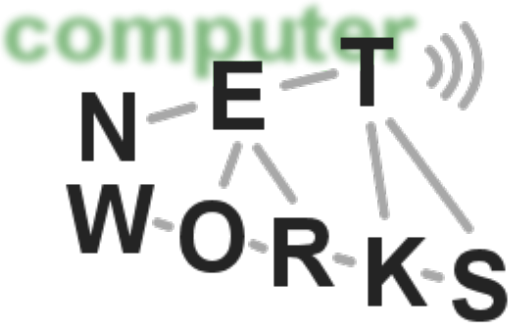
Introduction to Social Networks

Advanced Computer Networks

Summer Semester 2012



computer NETWORKS

Social Network

- Social Network
  - A network made up by a set of individuals interconnecting with each other basing on social relationships (such as friendships, partnerships, etc.)
- Entity: a basic unit of network
- Link: interconnection between entities
- Behavior and dynamics
  - Each individual's actions have implicit consequences for the outcomes of everyone in the system
  - Individual actions are not in isolation: cause-effect
    - Changes in a product, a Web site, or a government program
    - The rich get richer; winners take all; small advantages are magnified to a critical mass; new ideas get attention that becomes viral

Network: friendship

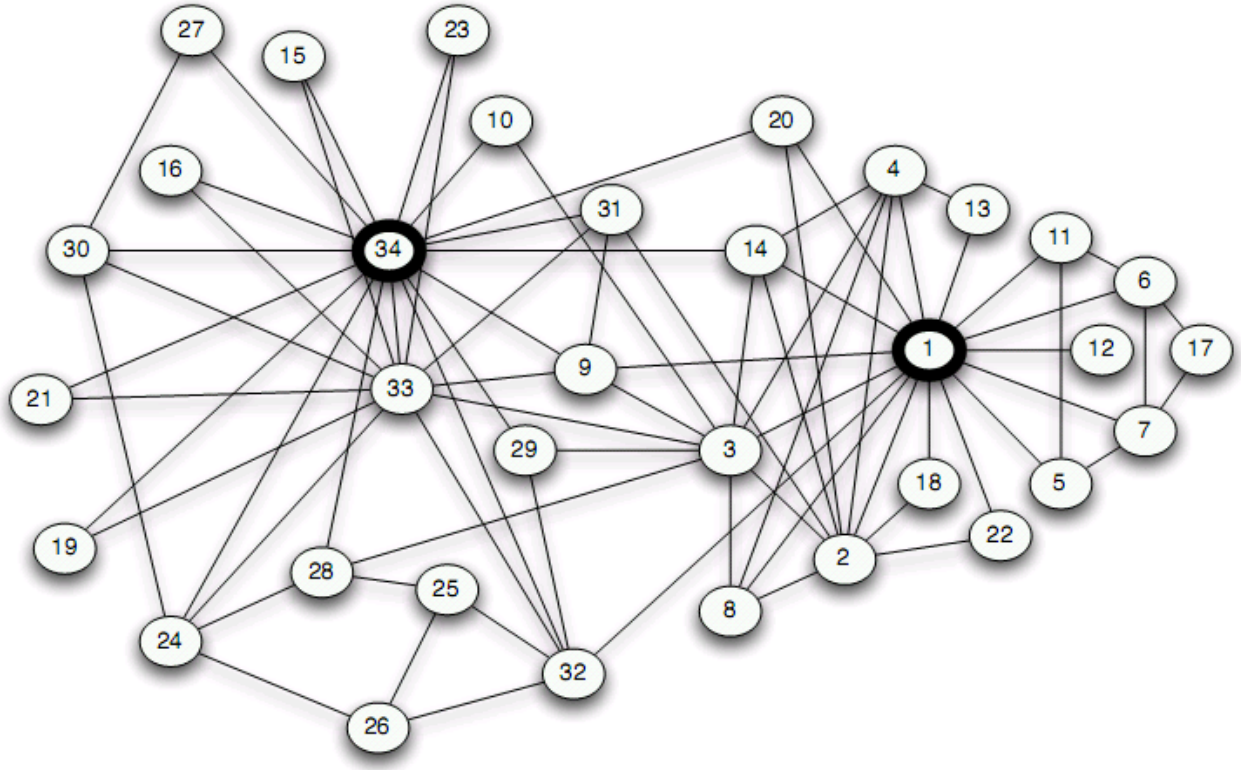
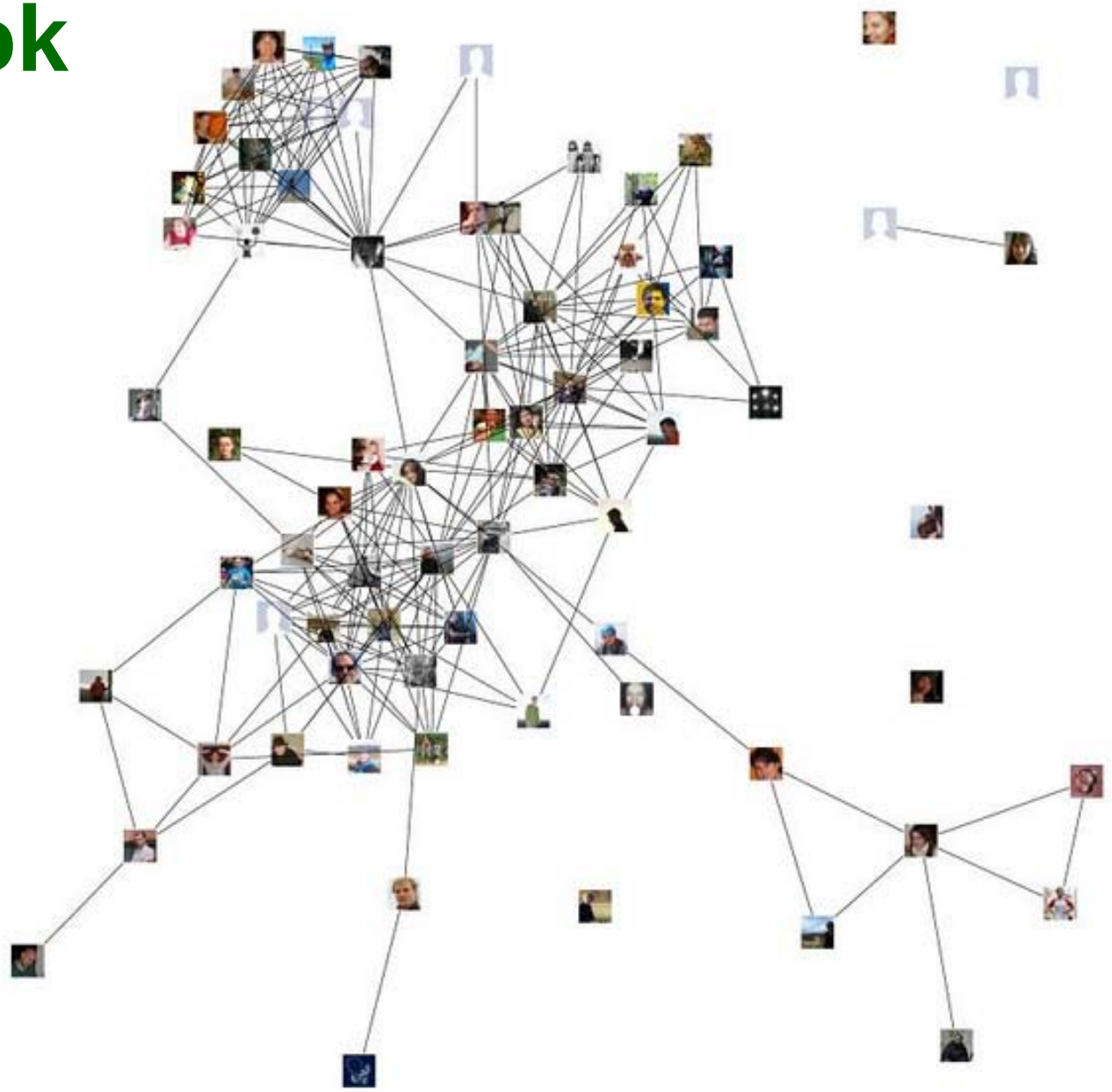


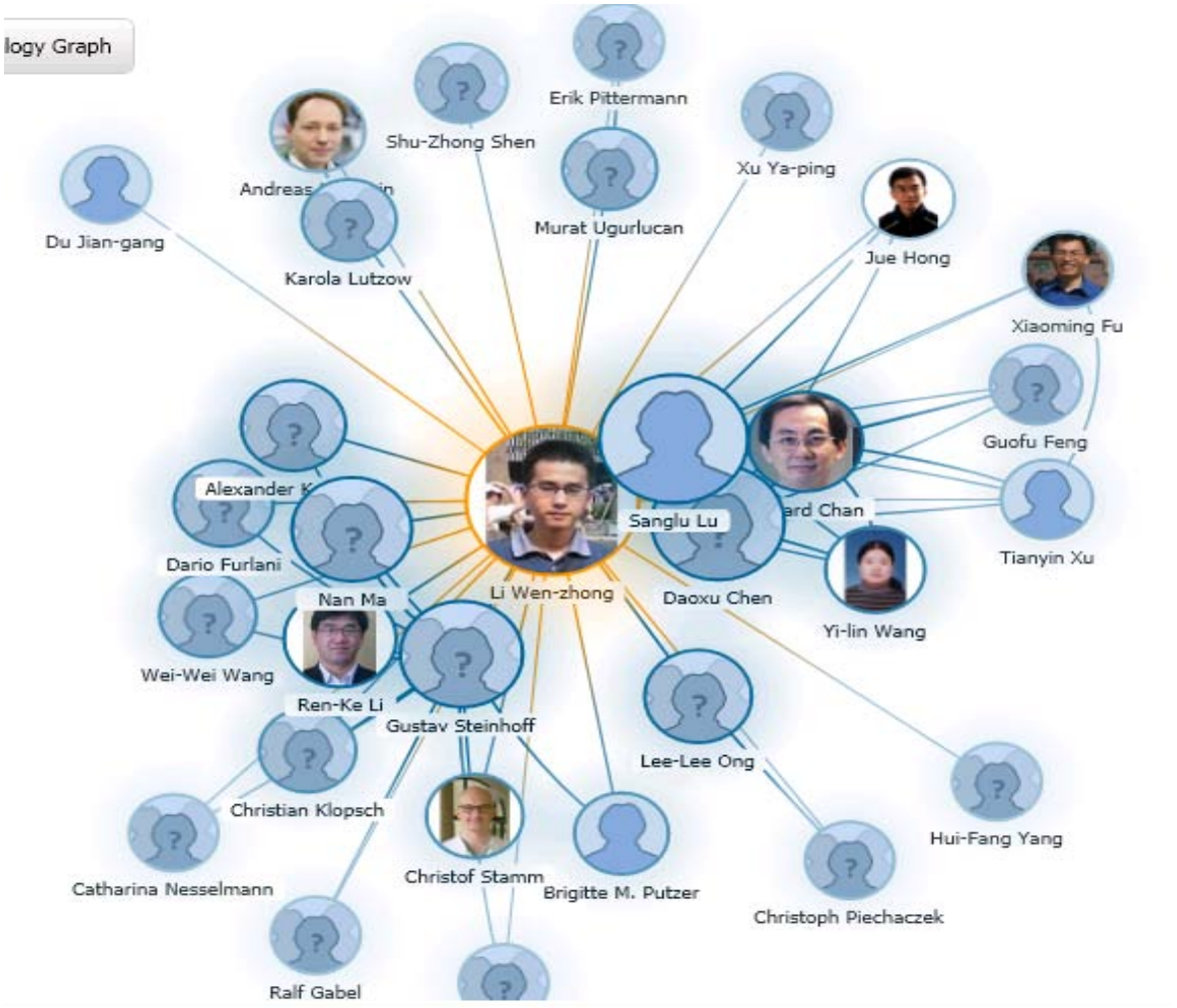
Figure 1.1: The social network of friendships within a 34-person karate club [421].

Facebook

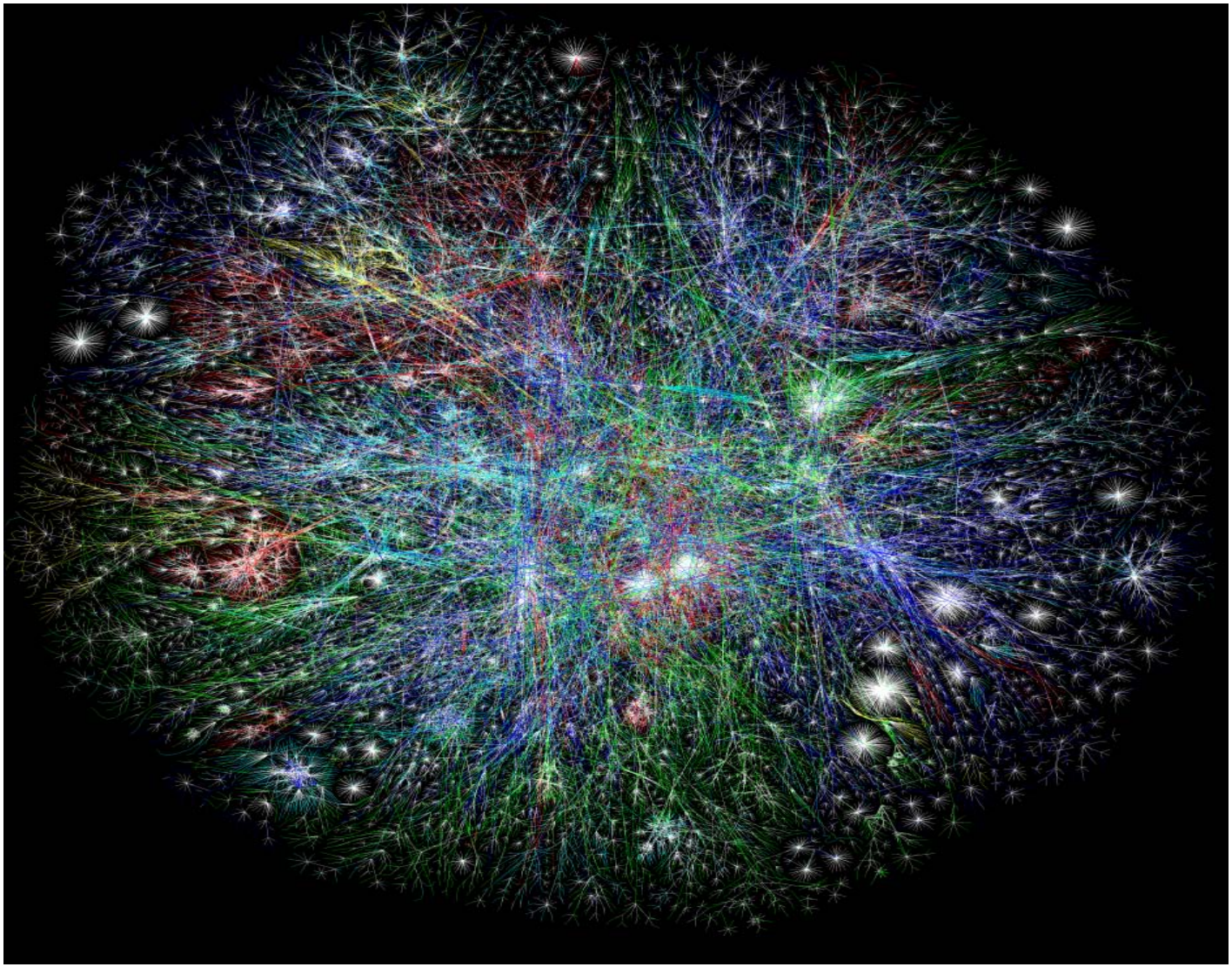


http://revolution-computing.typepad.com/.a/6a010534b1db25970b016760ccd666970b-pi

Network: Co-authorship



Network: Communication



Information

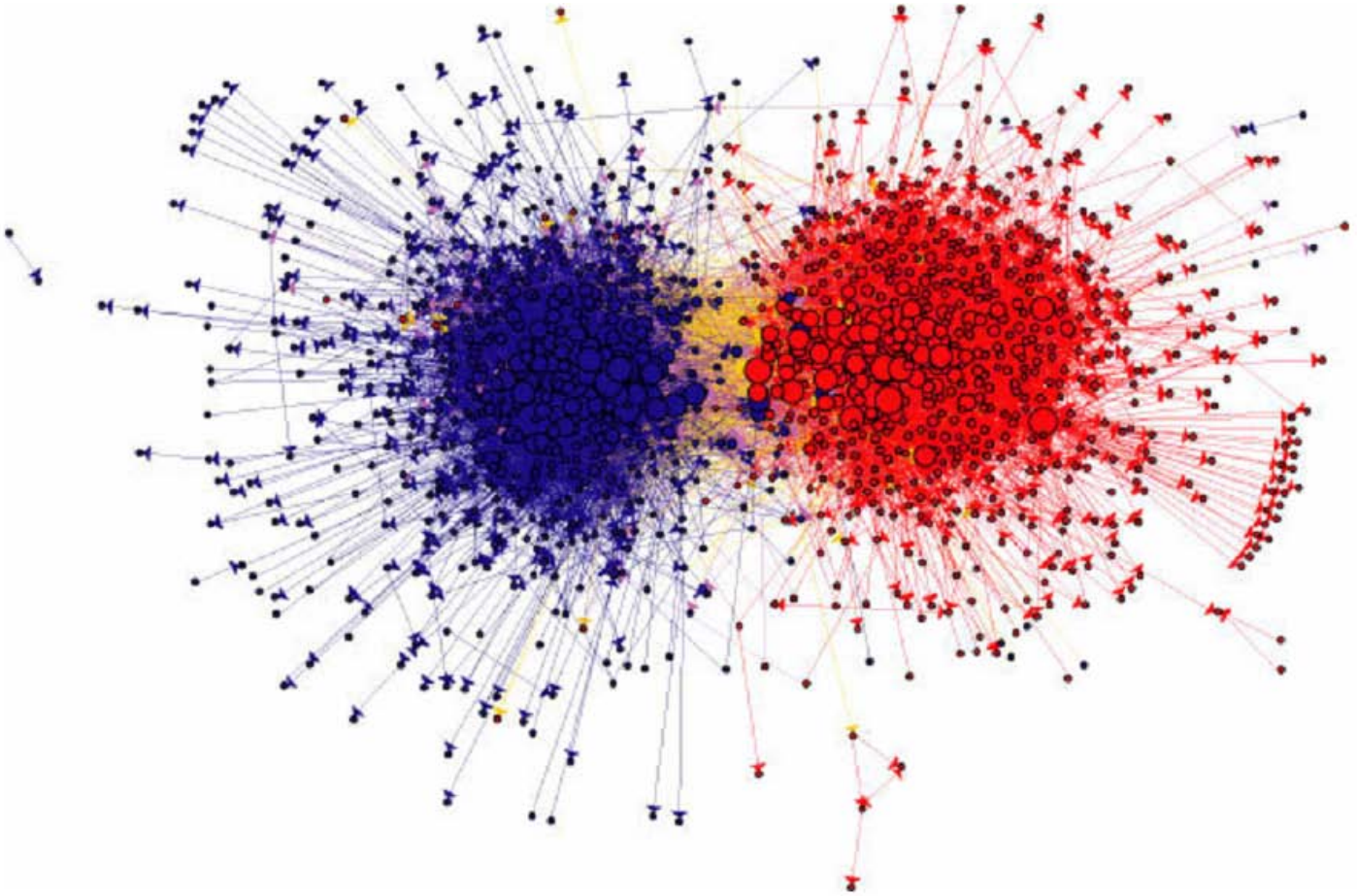
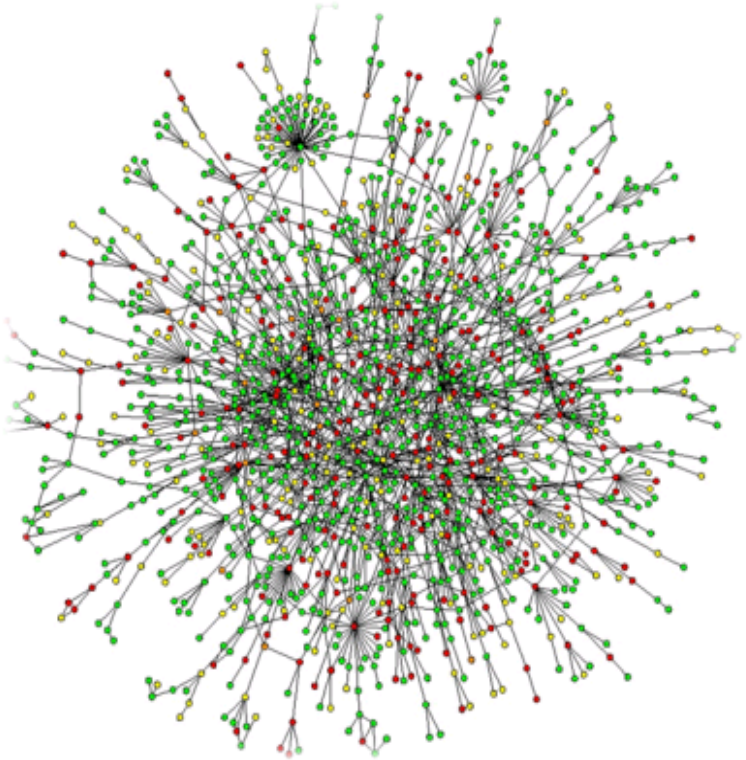


Figure 1.4: The links among Web pages can reveal densely-knit communities and prominent sites. In this case, the network structure of political blogs prior to the 2004 U.S. Presidential election reveals two natural and well-separated clusters [5]. (Image from http://www-personal.umich.edu/ ladamic/img/politicalblogs.jpg)

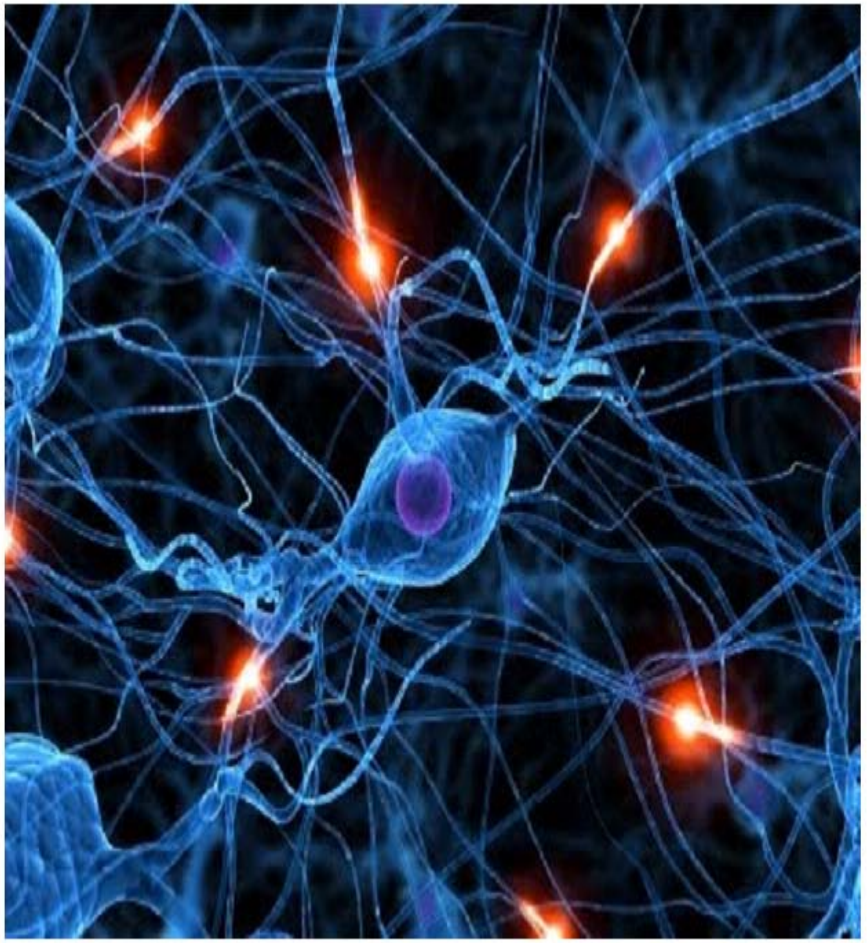
German Rail Network



Protein-Protein Interaction Networks:

Nodes: Proteins

Edges: 'physical' interactoins



Human brain has between
10-100 billion neurons

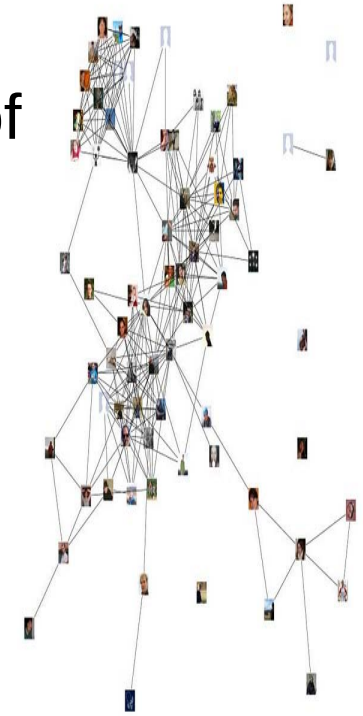
- Something in common
  - A network that defines the interactions between the components
  - Seems random, but display signatures of order and self-organization
- Characteristics
  - Virtual: it is not physically exists
  - Complex: it consists of a large scale number of nodes
  - Grouping: it forms communities due to different interests
  - Dynamic: it’s structure is evolving over time

Social Network Analysis

- Methods
- Empirical:
  - Study network data to find organizational principles
- Mathematical models:
  - Probabilistic, graph theory
- Algorithms
  - Algorithms for analyzing graphs

Research Questions

- Structure and evolution
  - What is the structure of a network?
  - Why and how did it become to have such structure?
- Processes and dynamics
  - Networks provide “skeleton” for spreading of information, behavior, diseases



Targets

- Patterns and statistical properties of network data
- Design principles and models
- Understand why networks are organized the way they are (prediction)

Implications

- Structure of network
  - Connectivity
  - Community
- Power law and small world phenomenon
  - Decentralized search in networks
  - Distributed routing strategies
- Searching the web
  - PageRank
- Epidemics
  - Spreading of disease
  - Information propagation in social networks
- …

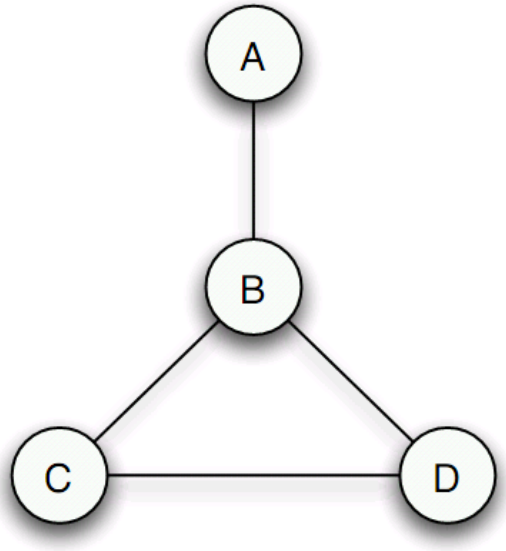
Research Work Done by the Lab

- Cuckoo: Scaling Microblogging Services with Divergent Traffic Demands
- GEMSTONE: Empowering Decentralized Social Networking with High Data Availability
- LENS: Leveraging Social Networking and Trust to Prevent Spam Transmission
- Exploring Regional and Global Population Growth in
- Online Social Networks
- Exploring User Social Behaviors in Mobile Social Applications
- Rethinking Routing Information in Mobile Social Networks: Location-based or Social-based?

Modeling Social Networks

Network as a graph

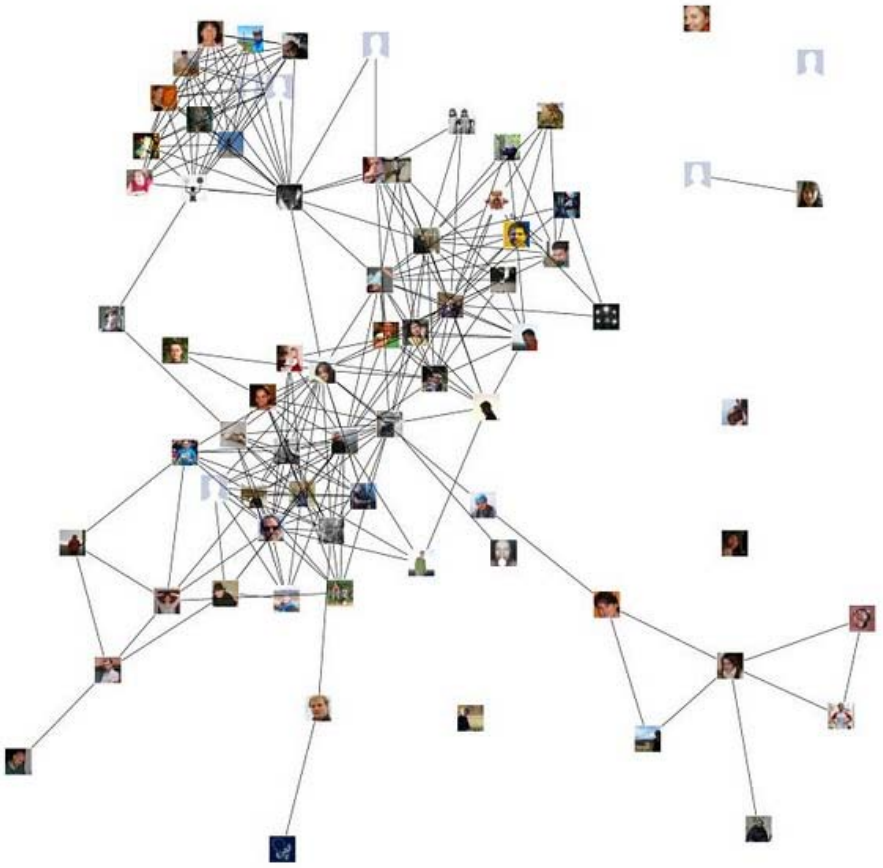
- A network can be represented by a graph mathematically
- Node: an object in the network
- Edge: a link between objects
- Neighbors: nodes connected by edg



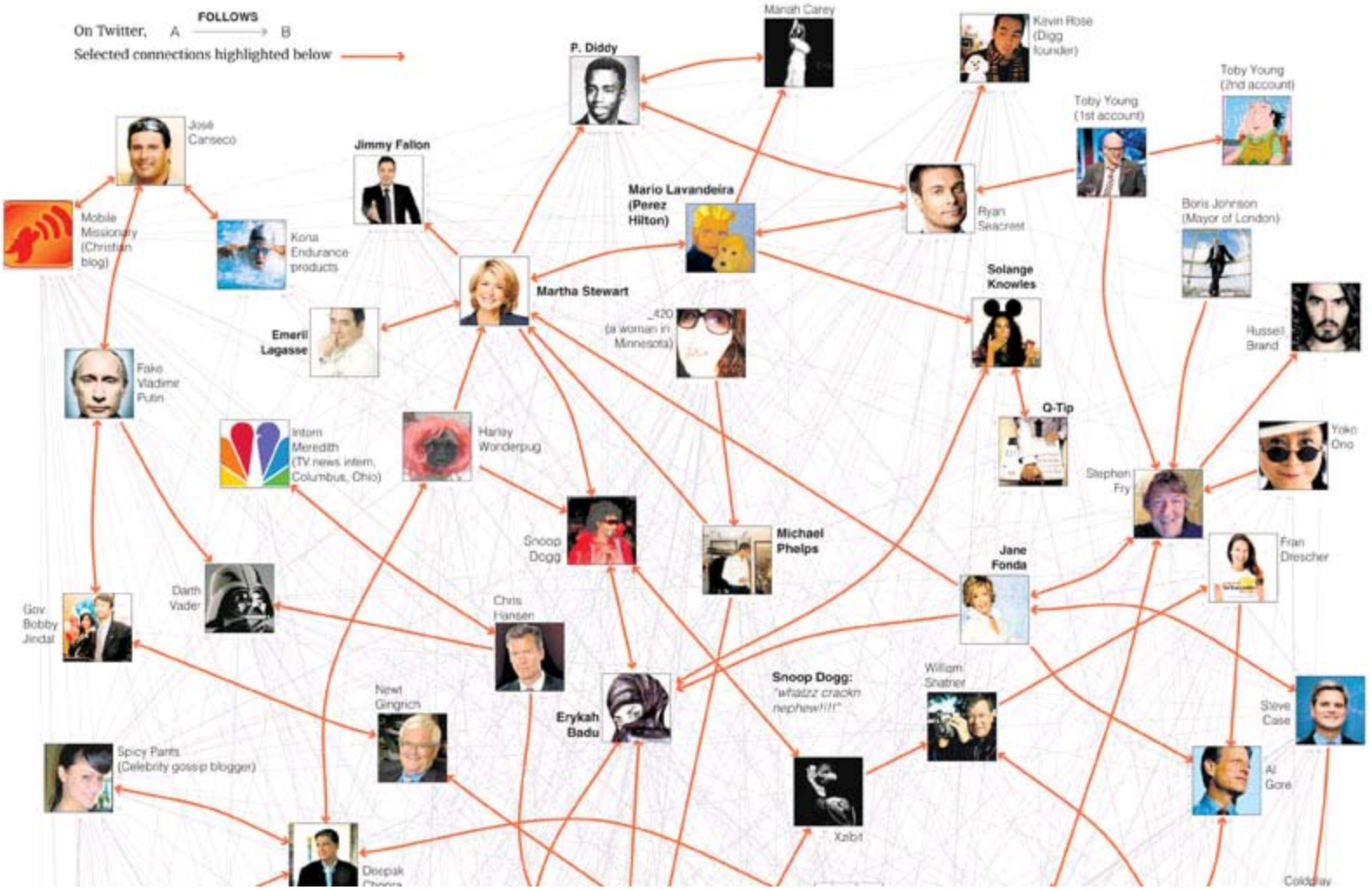
(a) *A graph on 4 nodes.*

Directed Graph

- Facebook friendship network

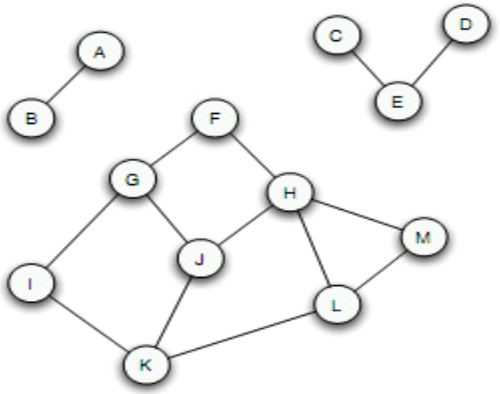


Undirected Graph

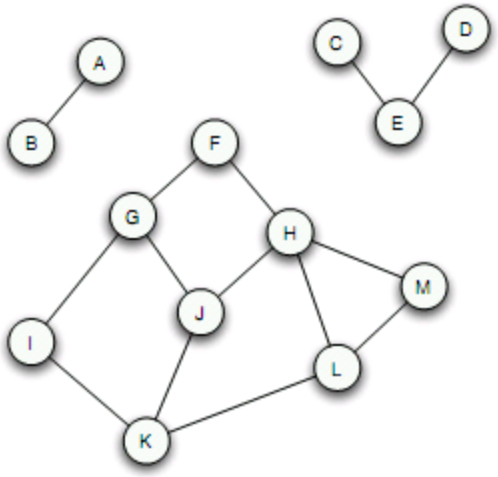


computer NETWORKS

- Path: a sequence of interconnected nodes
- Cycle: a path, the first and last nodes are the same, but other nodes are distinct.
- Connectivity:
  - A graph is connected if for every pair of nodes, there is a path between them



- Components
  - If a graph is not connected, it breaks apart into several connected subgraphs
  - A connected component is a subset of the nodes such that (i) every node in the subset has a path to every other; and (ii) the subset is not part of some larger set with the property that every node can reach every other



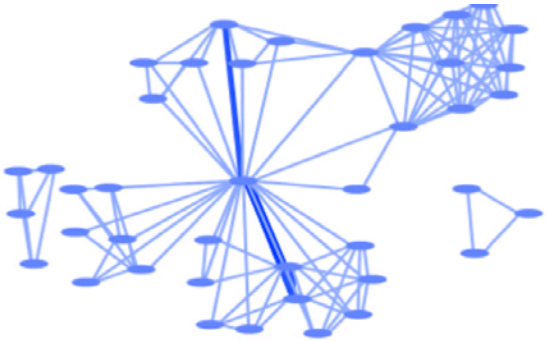
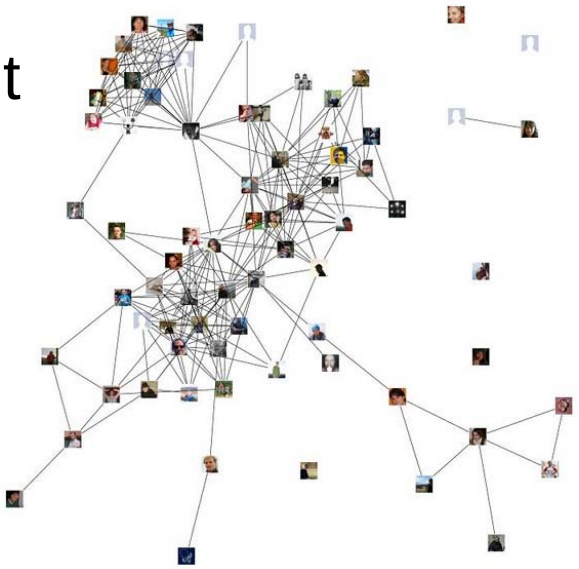


Figure 2.6: The collaboration graph of the biological research center *Structural Genomics of Pathogenic Protozoa (SGPP)* [134], which consists of three distinct connected components. This graph was part of a comparative study of the collaboration patterns graphs of nine research centers supported by NIH's Protein Structure Initiative; SGPP was an intermediate case between centers whose collaboration graph was connected and those for which it was fragmented into many small components.

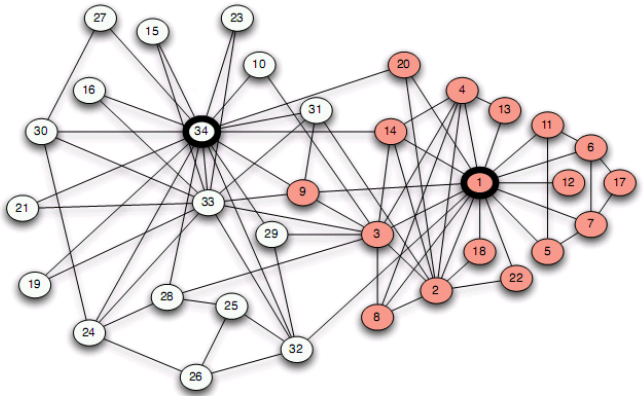
Example: Giant Component

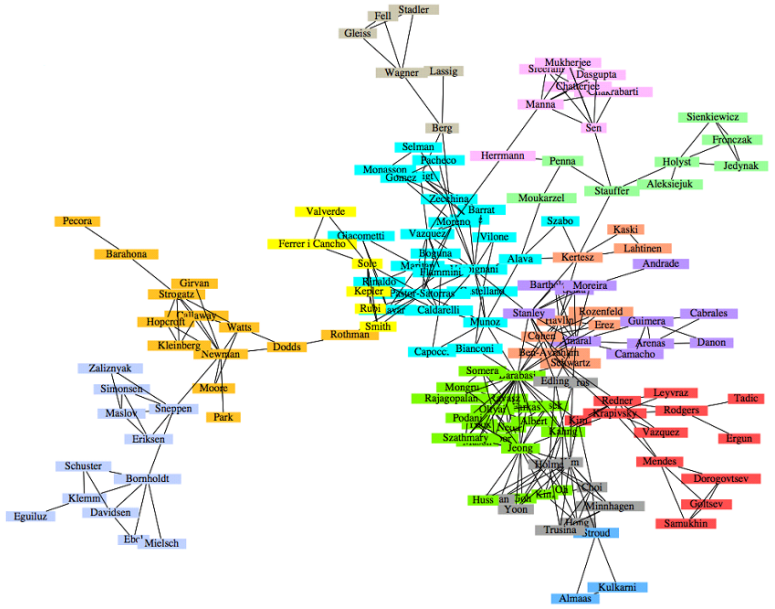
- Is the global friendship network connected?
  - Not necessary, some node maybe have no friend
  - Large complex networks often have a giant component, a connected component that contains a significant fraction of all the nodes
  - Why only one?
    - If there are two, it requires no single link between nodes in the two components, which is unlikely.



Community

- Social network tends to group into clusters due to different interests
- Communities
  - Sets of nodes with lots of connections inside and few to outside

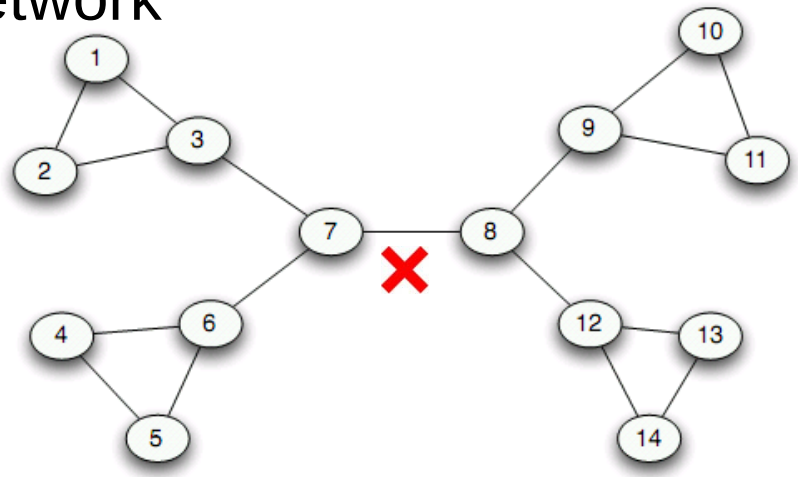




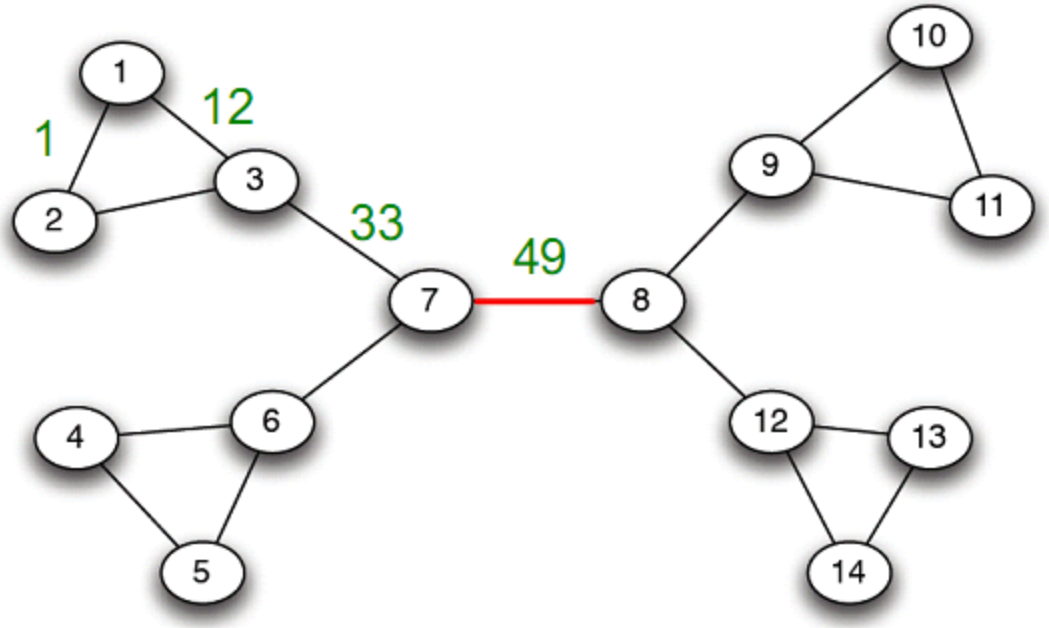
computer NETWORKS

Community Detection

- How to divide a network into communities?
  - By observation?
  - Automatically?
- A possible idea
  - Finding the most important edges to divide the network
  - Imagine traffic flows in the network



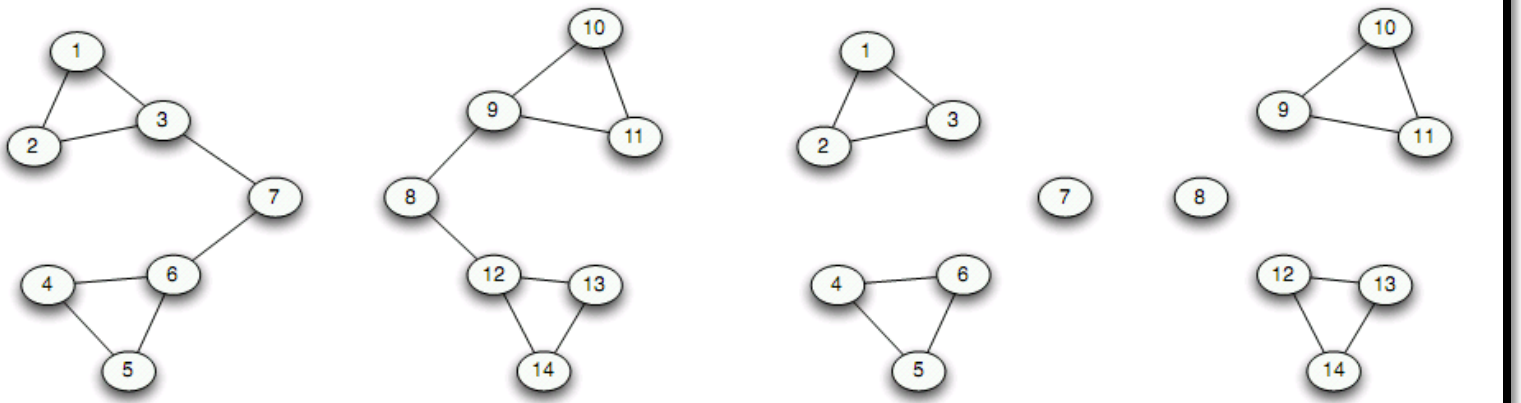
- Betweenness
  - The number of shortest paths passing through the edge
-



Girvan-Newman Algorithm

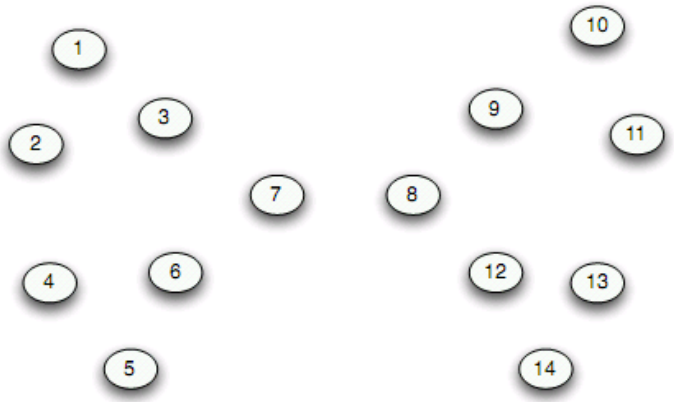
- Process
  - 1. Calculate betweenness of each edges
  - 2. remove edges with highest betweenness
  - 3. repeat 1,2 until the number of communities reach to a threshold or no edges are left
- Works for undirected unweighted graph
- Gives a hierarchical decomposition of the network

Example

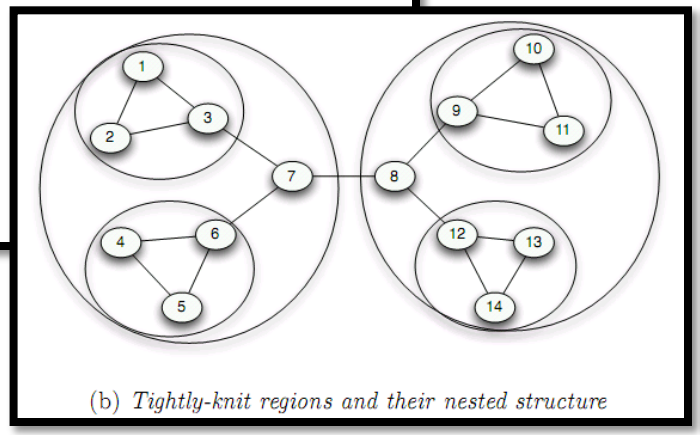


(a) *Step 1*

(b) *Step 2*



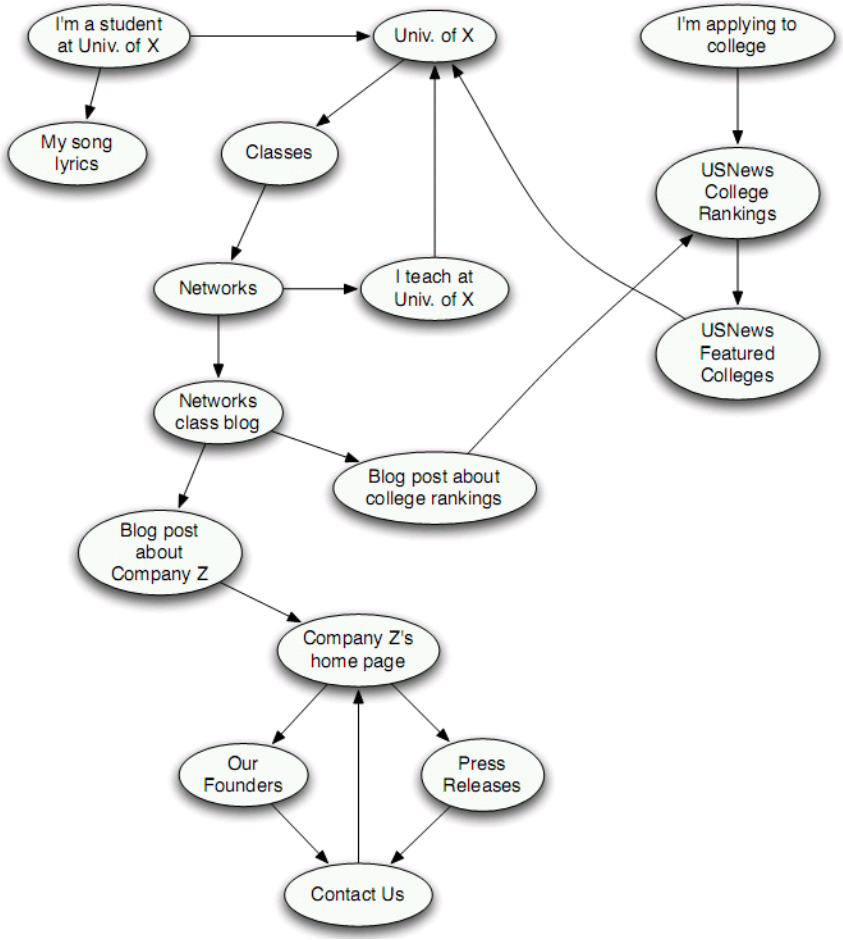
(c) *Step 3*



(b) *Tightly-knit regions and their nested structure*

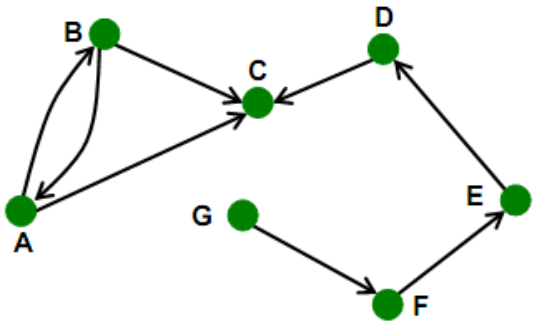
The Structure of the Web

- Web as a directed graph
  - Nodes: pages
  - Edges: hyperlinks (directed)



- Question: What does Web look like at a global level?
  - Giant component?
  - Small Communities?

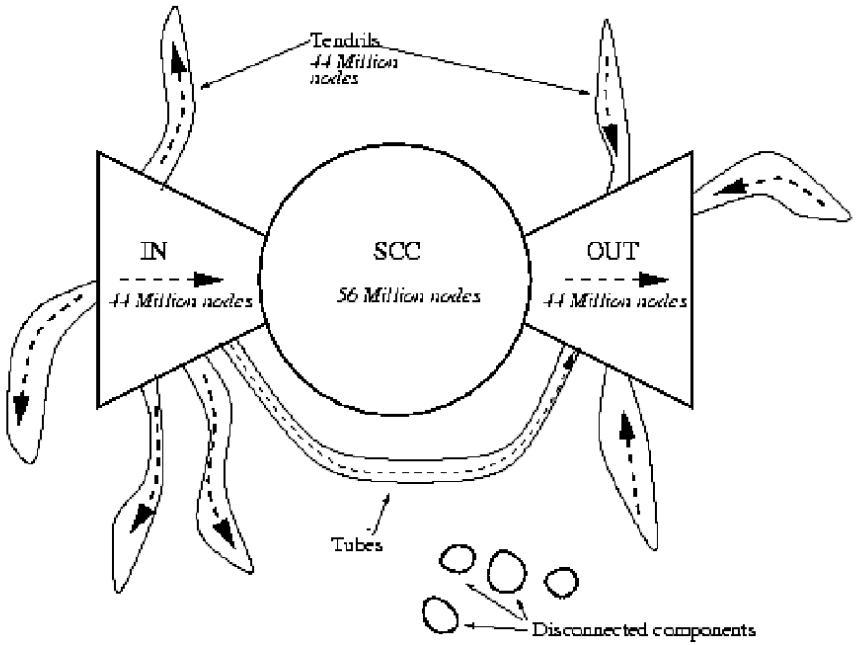
Directed Graph



- Path: directional
- Strong connectivity
  - A directed graph is strongly connected if there is a path from every node to every other node
- Strongly connected component (SCC)
  - A subset of the nodes that (i) (i) every node in the subset has a path to every other; and (ii) the subset is not part of some larger set with the property that every node can reach every other.

The Bow-Tie Structure of the Web

- 250 million pages, 1.5 billion links (1999)
- A giant SCC (56 million nodes)
- IN set (44 million nodes)
  - Nodes that can reach the giant SCC but cannot be reached from it
- OUT set (44 million nodes)
  - Nodes that can be reached from the giant SCC but cannot reach it
- Tendrils (44 million nodes)
  - The nodes reachable from IN that cannot reach the giant SCC
  - The nodes that can reach OUT but cannot be reached from the giant SCC.
- Tubes
  - The nodes reachable from IN to OUT
- Disconnected



Presentation Topics

- 1. The little engine(s) that could: scaling online social networks. SIGCOMM 2010.
  - http://ccr.sigcomm.org/online/?q=node/642
- 2. An analysis of social network-based Sybil defenses. SIGCOMM 2010.
  - http://ccr.sigcomm.org/online/?q=node/643
- 3. Link Analysis and Web Search, Chapter 14 of the book
  - David Easley and Jon Kleinberg, Networks, Crowds, and Markets: Reasoning About a Highly Connected World, Cambridge University Press, 2010.