social worlds (e.g. [Second Life](#))



| Social network   | Social media                        |
|--|-------------------------------------|
| Groups of people   | Broad audience                      |
| Interaction and engagement between actors that are connected           | A system to disseminate information |
| Two way and direct communication between the actors that are connected | One to many communication           |



## Top 15 Most Popular Social Networking Sites





# Popularity varies with location

WORLD MAP OF SOCIAL NETWORKS  
January 2019

Facebook is the leading social network in 153 out of 167 countries.



credits: Vincenzo Cosenza [vincos.it](http://vincos.it)

license: CC-BY-NC

source: Alexa/SimilarWeb

From <https://vincos.it/world-map-of-social-networks/>



# Popularity varies with location

## WORLD MAP OF SOCIAL NETWORKS

January 2018

Facebook is the dominant social network 2.2 billion users

Russia is dominated by Odnoklassniki , VK and China is dominated by Qzone.



Japan was the country where Twitter was the leader in 2017 but Facebook has now swallowed up Japan.

- Facebook
- QZone
- VKontakte
- Odnoklassniki
- Instagram

credits: Vincenzo Cosenza vincos.it

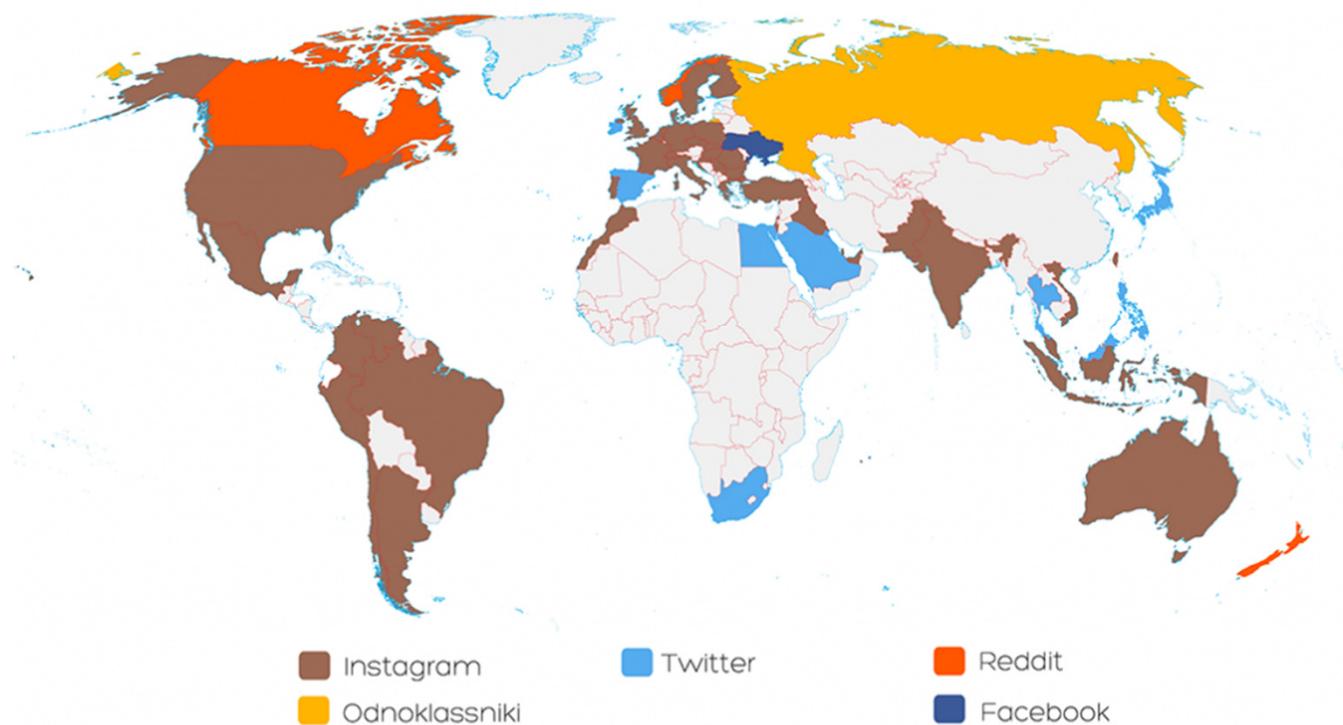
license: CC-BY-NC

source: Alexa/SimilarWeb



# WORLD MAP OF SOCIAL NETWORKS

Ranked 2nd - January 2019



credits: Vincenzo Cosenza [vincos.it](http://vincos.it)

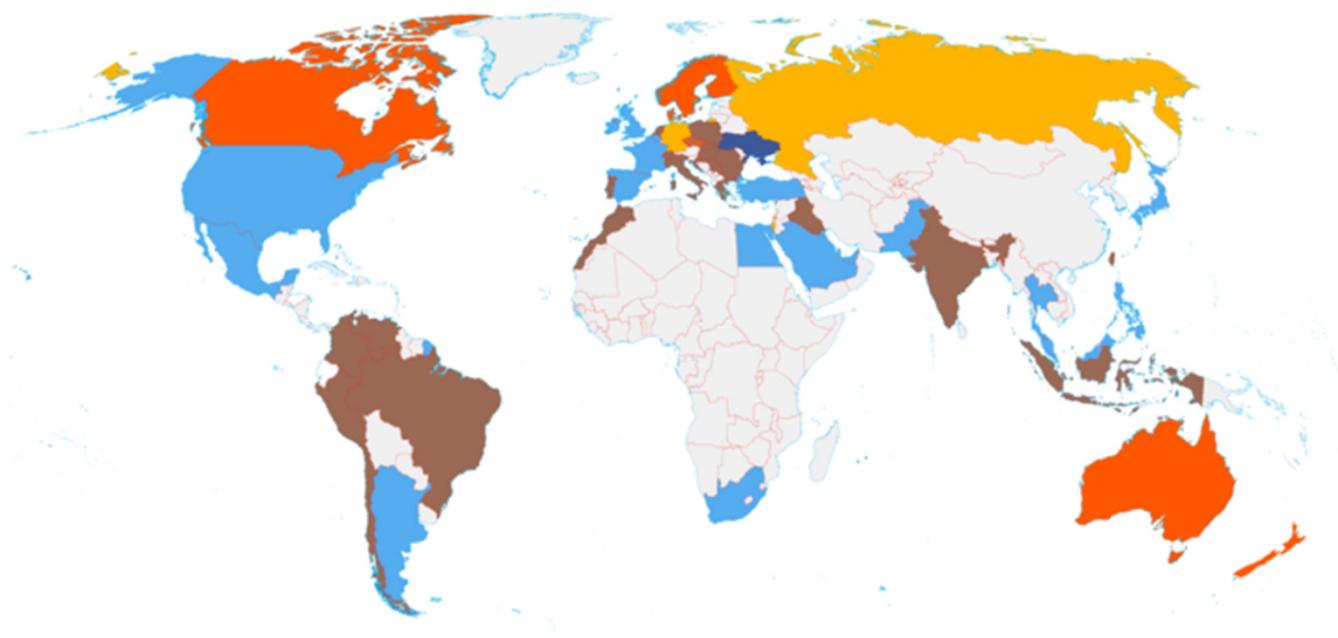
license: CC-BY-NC

source: SimilarWeb/Alexa



# WORLD MAP OF SOCIAL NETWORKS

Ranked 2nd - January 2018



Instagram

Odnoklassniki

Twitter

V Kontakte

Reddit

Facebook

credits: Vincenzo Cosenza vincos.it

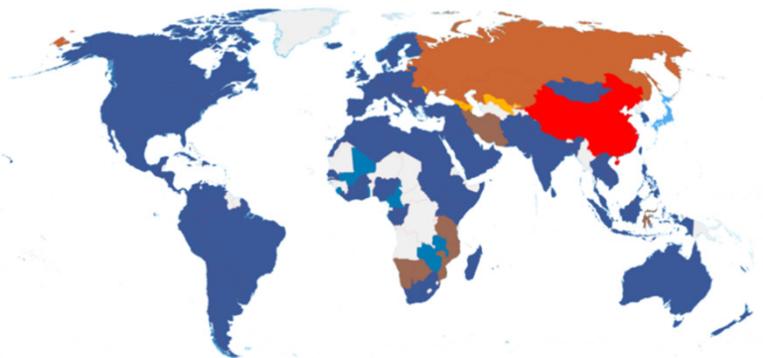
license: CC-BY-NC

source: SimilarWeb/Alexa



## WORLD MAP OF SOCIAL NETWORKS

January 2017



credits: Vincenzo Cosenza vincos.it

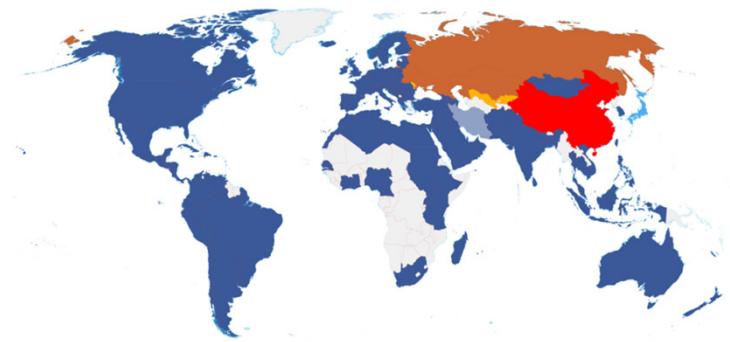
license: CC-BY-NC

source: Alexa/SimilarWeb

January 2016: a new edition of my World Map of Social Networks, showing the most popular social networking sites by country, according to Alexa & SimilarWeb traffic data (caveat: it's hard to understand the impact of Google+ because it is part of Google domain traffic).

## WORLD MAP OF SOCIAL NETWORKS

January 2016



credits: Vincenzo Cosenza vincos.it

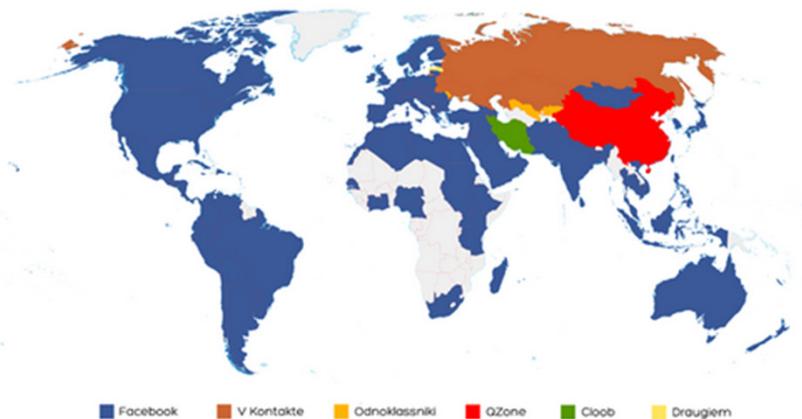
license: CC-BY-NC

source: Alexa/SimilarWeb

Second Term 19/20

## WORLD MAP OF SOCIAL NETWORKS

December 2013



Facebook      Vkontakte      Odnoklassniki      QZone      Cloob      Draugiem

June 2009



Facebook      Vkontakte      Odnoklassniki      Lidé      Hyves      Zing      HIS  
Orkut      Nasza-Klasa      Ozone      Iwiw      Maktoob      One      Mod  
Friendster      Wretch      Cyworld



## <http://www.ebizmba.com/articles/social-networking-websites>

### Top 15 Most Popular Social Networking Sites | July 2017

Here are the top 15 Most Popular Social Networking Sites as derived from our eBizMBA Rank which is a continually updated

average of each website's U.S. Traffic Rank from Quantcast and Global Traffic Rank from both Alexa and SimilarWeb.\*\*#\*\*

Denotes an estimate for sites with limited data.



#### 1 | facebook

3 - eBizMBA Rank | **1,500,000,000** - Estimated Unique Monthly Visitors

Alexa Rank | 2 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 2 | YouTube

3 - eBizMBA Rank | **1,499,000,000** - Estimated Unique Monthly Visitors

Alexa Rank | 3 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA

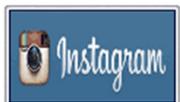


#### 3 | Twitter

11 - eBizMBA Rank | **400,000,000** - Estimated Unique Monthly Visitors

Alexa Rank | 7 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 4 | Instagram

26 - eBizMBA Rank | **275,000,000** - Estimated Unique Monthly Visitors

Alexa Rank | 15 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 5 | LinkedIn

26 - eBizMBA Rank | **250,000,000** - Estimated Unique Monthly Visitors

Alexa Rank | 38 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 6 | Reddit

30 - eBizMBA Rank | **125,000,000** - Estimated Unique Monthly Visitors

Alexa Rank | 28 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 7 | VK

32 - eBizMBA Rank | **120,000,000** - Estimated Unique Monthly Visitors | \*75\* - Quantcast Rank | 15

- Alexa Rank | 5 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 8 | Tumblr

39 - eBizMBA Rank | **110,000,000** - Estimated Unique Monthly Visitors | 30 - Quantcast Rank | 48

- Alexa Rank | 40 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 9 | Pinterest

40 - eBizMBA Rank | **105,000,000** - Estimated Unique Monthly Visitors | 13 - Quantcast Rank | 67

- Alexa Rank | 39 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 10 | Google Plus

42 - eBizMBA Rank | **100,000,000** - Estimated Unique Monthly Visitors | \*30\* - Quantcast Rank |

\*45\* - Alexa Rank | \*50\* - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 11 | Flickr

219 - eBizMBA Rank | **80,000,000** - Estimated Unique Monthly Visitors | 147 - Quantcast Rank | 367

- Alexa Rank | 144 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 12 | meetup

579 - eBizMBA Rank | **42,000,000** - Estimated Unique Monthly Visitors | 312 - Quantcast Rank | 666

- Alexa Rank | 760 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 13 | Ask.fm

585 - eBizMBA Rank | **40,000,000** - Estimated Unique Monthly Visitors | 919 - Quantcast Rank | 555

- Alexa Rank | 282 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 14 | LiveJournal

648 - eBizMBA Rank | **37,000,000** - Estimated Unique Monthly Visitors | 1415 - Quantcast Rank |

252 - Alexa Rank | 277 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



#### 15 | myspace

2531 - eBizMBA Rank | **10,000,000** - Estimated Unique Monthly Visitors | 1728 - Quantcast Rank |

3197 - Alexa Rank | 2668 - SimilarWeb Rank | *Last Updated: July 1, 2017.*

The Most Popular Social Networking Sites | eBizMBA



# Some popular social media sites of the world

## General

- [Facebook](#) : for users to connect and share content
- [Google+](#) : launched in 2011 for staying in touch with friends
- [Myspace](#) : launched in 2003
- [Twitter](#) : a service for posting and receiving messages to a network of contacts, as opposed to send bulk email messages.
- [Reddit](#) : a social news and entertainment website
- [Qzone](#) : popular in mainland China
- [Sina Weibo](#) : Social microblogging in mainland China
- [Orkut](#) : most popular in Brazil and India
- [VK](#) : most popular in Russia

## Business and Professional networking

- [LinkedIn](#): an online social network for professional networking, to help business professionals find a job and connect with potential business partners. It does not focus on making friends or sharing media like photos, videos and music.



# Some popular social media sites of the world

## Special Interests

- [Pinterest](#) : a virtual pinboard that users share images that are related to a particular interest or hobbies.
- [Instagram](#) : launched in 2010 . It is a smartphone application for sharing photos with other users.
- [Flickr](#) : online photo sharing service
- [Snapchat](#) : a photo messaging app launched in 2011
- [Youtube](#) : video sharing platform
- [Vine](#) : a mobile app for posting video clips
- [Soundcloud](#) : an online audio distribution platform
- [Yelp](#) : local business review

## Microblogging

- [Tumblr](#) : microblogging platform
- [Wordpress](#) : microblogging platform
- [Xanga](#) : microblogging platform

## Location

- [Foursquare](#) : for users to checkin at various places



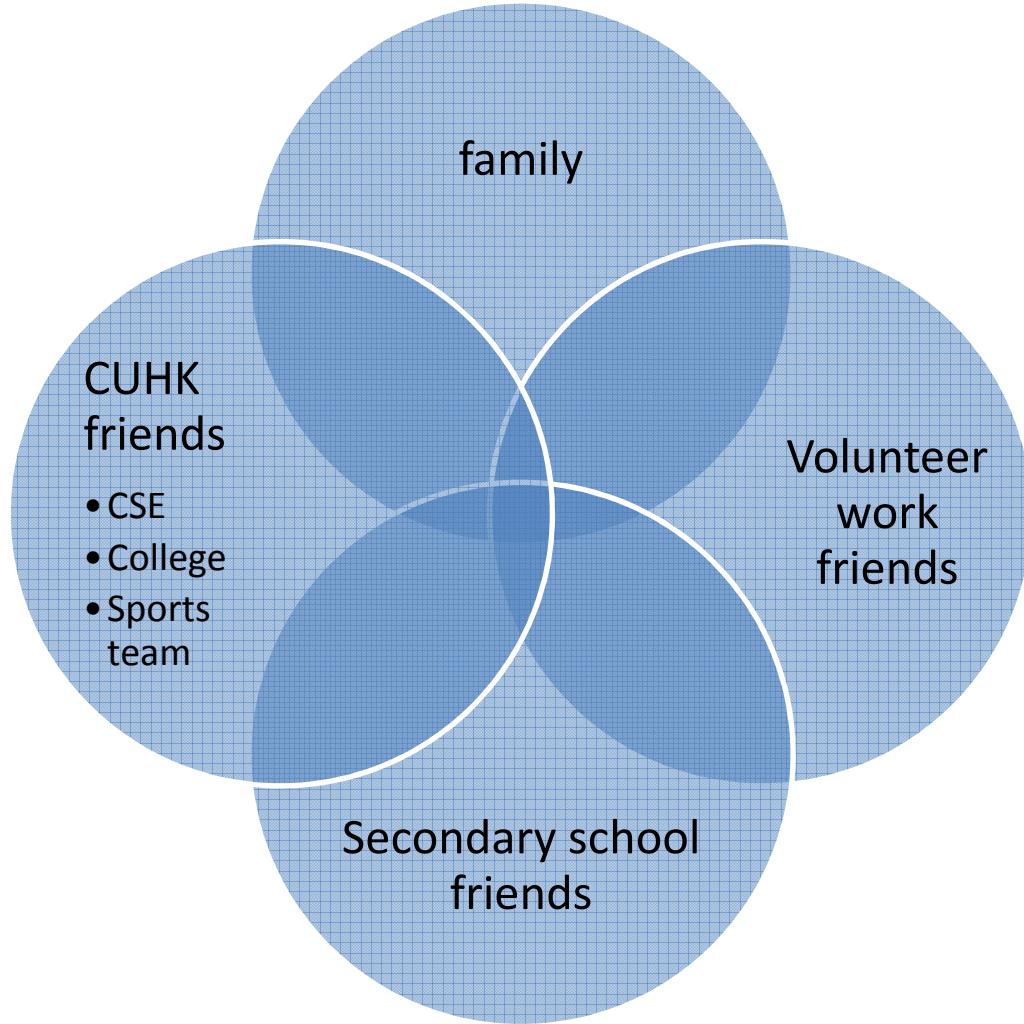
# Reasons of social networking ?

- Keep in touch with friends and family members
- Share and receive information
- Create new friends, business contacts and job hunting



# Features

- Profile
- Groups
- Search for friends
- Messages-mails
- Notification
- Sharing of ideas, activities, multimedia (pictures and video), articles etc.
- Tagging
- Posts/comments/like/unlike
- Apps/games
- Events





- Connectedness at the level of **structure** – who is linked to whom
- Connectedness at the level of **behavior** –each individual's actions have implicit consequences for the outcomes of everybody in the system (**behavior and interaction** in network contexts) .



# Main themes in this course

- **Network Structure** : who is linked to whom
  - Communities
  - Close versus casual social contacts
  - Structure balance
  - Small world phenomenon
- **Network Dynamics** : Information diffusion or cascading behavior, *i.e.* when individuals adopt the behavior of their neighbors in the network.
  - Network effects
  - Rich-get-richer
  - Epidemics



# Network structure

- Dynamic (changes with time)
- Snapshot of the nodes and edges at a particular moment in time – static structures
- Paths, components, distances etc

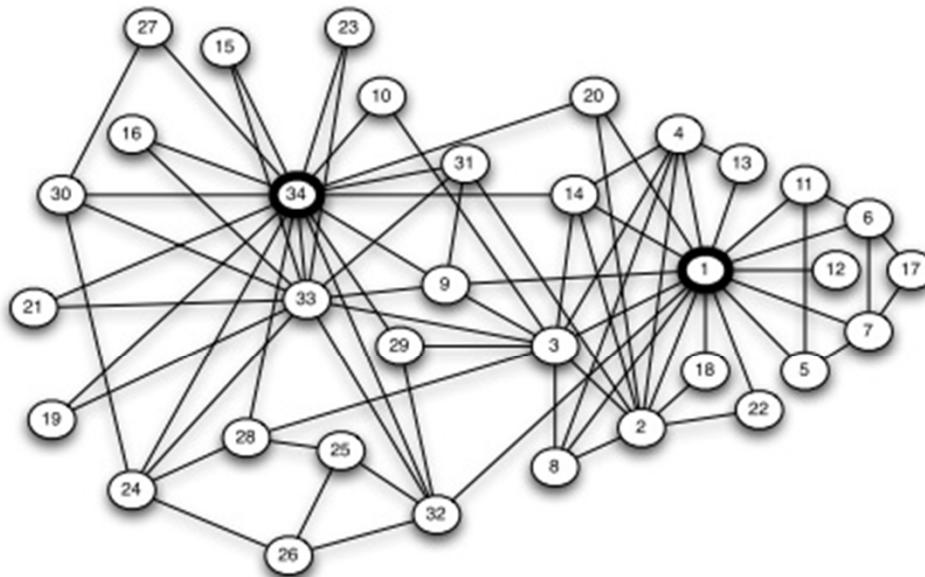


Figure 1.1: The social network of friendships within a 34-person karate club [421].

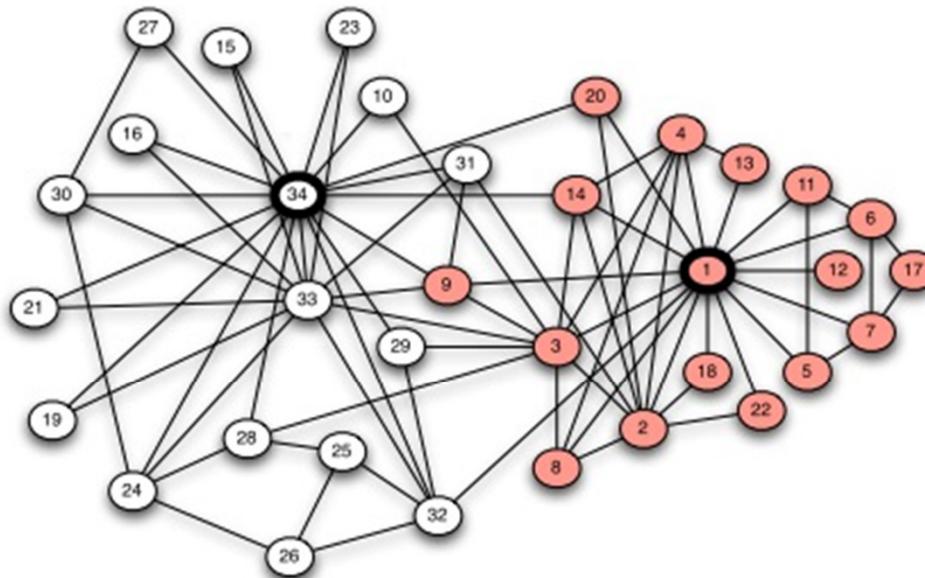


Figure 1.7: From the social network of friendships in the karate club from Figure 1.1, we can find clues to the latent schism that eventually split the group into two separate clubs (indicated by the two different shadings of individuals in the picture).

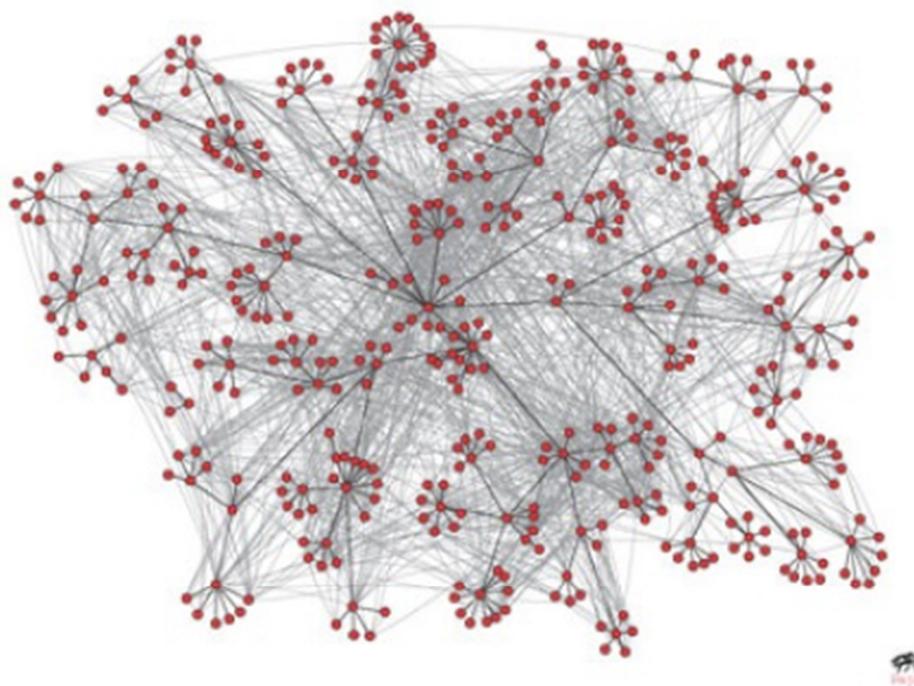


Figure 1.2: Social networks based on communication and interaction can also be constructed from the traces left by on-line data. In this case, the pattern of e-mail communication among 436 employees of Hewlett Packard Research Lab is superimposed on the official organizational hierarchy [6]. (Image from <http://www-personal.umich.edu/~ladamic/img/hplabsemailhierarchy.jpg>)



# Network Dynamics

- Promotion or recommendation of a product or idea on a social network; the rich get richer.
- In March 2011, fearing the nuclear crisis in Japan, people in China and Hong Kong had a panic buying of salt.
- In November 2013, panic buying of salt happened in a number of cities in India, due to some rumors on supply and price rise.
- In October 2014, large number of people in Hong Kong hit the “unfriend” button and lots of arguments on social media.



# Aspects of networks

- A network is any collection of objects in which some pairs of these objects are connected by links.
- Social Networks are typically networks which individuals are linked if they are “friends”.
- Other networks include
  - network of loans among financial institutes.
  - www which web pages are linked.



# Graphs

- A graph is a way of specifying relationships among a collection of items
- Graph  $G = (V, E)$
- $V$  = a set of nodes or vertices
- $E$  = a set of edges
- Two nodes are neighbors if they are connected by an edge
- An edge could be directed edge or undirected edge.
  - Directed edge : Followers on Twitter
  - Undirected edge : Friends on facebook



# Graphs

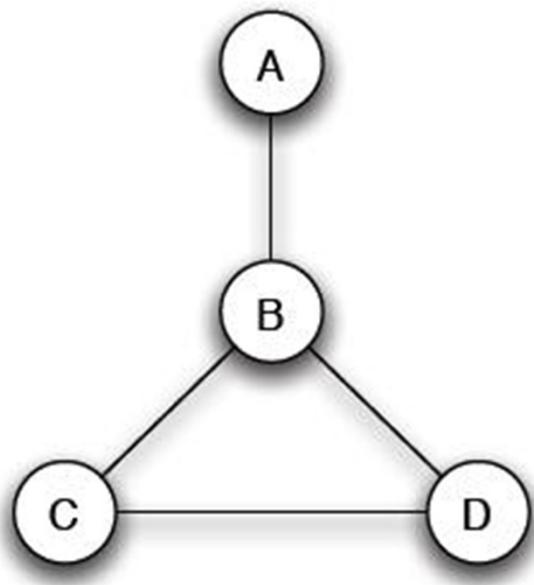
- A graph is a way of specifying relationships among a collection of items

$$G = (V, E)$$

$V = \{i\}$  : a set of nodes or vertices

$E = \{(i,j)\}$  : a set of edges

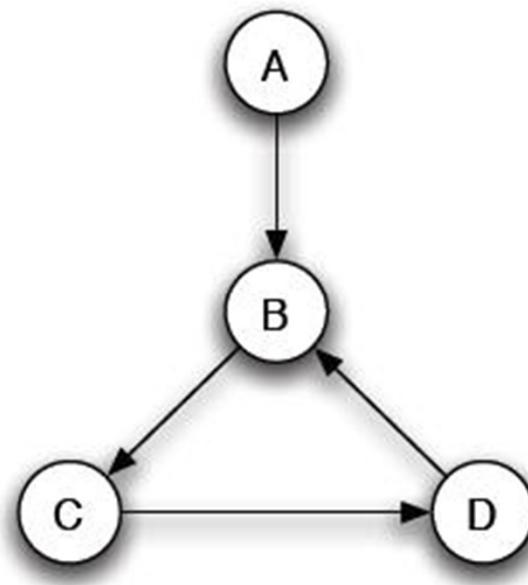




**Undirected Graph**

$$V = \{A, B, C, D\}$$

$$E = \{(A, B), (B, C), (C, D), (B, D)\}$$



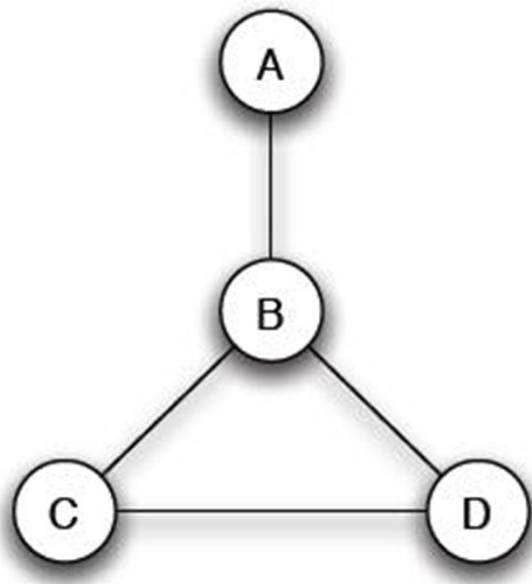
**Directed Graph**

$$V = \{A, B, C, D\}$$

$$E = \{(A, B), (B, C), (C, D), (D, B)\}$$



# Adjacency Matrix



**Undirected Graph**

$M_{ij} = 1$  if there is an edge from  $i$  to  $j$   
 $M_{ij} = 0$  otherwise

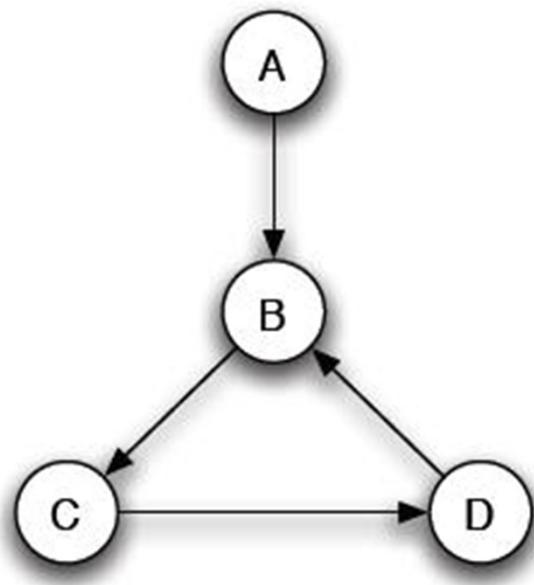
$$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

$$V = \{A, B, C, D\}$$

$$E = \{(A, B), (B, C), (C, D), (B, D)\}$$



# Adjacency Matrix



**Directed Graph**

$M_{ij} = 1$  if there is an edge from  $i$  to  $j$   
 $M_{ij} = 0$  otherwise

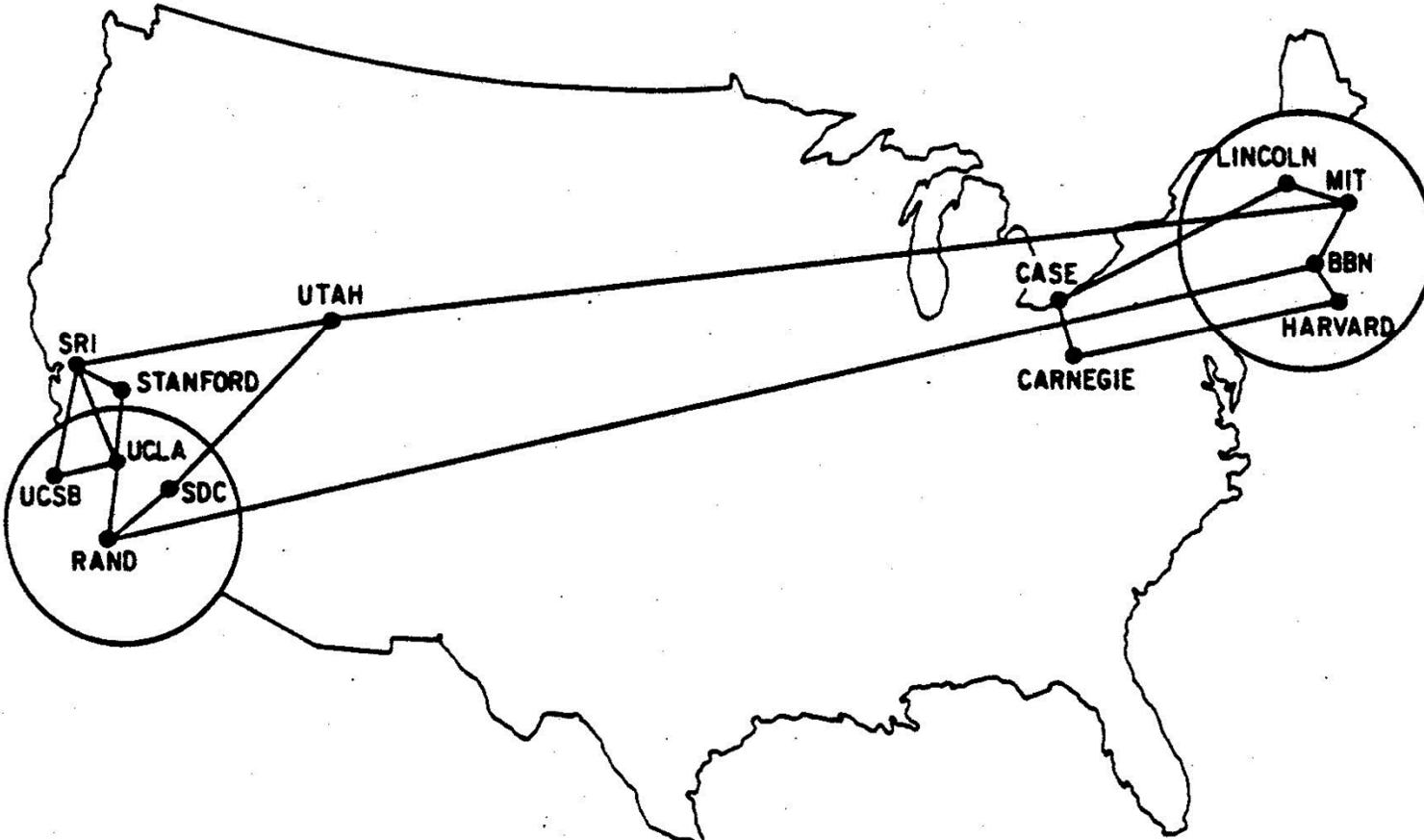
$$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$V = \{A, B, C, D\}$$

$$E = \{(A, B), (B, C), (C, D), (D, B)\}$$

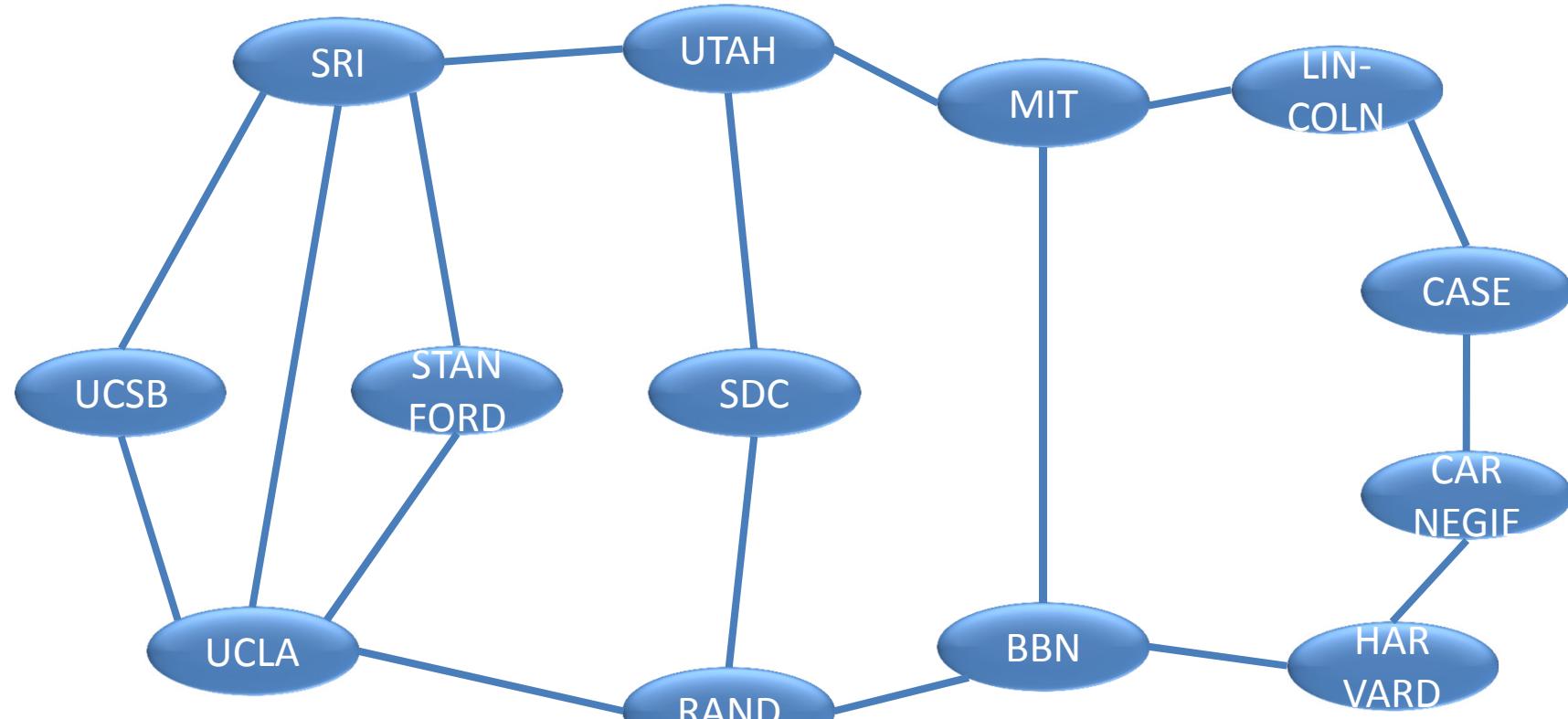


# Arpanet (Advanced Research Project Agency Network) – the first internet





# Example: Arpanet



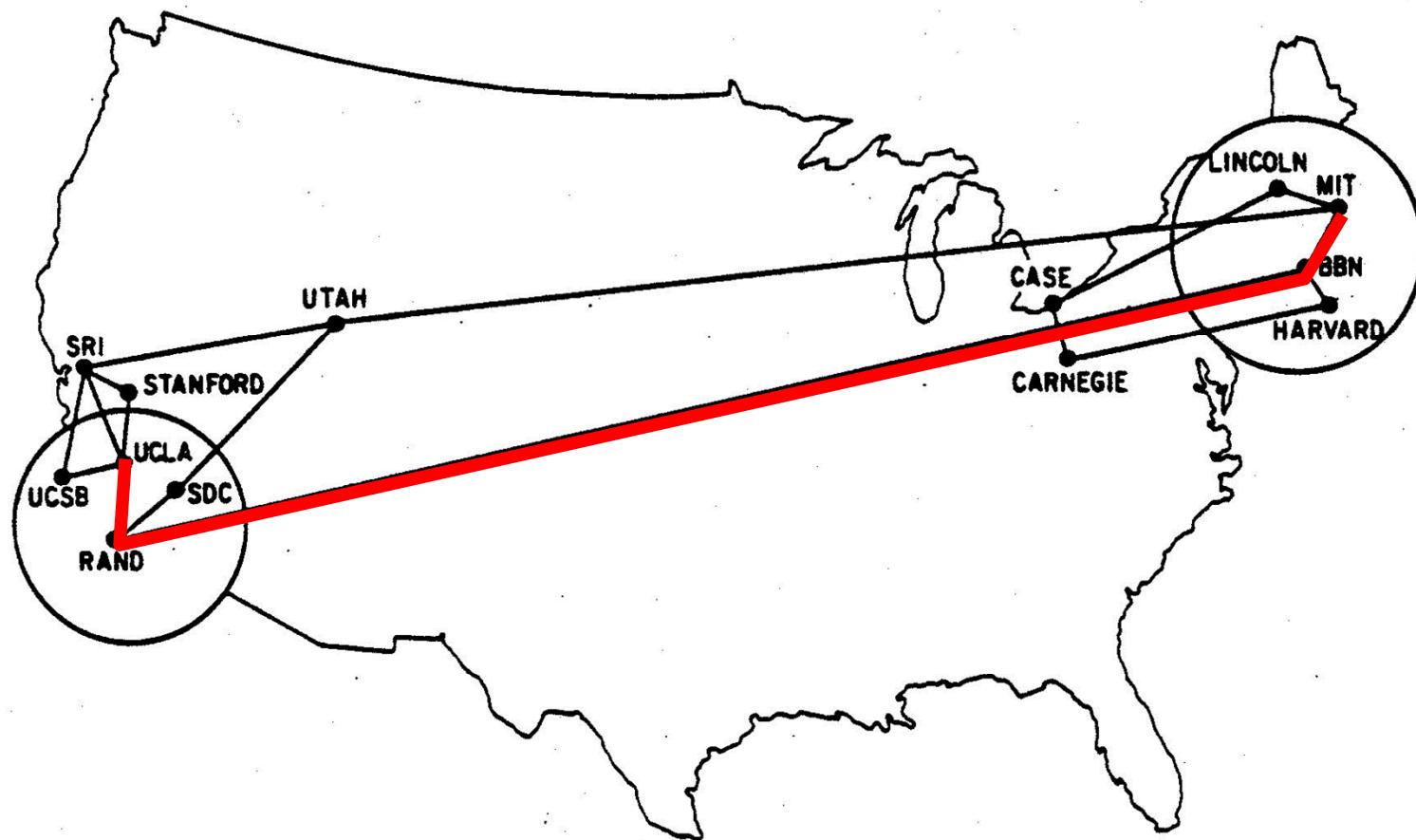


# Path

- A path in a graph is a sequence of nodes with the property that each consecutive pair in the sequence is connected by an edge.
- E.g. the sequence of nodes MIT, BBN, RAND, UCLA

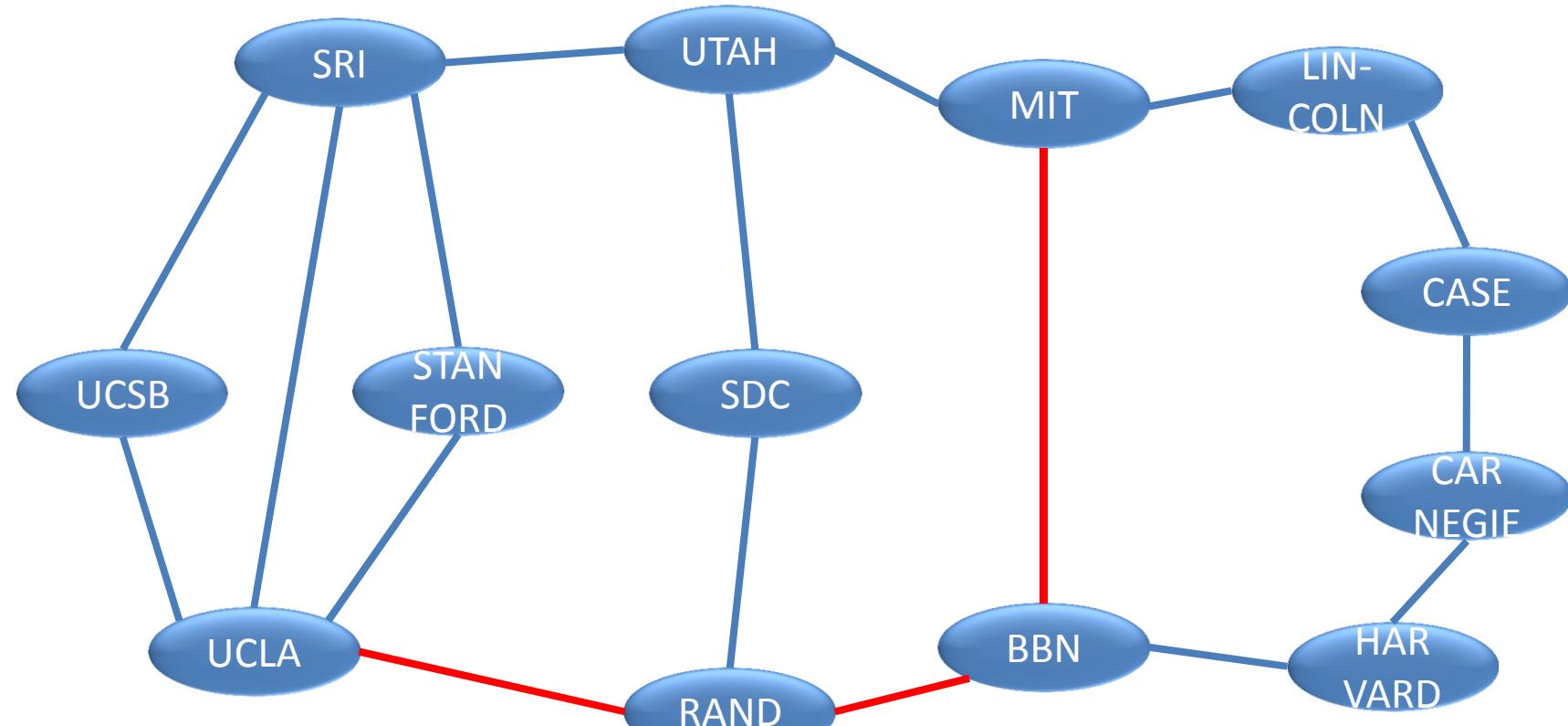


# Arpanet





# Example: Arpanet



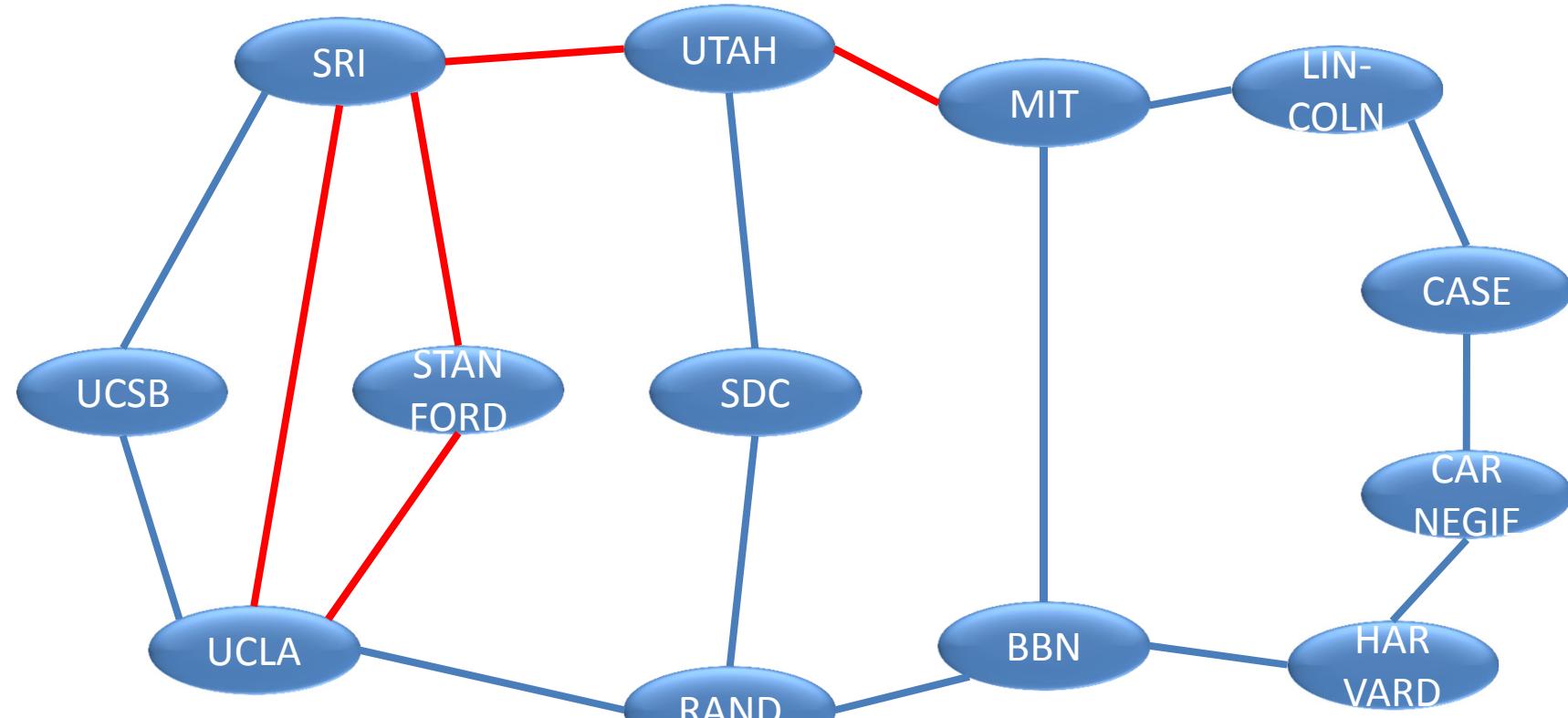


# Path

- A **path** in a graph is a sequence of nodes with the property that each consecutive pair in the sequence is connected by an edge.
- E.g. the sequence of nodes MIT, BBN, RAND, UCLA
- A path can have repeat nodes; e.g. SRI, STAN, UCLA, SRI, UTAH, MIT
- Path without repeat nodes is a **simple path**



# Example: Arpanet



SRI, STAN, UCLA, SRI, UTAN, MIT



# Length and distance

- The **length** of a path is the number of steps it contains from beginning to the end
- The length of the path “MIT, BBN, RAND, UCLA” is 3.
- The **distance** between two nodes in a graph is the length of the shortest path between them.

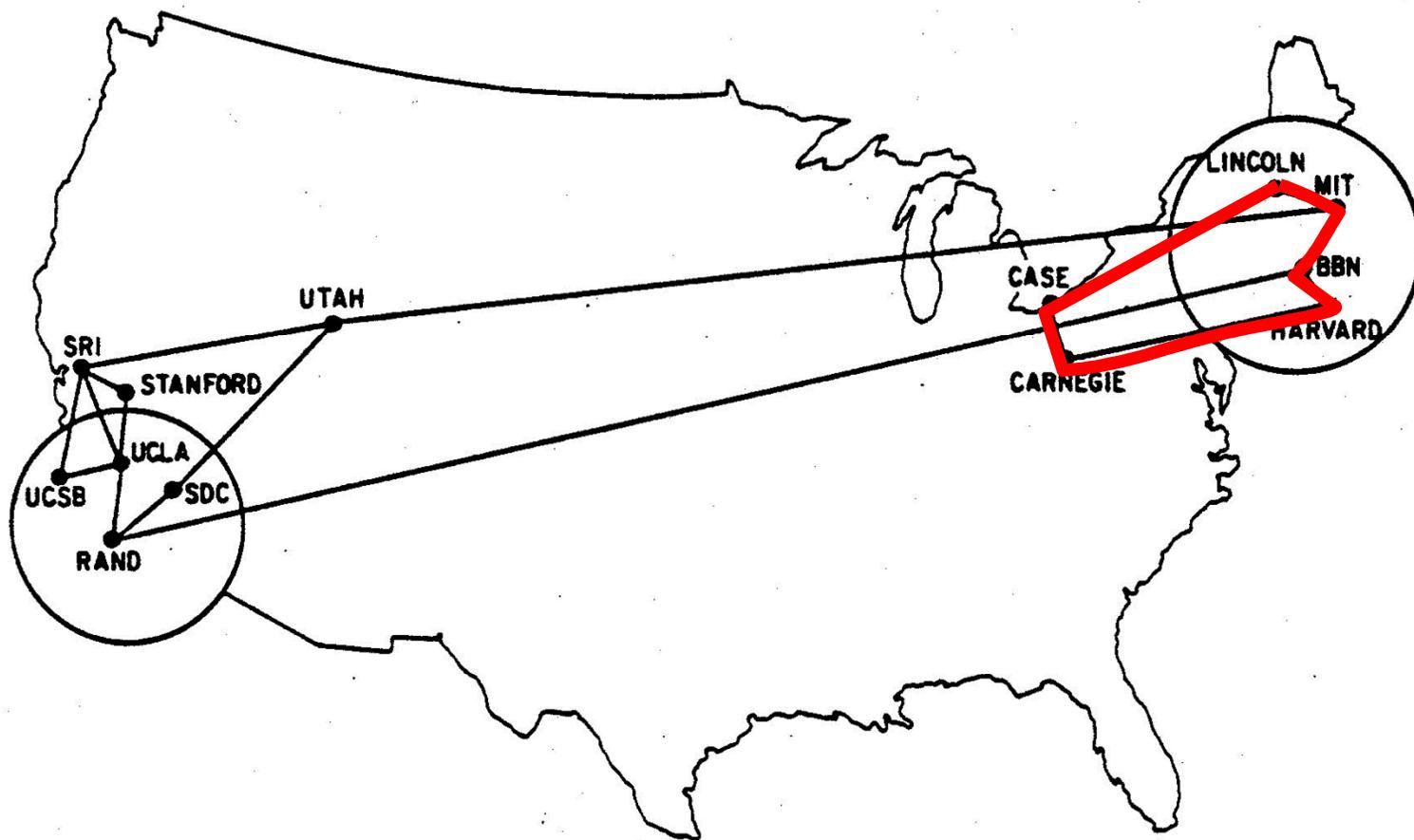


# Cycle

- Informally a “ring” structure
- A **cycle** is a path with at least three edges, in which the first and last nodes are the same, but otherwise all nodes are distinct.
- E.g. LIN, CASE, CARN, HARV, BBN, MIT, LIN

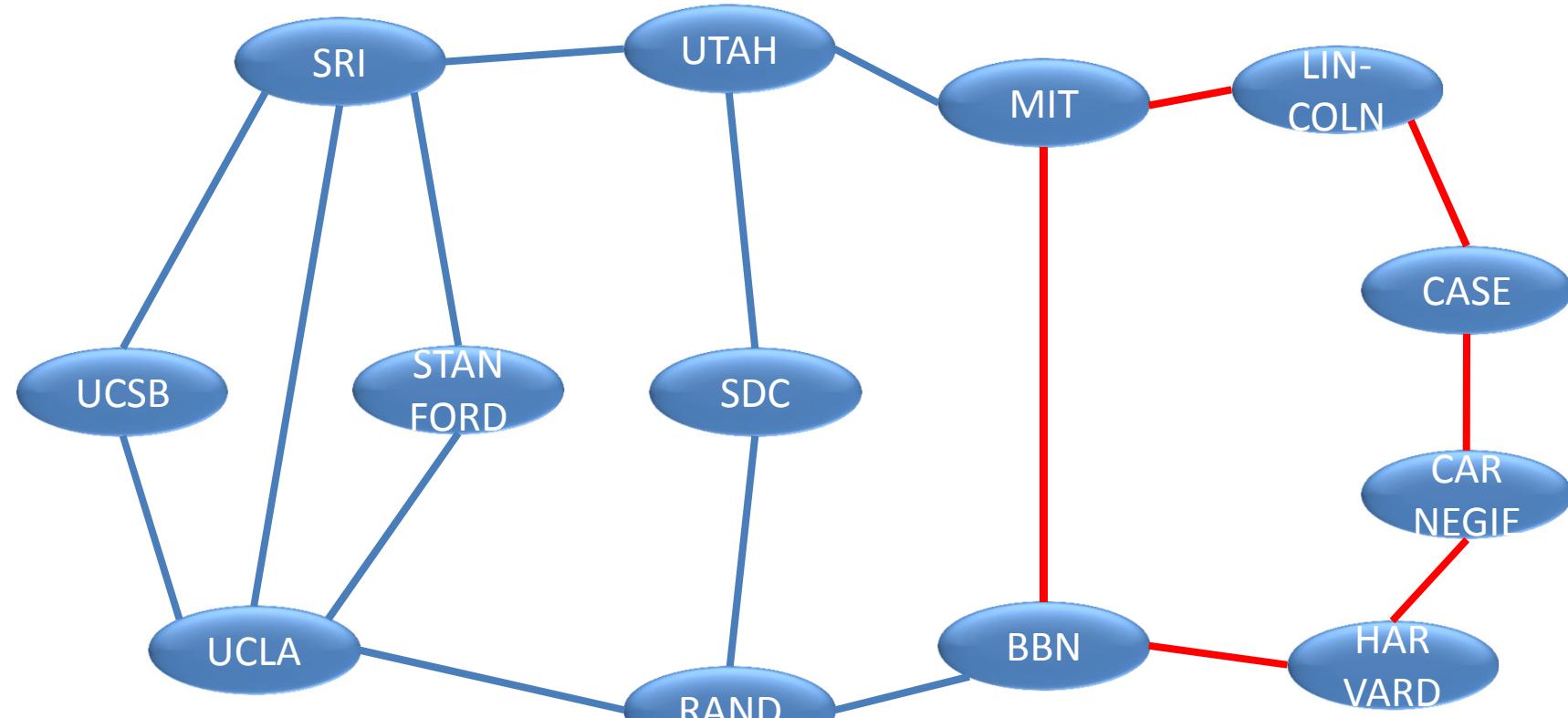


# Arpanet





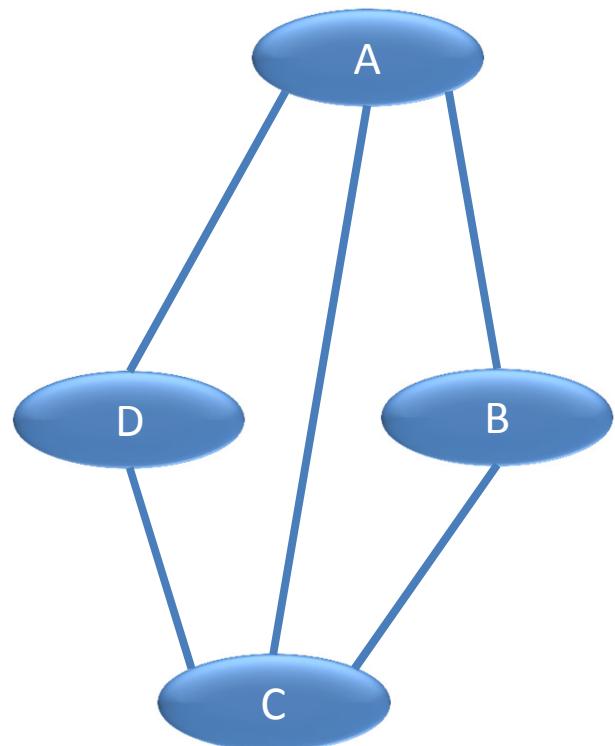
# Example: Arpanet



LIN, CASE, CARN, HARV, BBN, MIT, LIN



# Degree (undirected graph)

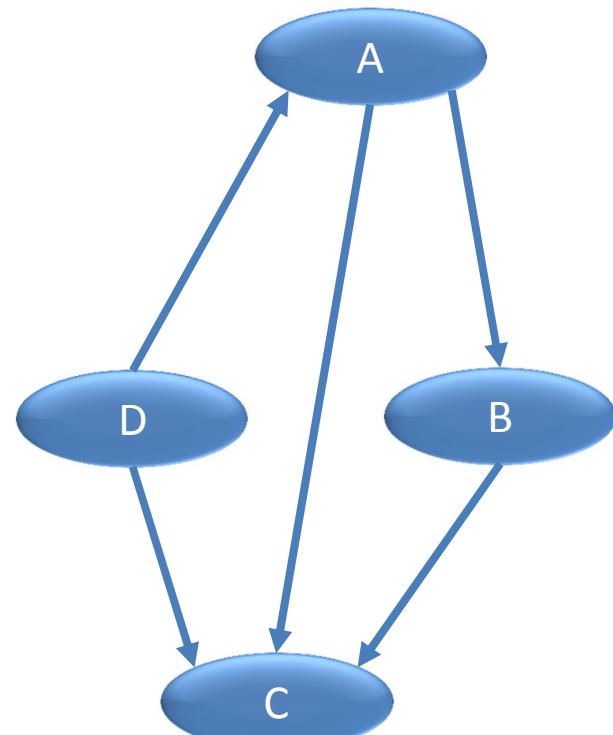


- The degree of a vertex of a graph is the number of edges incident on the vertex.

| vertex | degree |
|--------|--------|
| A      | 3      |
| B      | 2      |
| C      | 3      |
| D      | 2      |



# In-degree (directed graph)

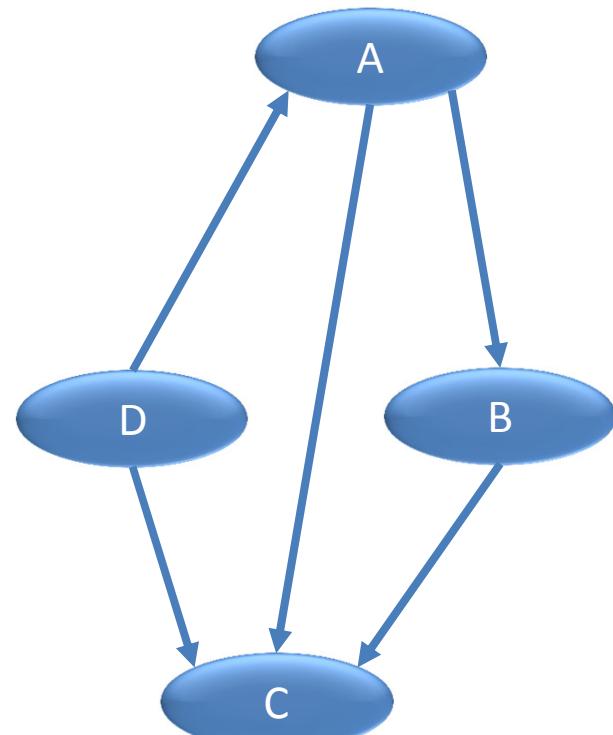


- The in-degree of a vertex of a graph is the number of edges who point to that vertex.

| vertex | In-degree |
|--------|-----------|
| A      | 1         |
| B      | 1         |
| C      | 3         |
| D      | 0         |



# Out-degree (directed graph)



- The out-degree of a vertex of a graph is the number of edges who originate from that vertex.

| vertex | Out-degree |
|--------|------------|
| A      | 2          |
| B      | 1          |
| C      | 0          |
| D      | 2          |

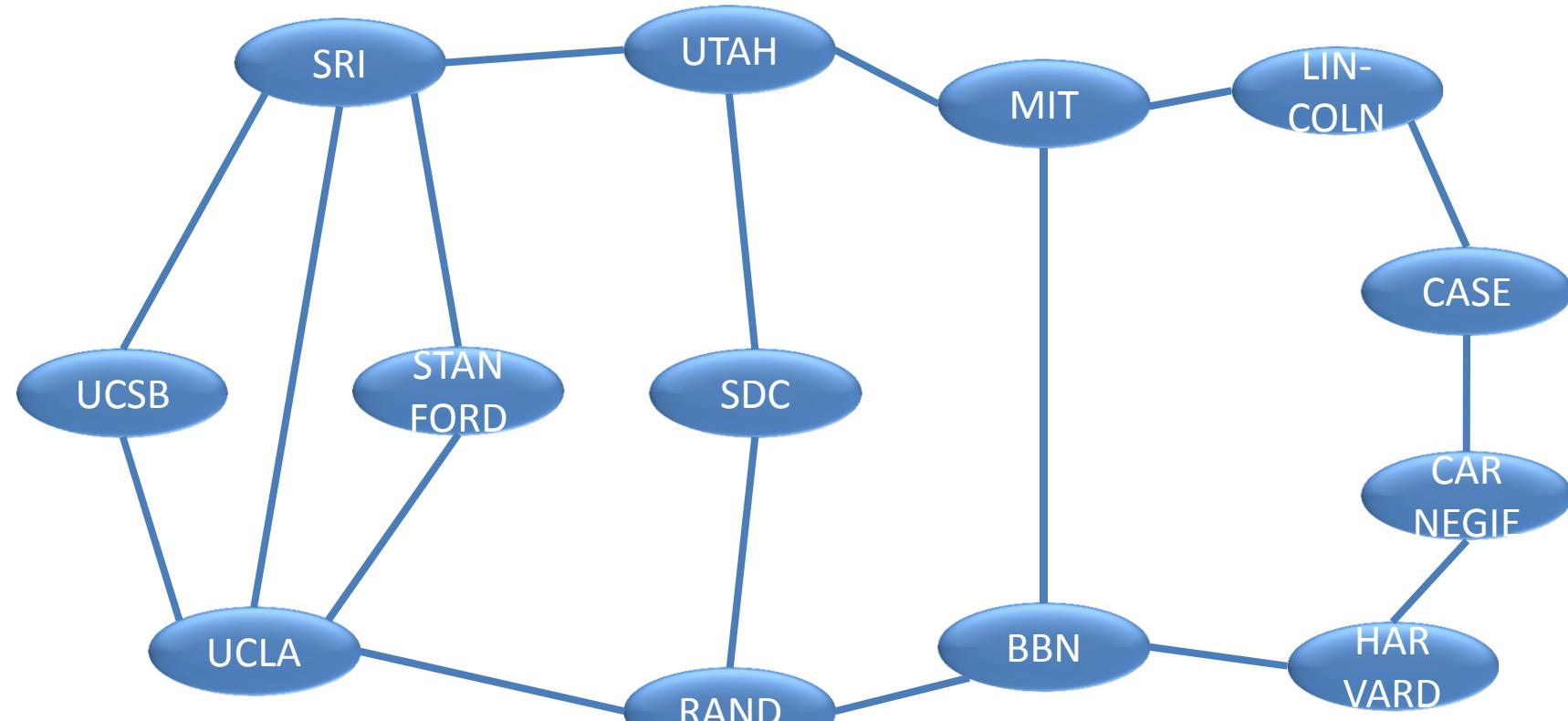


# Connectivity

- A graph is connected if for every pair of nodes, there is a path between them

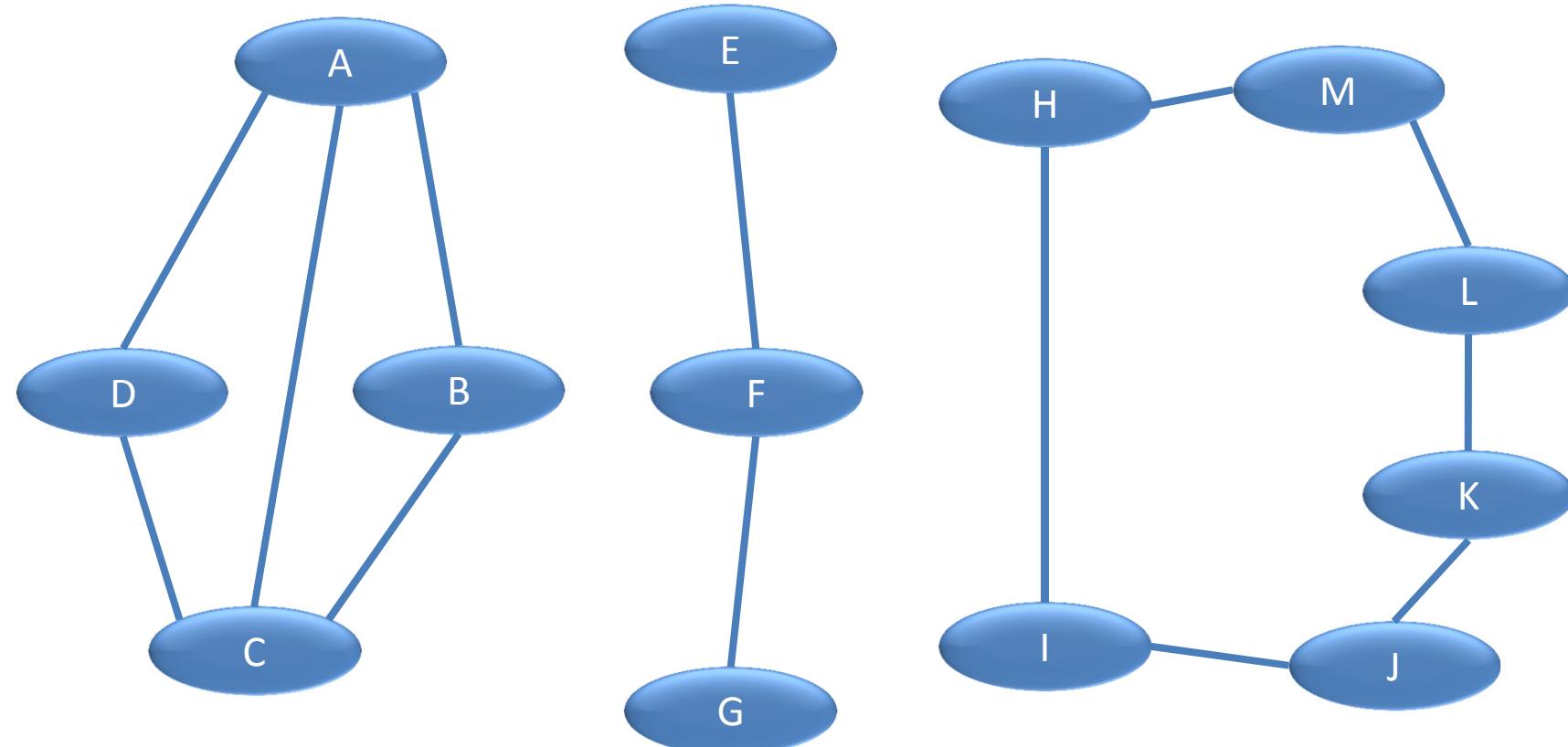


# Connected graph





# Disconnected graph (with three components)





# Connected Component

- A **connected component** (or **component**) of a graph is a sub-set of the nodes such that
  - every node in the subset has a path to every other
  - the subset is not part of some larger set with the property that every node can reach every other



# Giant Component

- Giant component : a connected component that contains a significant fraction of all the nodes.
- Social networks are large, complex networks.

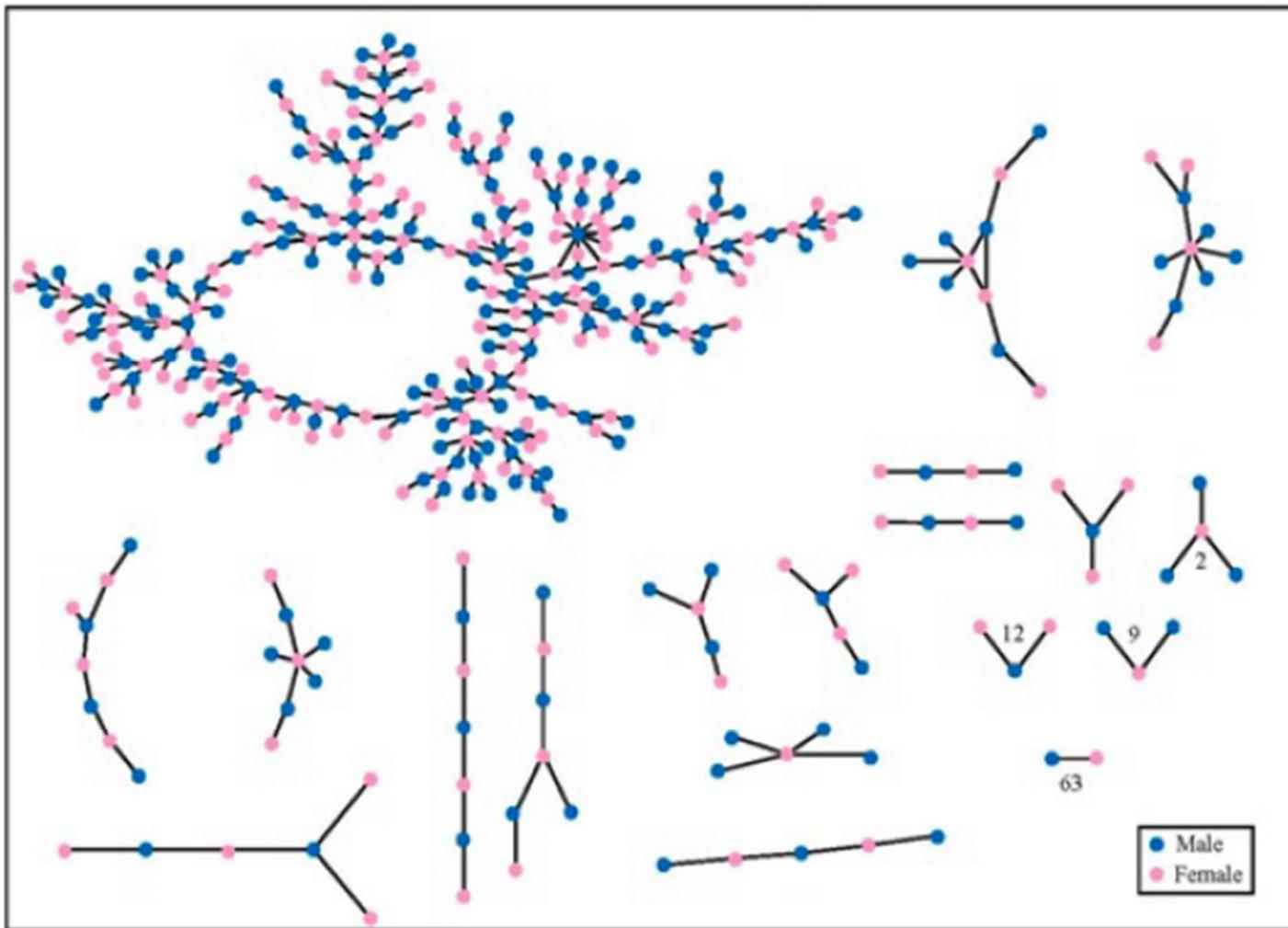
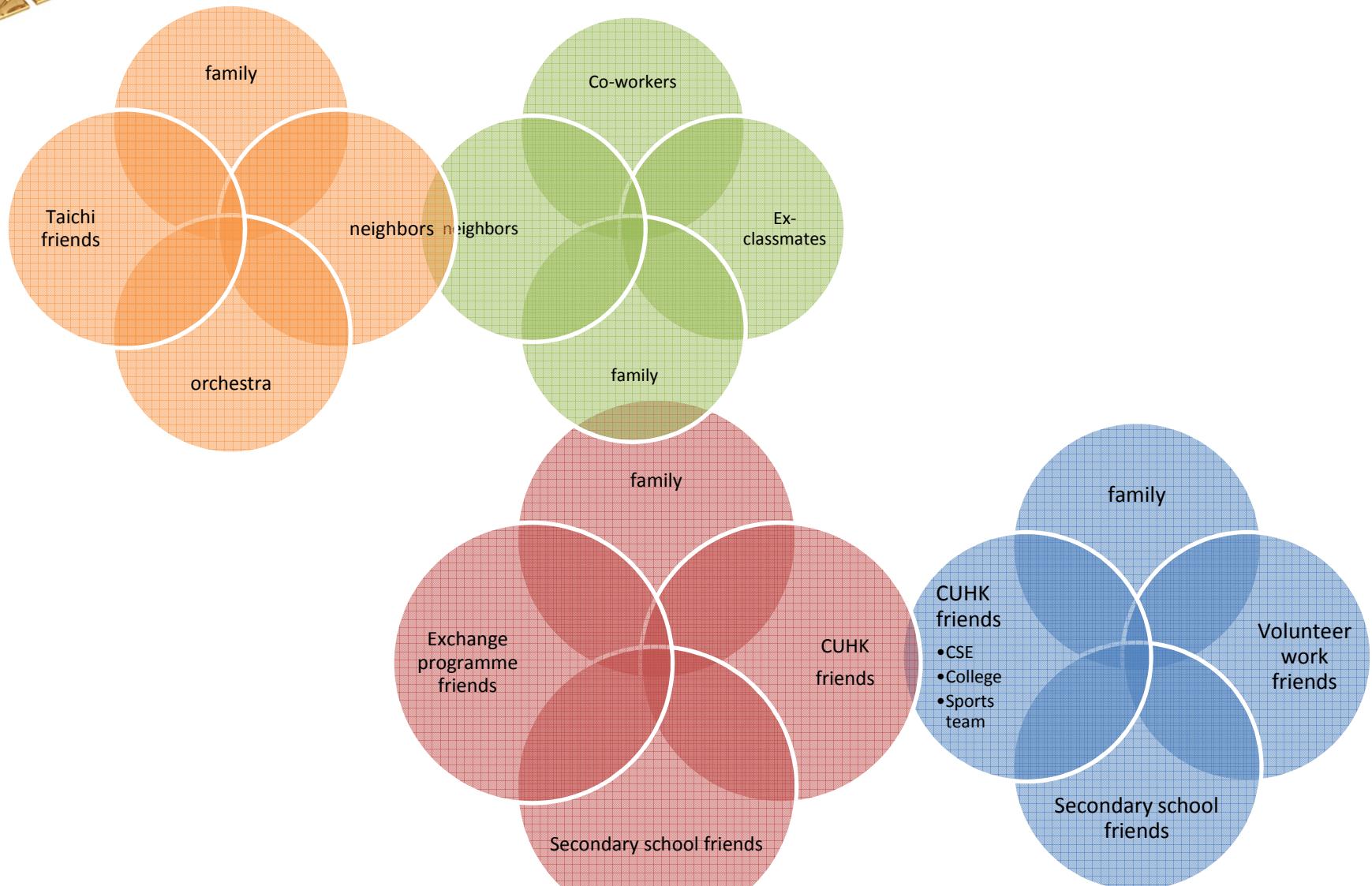


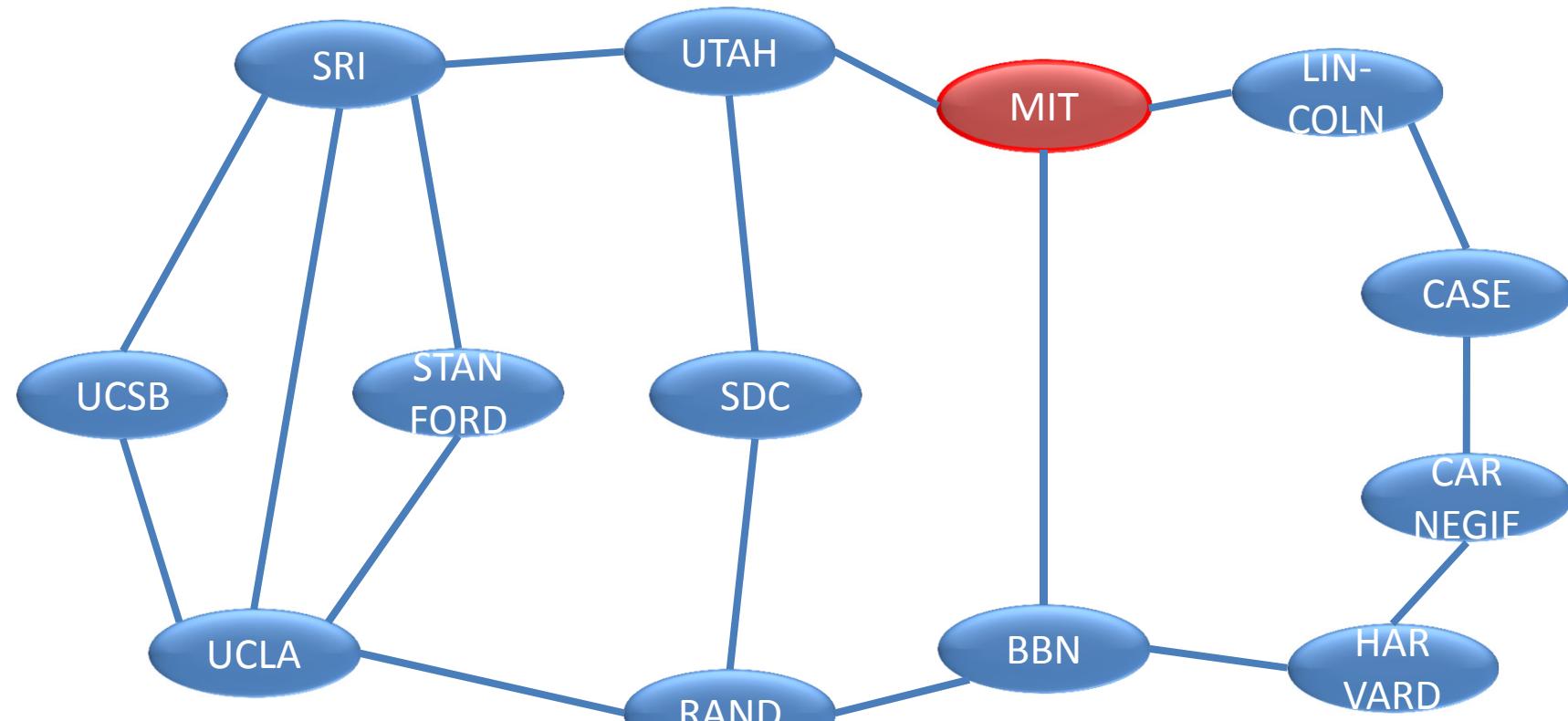
Figure 2.7: A network in which the nodes are students in a large American high school, and an edge joins two who had a romantic relationship at some point during the 18-month period in which the study was conducted [49].





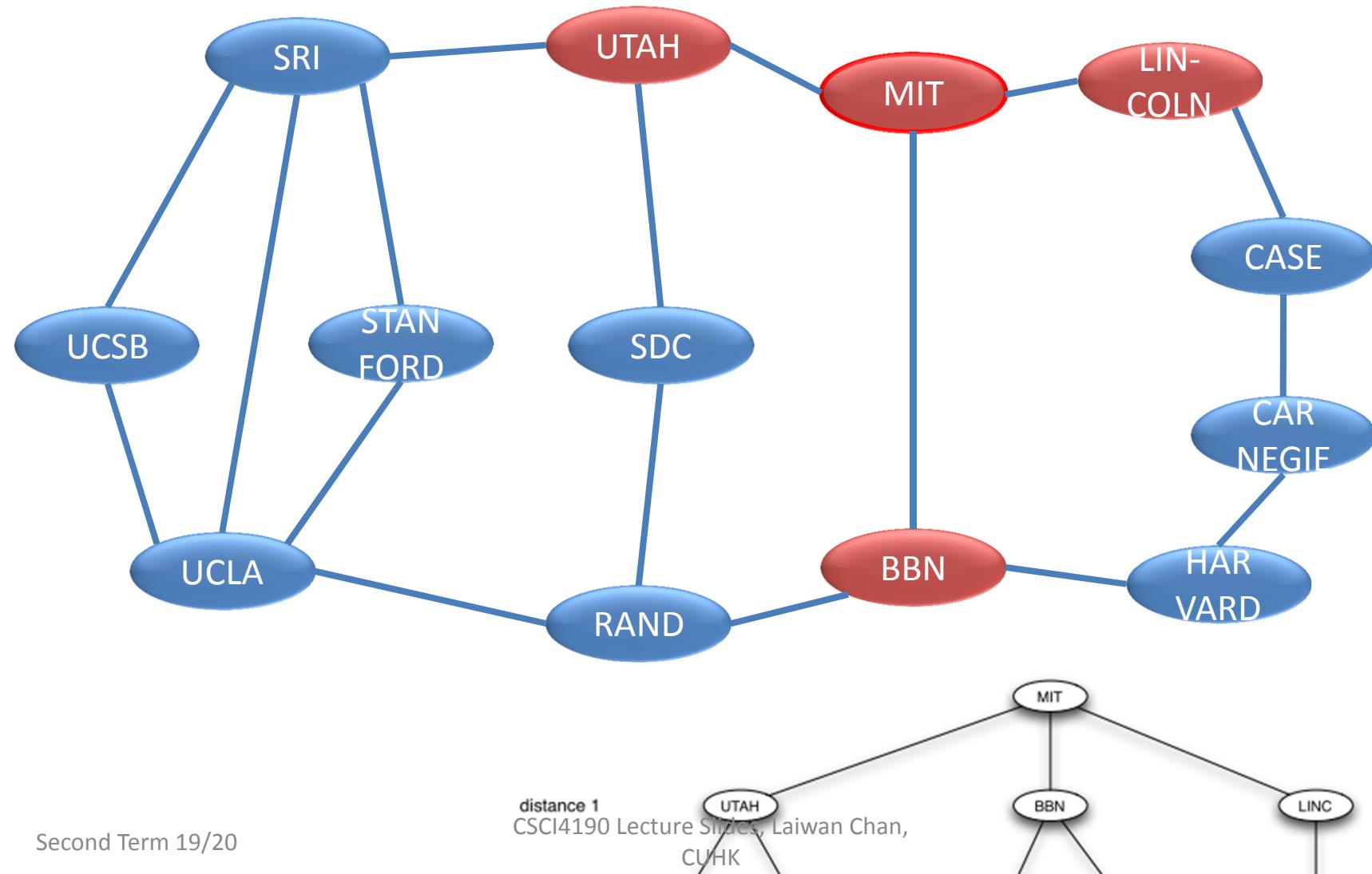
# Breadth-first-search

for finding the distance between two nodes



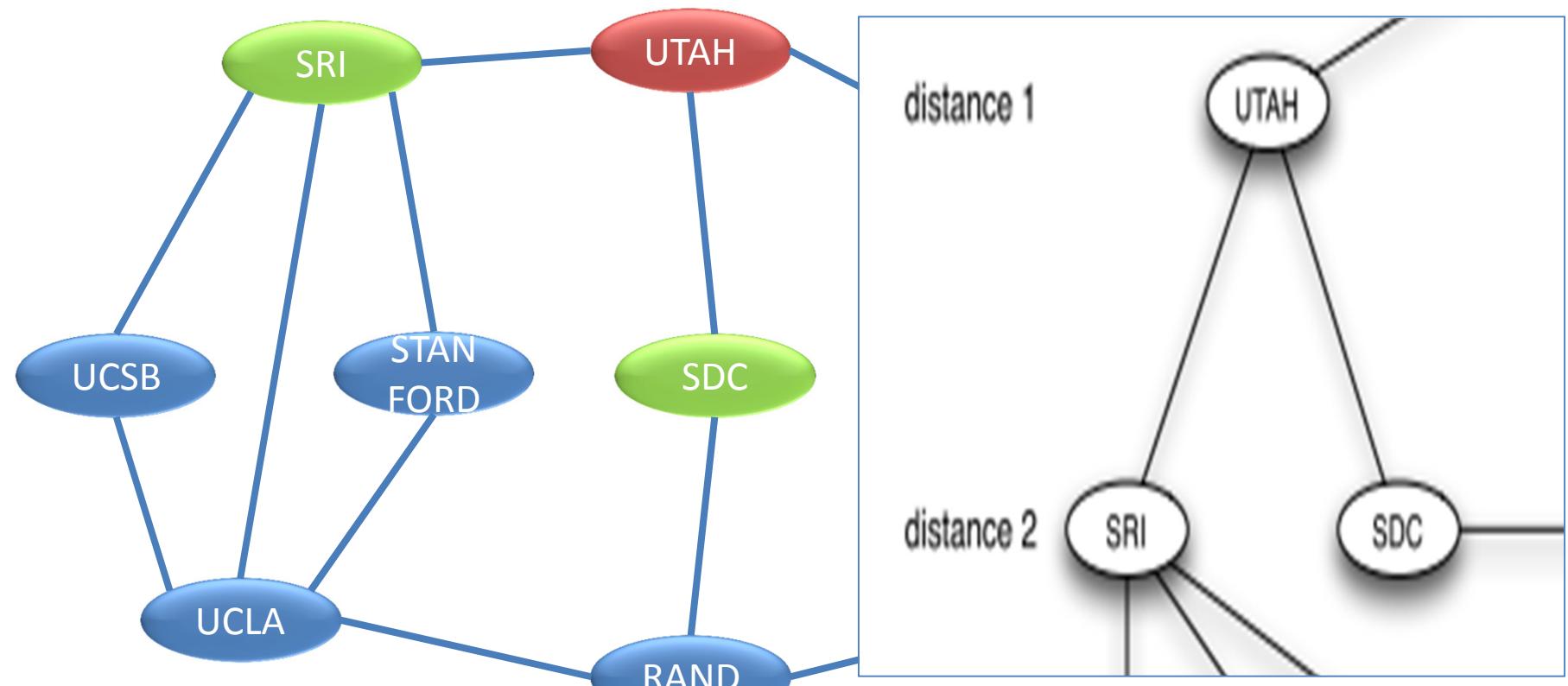


# Breadth-first-search



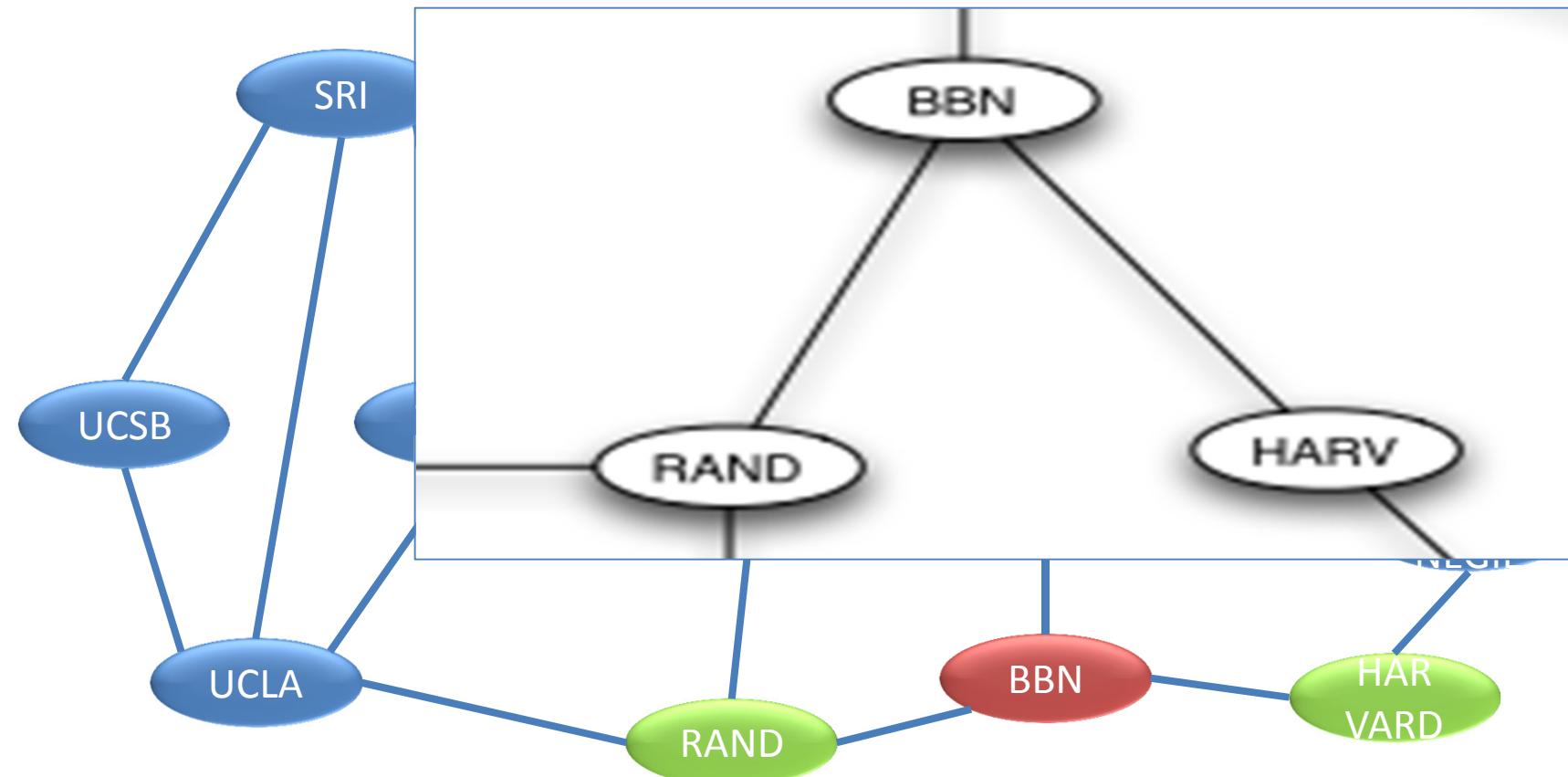


# Breadth-first-search



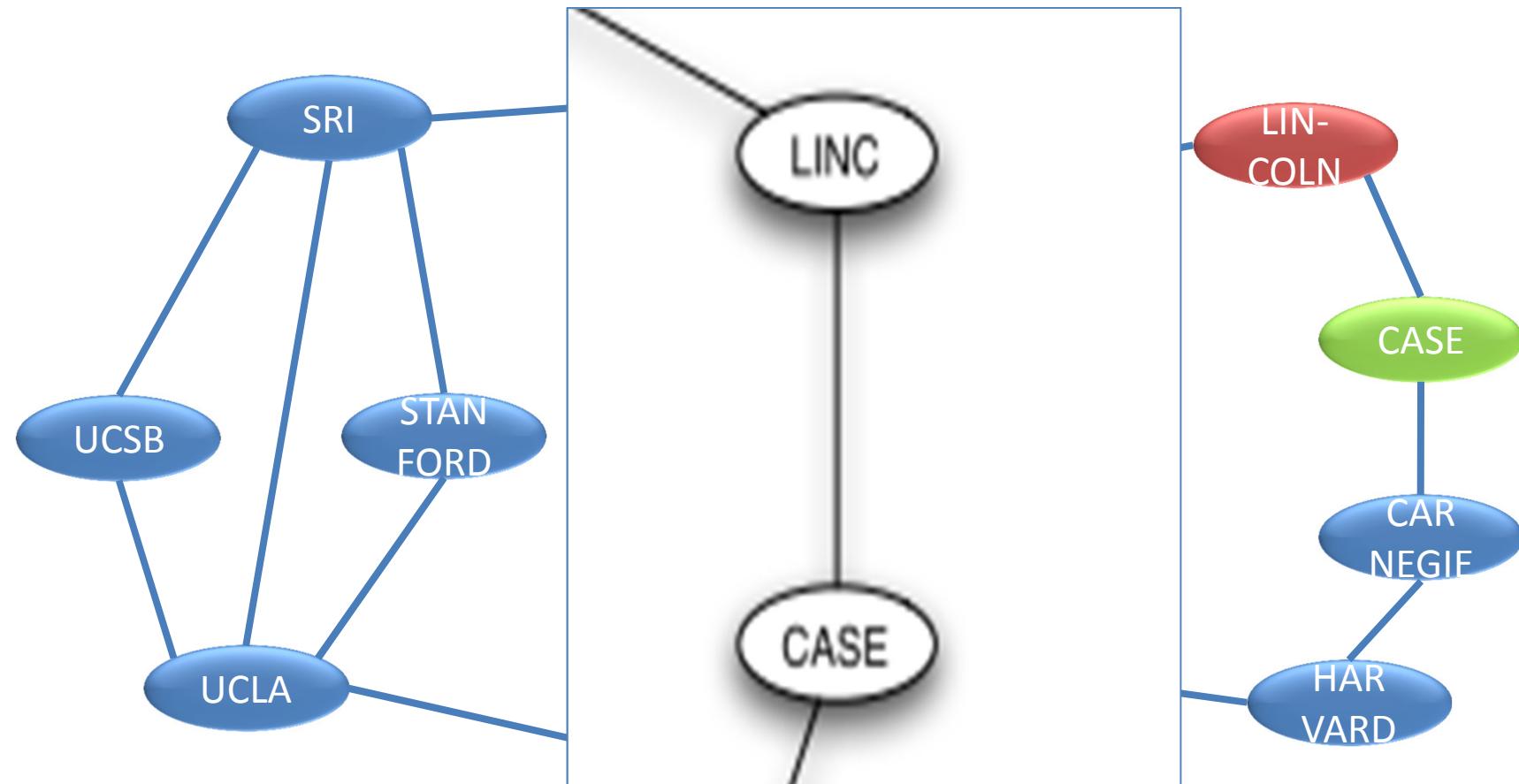


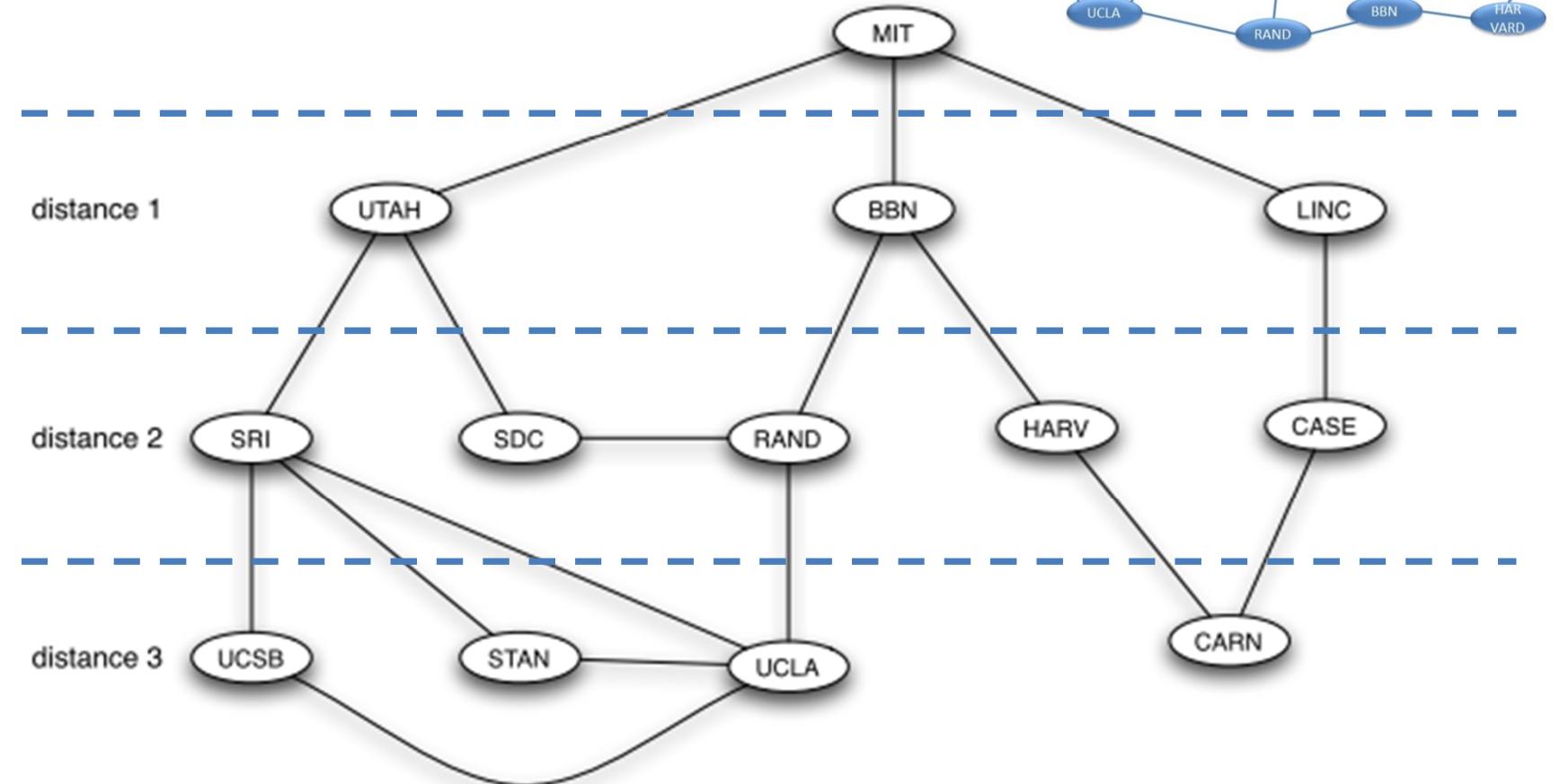
# Breadth-first-search

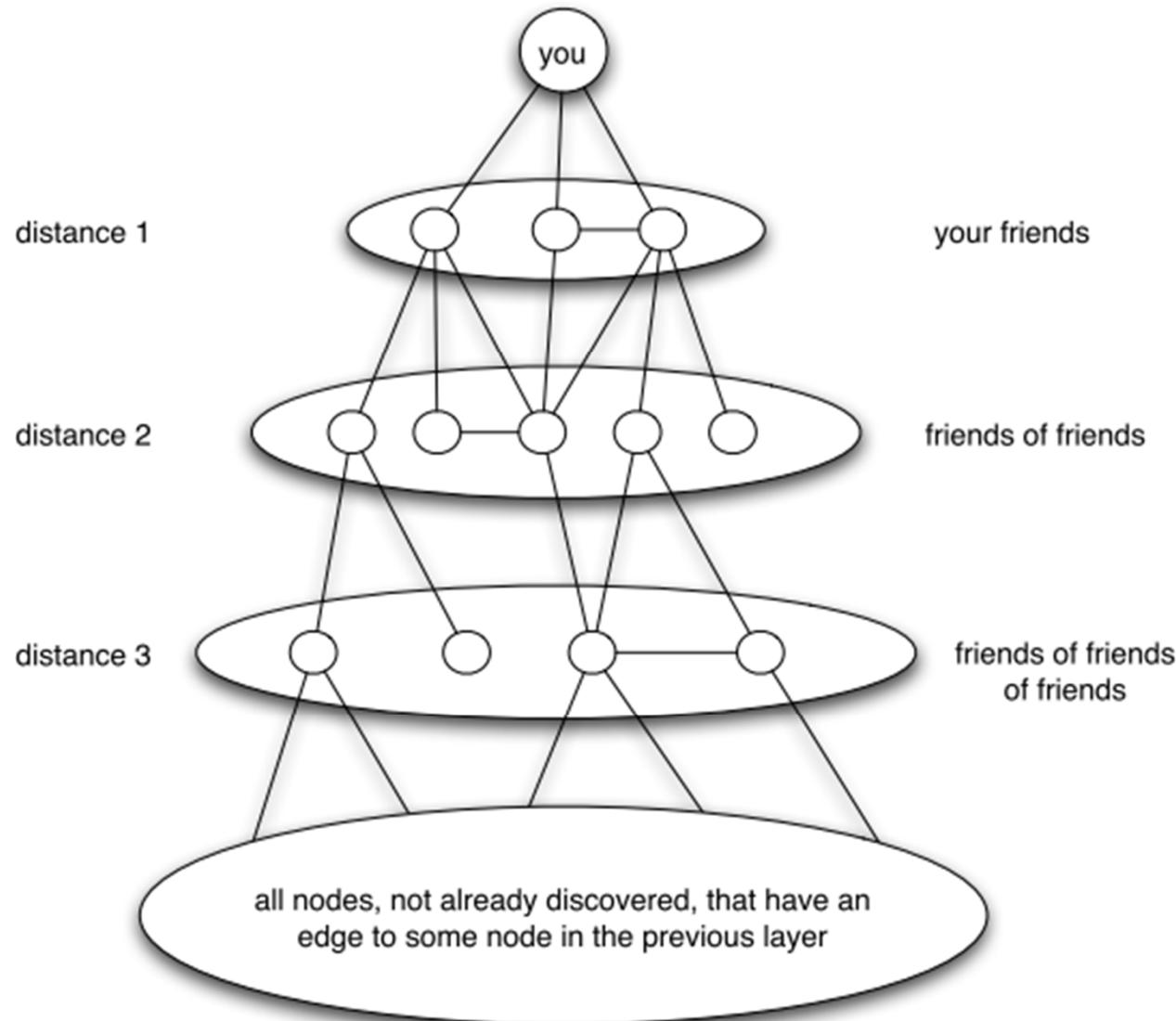




# Breadth-first-search









- Social networks tend to have very short paths between essentially arbitrary pairs of people.

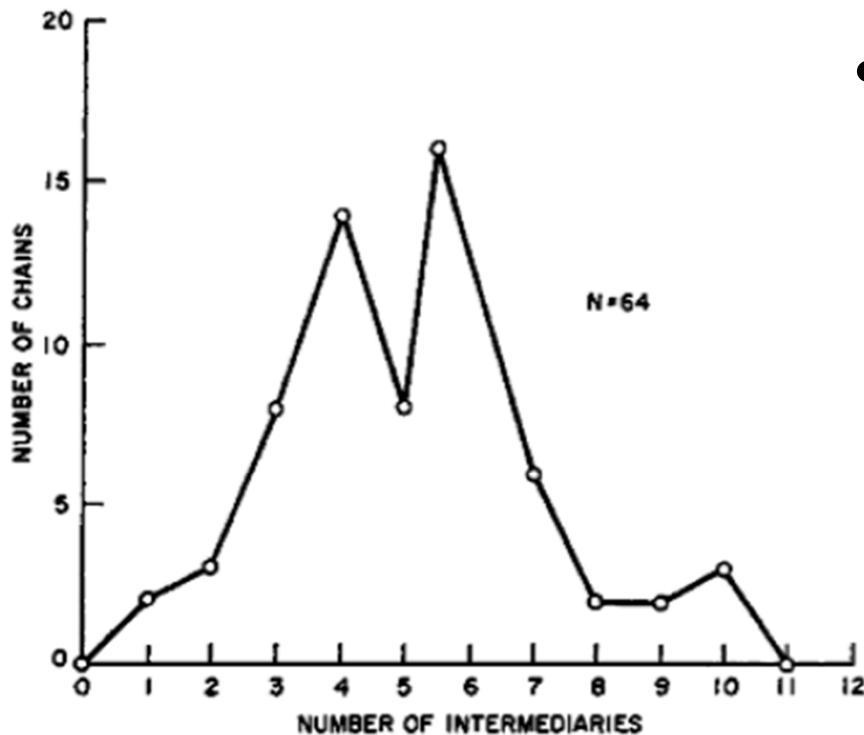


# Small world phenomenon

- Six degrees of separation
- “I read somewhere that everybody on this planet is separated by only six other people.”
- Experiments
  - Letter forwarding (Travers and Milgram)
  - 240 million active users accounts on Microsoft Instant Messenger (Jure Leskovec and Eric Horvitz)
  - Erdös number : Collaboration networks
  - Bacon number : movie actors and actresses



# Letter forwarding



- 296 randomly chosen “starters” to try forwarding a letter to a “target” person.

Figure 2.10: A histogram from Travers and Milgram’s paper on their small-world experiment [391]. For each possible length (labeled “number of intermediaries” on the  $x$ -axis), the plot shows the number of successfully completed chains of that length. In total, 64 chains reached the target person, with a median length of six.



# Microsoft Instant Messenger

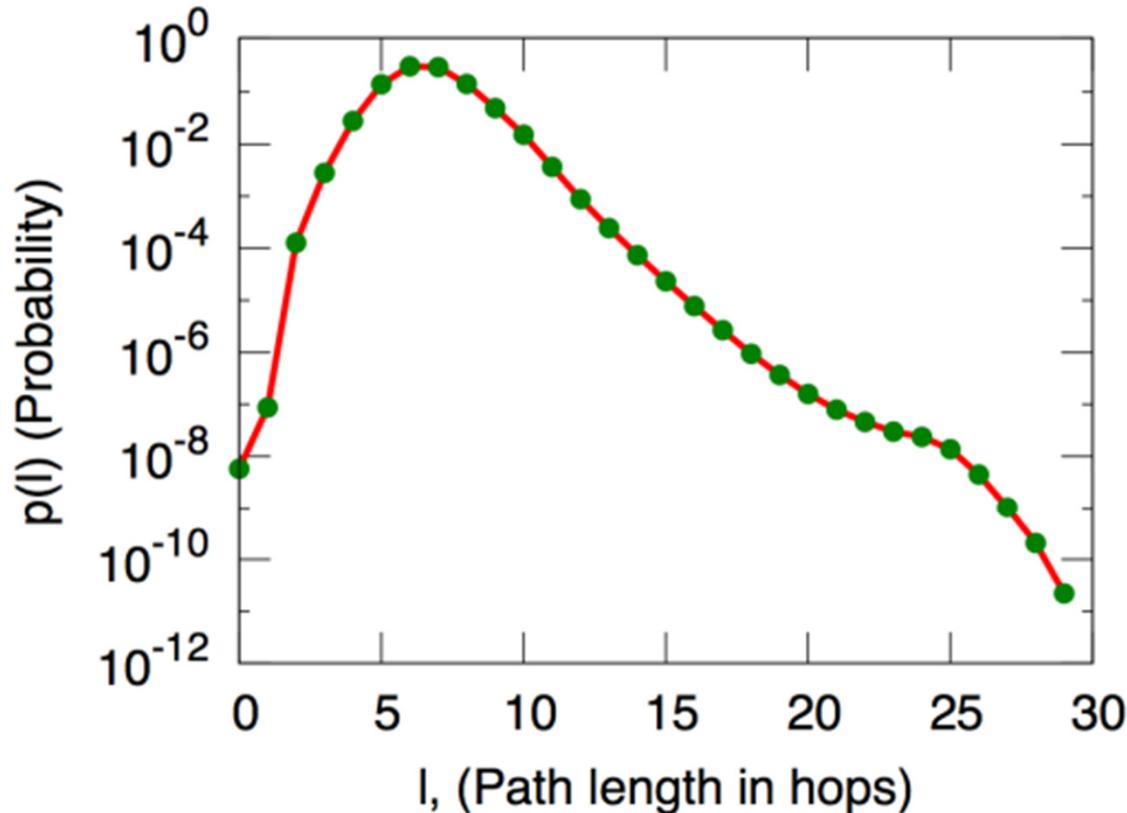


Figure 2.11: The distribution of distances in the graph of all active Microsoft Instant Messenger user accounts, with an edge joining two users if they communicated at least once during a month-long observation period [273].

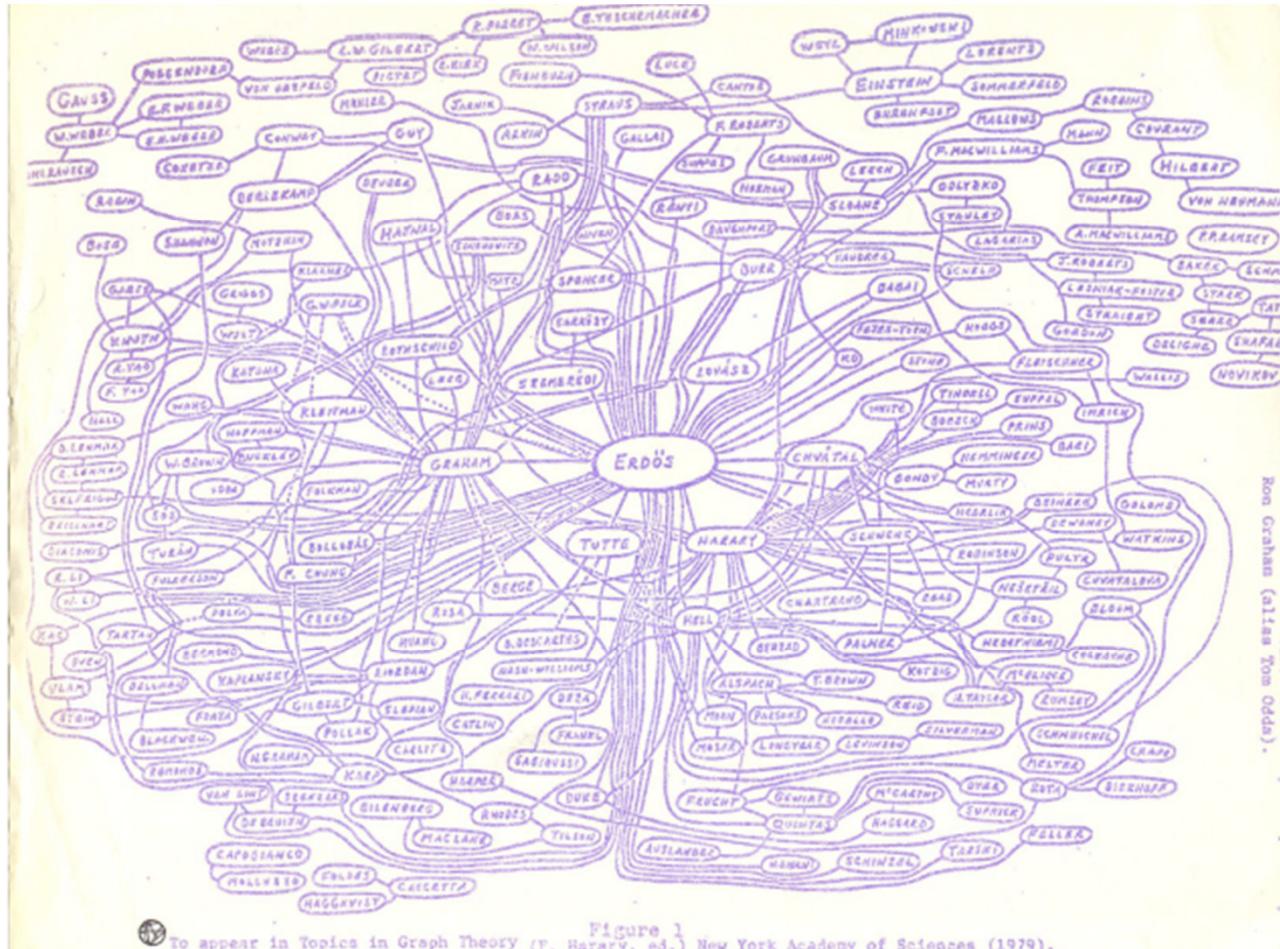
An estimated average distance of 6.6 and an estimated median of 7.



# Erdös number

- Paul Erdös
- A person's Erdös number is the distance from him or her to Erdös in the collaboration graph

|                 | Erdös number |
|-----------------|--------------|
| Albert Einstein | 2            |
| Enrico Fermi    | 3            |
| Noam Chomsky    | 4            |
| Linus Pauling   | 4            |
| Francis Crick   | 5            |
| James Watson    | 6            |



To appear in Topics in Graph Theory (F. Harary, ed.) New York Academy of Sciences (1979).

Figure 2.12: Ron Graham's hand-drawn picture of a part of the mathematics collaboration graph, centered on Paul Erdős [189]. (Image from <http://www.oakland.edu/enp/cgraph.jpg>)



# Bacon number

- Bacon number is his or her distance in this graph to Kevin Bacon
- The average Bacon number, over all performers in the IMDB, is approximately 2.9, and it's a challenge to find one that's larger than 5.