Dynamic and Social Network Analysis

Lecture-1
Miray Kas
Bilkent University
Computer Engineering Department

Welcome to Dynamic and Social Network Analysis!

A Little About Me

- Bilkent Alumni (BS and MS from CS Department)
  - During my masters, my advisors were Prof. Ibrahim Körpeoğlu and Prof. Ezhan Karasan
- PhD from Carnegie Mellon University
  - My research area was social network analysis
- Working at Google for ~8 years
- I joined a startup for about a year in between my Google experience
- I have been living in the USA for the past 13 years.

Tell me about yourself!

- Brief personal intro, which year/department are you in?

- What are your career plans (academia? industry? abroad?)

- What is your area of interest?

Class Logistics

Goals - What is this class about?

Gain a basic understanding of dynamic and social network analysis

- Learn the terminology and fundamental techniques of network analysis
- How and where it can be used
- Collection, representation, and conversion of social network data
- Learn how to interpret the results of an analysis
- Hands on practice

NOT about pure programming!

Topics

- Intro, Logistics & Basics
- Network Analysis Fundamentals
- Metrics & Centrality
- Structure: Groups and community
- Network Topology
- Dynamic Networks
- Network Evolution
- Applications
  - Authorship/Science/Citation Networks
  - Geospatial Networks
  - Social Media Networks and Analysis
  - Health and Disease Spread
  - Malicious Networks

Approach and Course Structure

- First half (until midterm):   Theory and Methods
- Second half (after midterm):   Applications
- 3 lectures a week
  - Completely remote, online.
  - Slide decks and lecture recordings will be shared.
  - Created a joint course for 429/529 on Moodle
- Hands on
  - Homeworks + midterm
  - Tools usage
  - Final Project (Write up + presentation)
  - No final exam!

Grading Criteria

Homeworks (25%):

- 5 homeworks equally weighted
- Only 1 or 2 of them will involve programming
- The rest will involve hands on problem sets, tool usage & analysis/interpretation

Midterm (25%):

- We will use one of our class sessions for ease of scheduling.
- Cameras turned on, will be proctored online. No other person in the room.
- Attendance in the exam without camera turned on will not be accepted

Final Project (50%)

More on the final project

Final Project (50%):

- Can be done individually or in pairs.
- 4 milestones to ensure progress. Each step is graded.
- **Project Milestones**
  - **P1: Project Short Proposal** (5%): A brief discussion of your proposed project: dataset, methods, and expected challenges. No less than 250 words, no more than 500 words.
  - **P2: Project Draft Intro, Background, and Methodology Sections (**15%**)**
  - **P3: Project Presentation:** 10-15 minute conference talk discussing your project's methods and key findings (10%)
  - **P4: Final Project Submission** (20%) Due Last Day of Class; Publication style write up of your project in the format of an ACM paper (6 pages max, 4 pages min including references; this page limit is **strict**). You should share your data and analysis files for evaluation as well

More on the final project

- **Choose and propose a topic**
  - You are welcome and encouraged to choose your topic before Project Short Proposal is due
- **Requirements**
  - Must be strongly related to the subject of the course (lectures + reading material)
  - Must strongly tie to the literature (No plagiarism, at least 10 references cited in final write up)
  - Must involve real or simulated network data
  - Must use network / node / link analysis techniques
  - Must be in the format of an ACM paper (6 pages max, 4 pages min including references; this page limit is **strict**).
  - Final submission should include the writeup, dataset, analysis files, and the presentation for the talk together.

Contact

- Email: miraykas@cs.bilkent.edu.tr, miraykas@gmail.com

- There will be 3 lectures in a regular week.

- Thursday 9:30 class is the spare hour.

  - I will use it as the office hour.

Introduction to Social Network Analysis

The fate of Saddam Hussain & Network Analysis
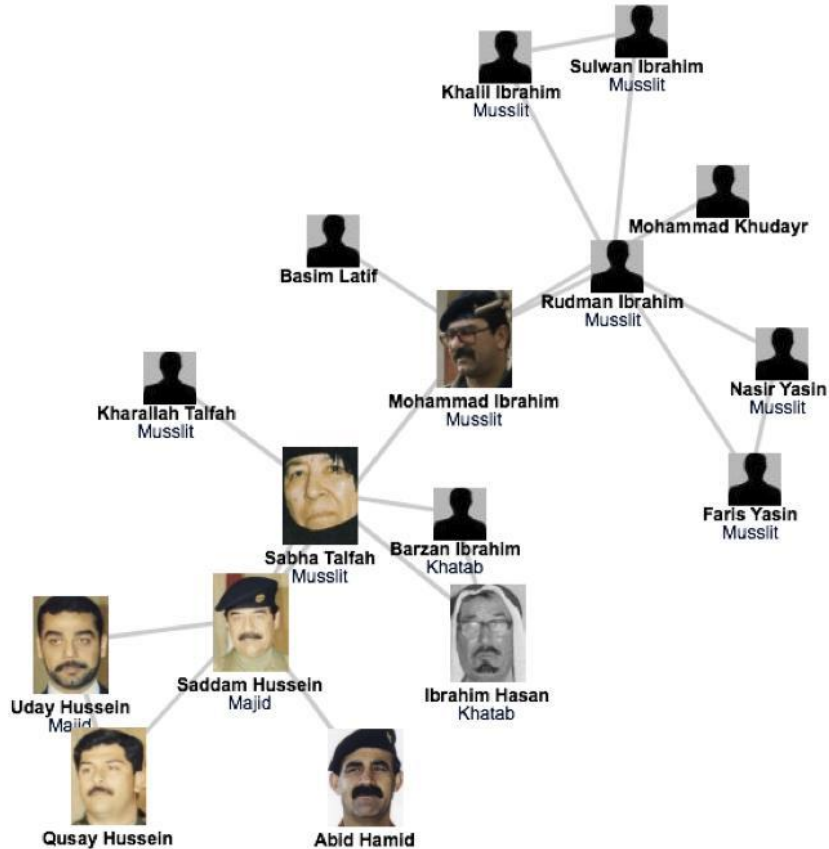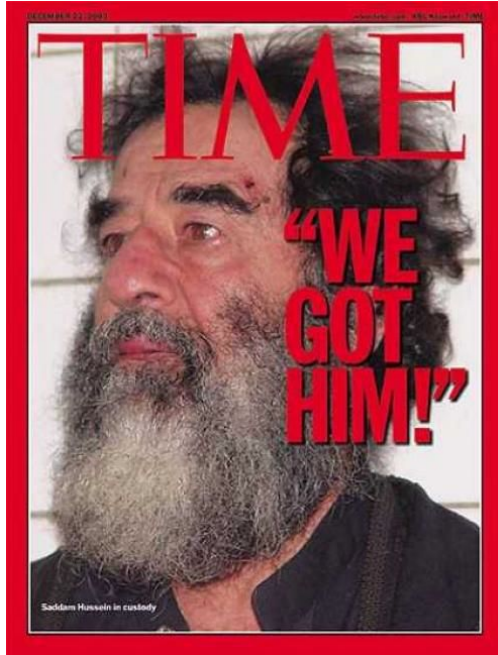
The fate of Saddam Hussain & Network Analysis

- Networks show strong predictive power
- His capture was not based on fresh data!
  - Hierarchical government network did not show his whereabouts
  - Social Network constructed from family photo album revealed his locations
  - Construction of social networks (e.g. mapping information) is very complicated!

- Your choice of network and how you analyze is super important!

Rise of Social Media: Billions of Users!

- Social Networks (Facebook)
- Discussion Forums (Reddit)
- Media Sharing Networks (Instagram)
- Blogging & Publishing Networks (Medium)
- Consumer Review Networks (Yelp)

Social Network Analysis of Your Friends in Facebook

JÜRGEN'S
SOCIAL
NETWORK

Based on Facebook Data
Spring 2013

Friends from
ETHIOPIA

Friends from
CHURCH
GRAZ

Friends from
CHURCH
STUTTGART

Friends from
UNIVERSITY
GRAZ

Flatmates and
neighbors from
GRAZ

Friends from
CANADA

Friends who don't know
anybody else here

Friends and Family from
MY HOMETOWN

Friends from
HIGHSCHOOL

Friends from
UNIVERSITY
STUTTGART

https://www.jroehm.com/2014/10/29/social-network-vizualiation/

Fast-Growing Scientific Area

Social Network Analysis (SNA) is studied since 1960s.

- Exponential growth in early 2000s

Government & Defense

- Saddam's case
- Trade tracking between countries

Social media and social network applications

Academia/Science

- Linguistics, health and disease spread
- Human genome and brain connectivity

Business Practitioners

- New organizational forms
- Management consultants
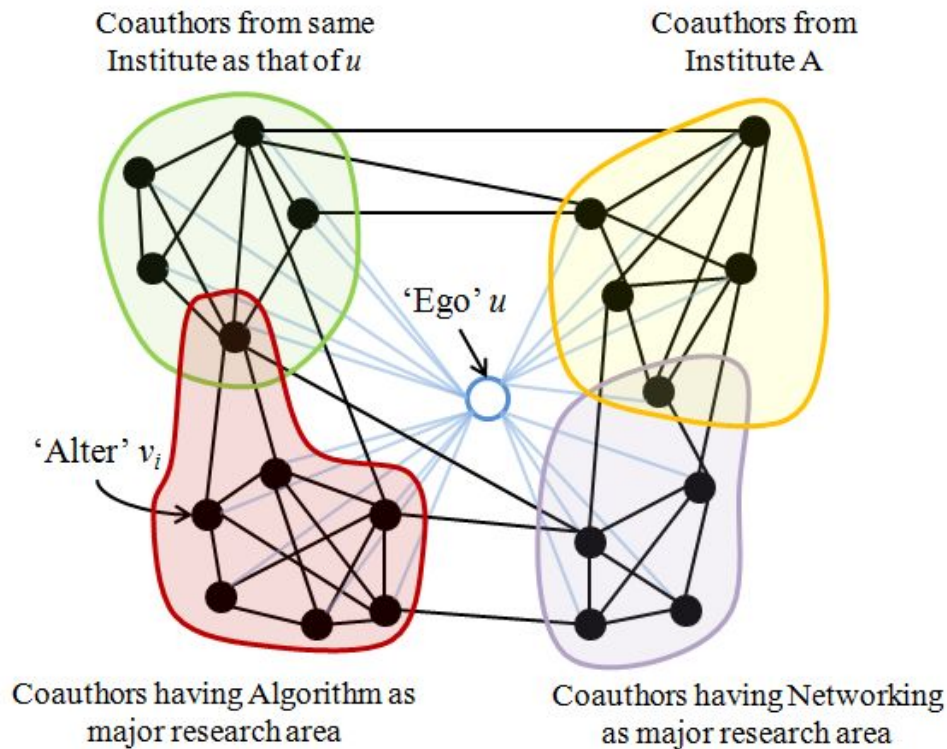
In 1960s, sociologists were studying a group of monks.

The techniques are enhanced to analyze networks with billions of nodes!

Network Perspective
&
Fundamentals of Dynamic and Social Network Analysis

The Network Perspective

- Holistic view & System level thinking
- Look at the elements, but also how they relate to one another
- Can be analyzed at multiple layers
  - No reductions, holistic.
  - Reduced for zooming/focus reasons

Coauthors from same Institute as that of u

Coauthors from Institute A

'Ego' u

'Alter' v_i

Coauthors having Algorithm as major research area

Coauthors having Networking as major research area

Fundamental Difference btw.
Network and Non-Network Perspective

Inclusion of information

about

relationships!

Mainstream Data vs Social Networks

Variables (Attributes)

Person	Age	Work Org	Alma Mater
p1	32	MS Research	MIT
p2	25	Google	CMU
p3	47	MS Research	CMU
p4	47	MS Research	Georgia Tech

Cases (Individuals)

Social Network

A social network is a snapshot!

- Description of relationships among a group of actors at a particular point in time
- Actors could be individuals, organizations, and with new applications, they can even be genes or scientific topics!
- The social networks key component is relationships - show how the actors are connected to one another

Fundamental Concepts - Elements of a Social Network

Actor

Actors are discrete individual, corporate, or collective social units (among others; also: node, vertex)

Individual: A Facebook friend, a romantic partner

Corporate: Companies, government agencies, universities

Collective social units: Groups that can be represented as a node on a graph

Definitions based on Wasserman & Faust, 1994)

Fundamental Concepts - Elements of a Social Network

Relational Tie

A relational tie establishes a linkage between a pair of actors

- Evaluation of one person by another (for example expressed friendship, liking, or respect)
- Transfers of material resources (for example business transactions, lending or borrowing things)
- Association or affiliation (for example jointly attending a social event, or belonging to the same social club)
- Behavioral interaction (talking together, sending messages)
- Movement between places or statuses (migration, social or physical mobility)
- Physical connection (a road, river, or bridge connecting two points)
- Formal relations (for example authority)
- Biological relationship (kinship or descent)

Definitions based on Wasserman & Faust, 1994

Fundamental Concepts - Elements of a Social Network

Dyad

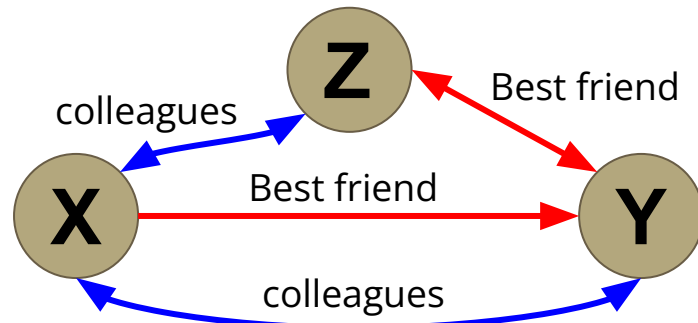- Dyad = actor_1 + actor_2 + possible relationships between them

Dyadic analysis focuses on whether
- the relations are bidirectional (i.e. reciprocated)
- certain types of relationships occur together

Triad

- Triad = actor_1 + actor_2 + actor_3 + possible relationships between them

Triadic analysis is crucial for understanding transitivity & balance

Definitions based on Wasserman & Faust, 1994)

Fundamental Concepts - Elements of a Social Network

Subgroup

- A subset of actors and all ties among them
- Dyads and Triads are special cases of subgroups

Group

- Collection of all actors on which ties are to be measured
  - A finite set or sets of actors
  - Must be able to argue theoretical, empirical, or conceptual criteria that brings the group of actors together

Definitions based on Wasserman & Faust, 1994)

Social Network

A social network is a snapshot!

- Description of relationships among a group of actors at a particular point in time
- Actors could be individuals, organizations, and with new applications, they can even be genes or scientific topics!
- The social networks key component is relationships - show how the actors are connected to one another

Real World Networks???

- Can pretty much model all real world interactions as networks

- *Main Groups:*
  - Social Networks (Longest history of study)
  - Information Networks
  - Technological Networks
  - Biological Networks

Set-I: Social Networks

- ***Vertices:*** Individuals or groups of people
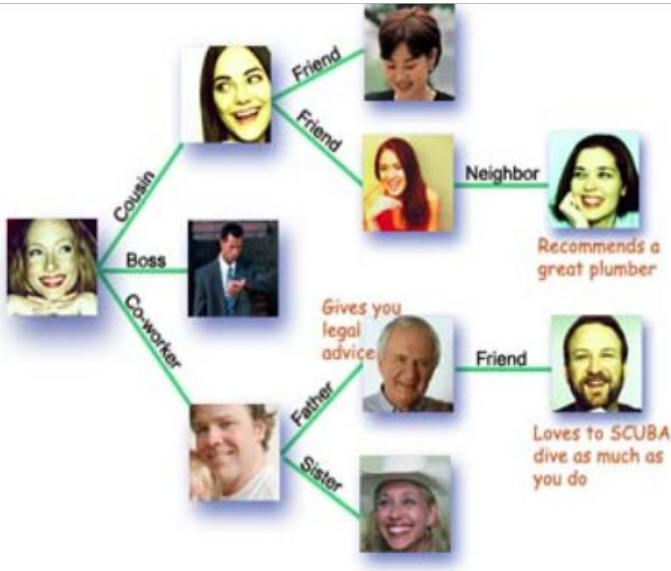- ***Links:*** Interactions or contacts

EXAMPLES:

Friendship
Business Relations
Intermarriages btw. Families
Collaboration btw. Film actors
*Co-authorship among academics*

Set-I: Social Networks

- Where to find reliable data about social networks?
    - Surveys are biased

- Potential Sources:
    - Communication records
        - Who calls who?
    - Facebook
        - Who is a friend of who?
    - Google Scholar
        - Who writes papers with whom?
    - IMDB
        - Who co-stars with whom?

Set-II: Information Networks

- *Vertices:* Information
- *Links:* Containment, reference, etc.
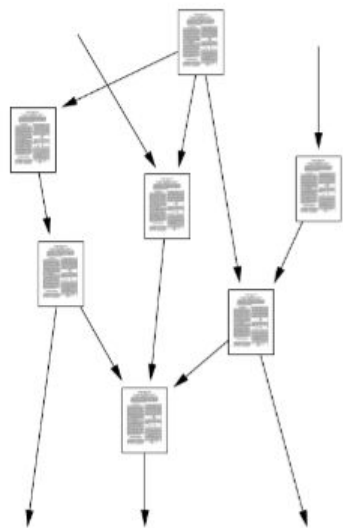
citation network
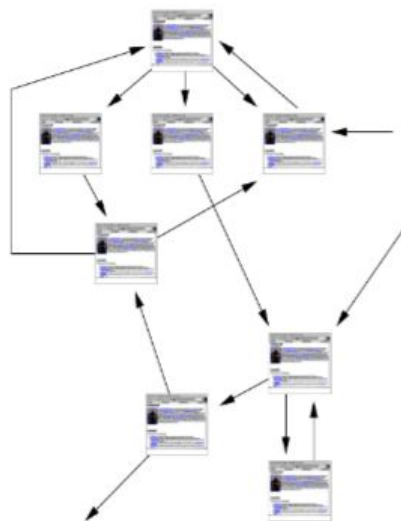
World-Wide Web

EXAMPLES:

Citation Networks

World Wide Web

Words in thesaurus

Color code:

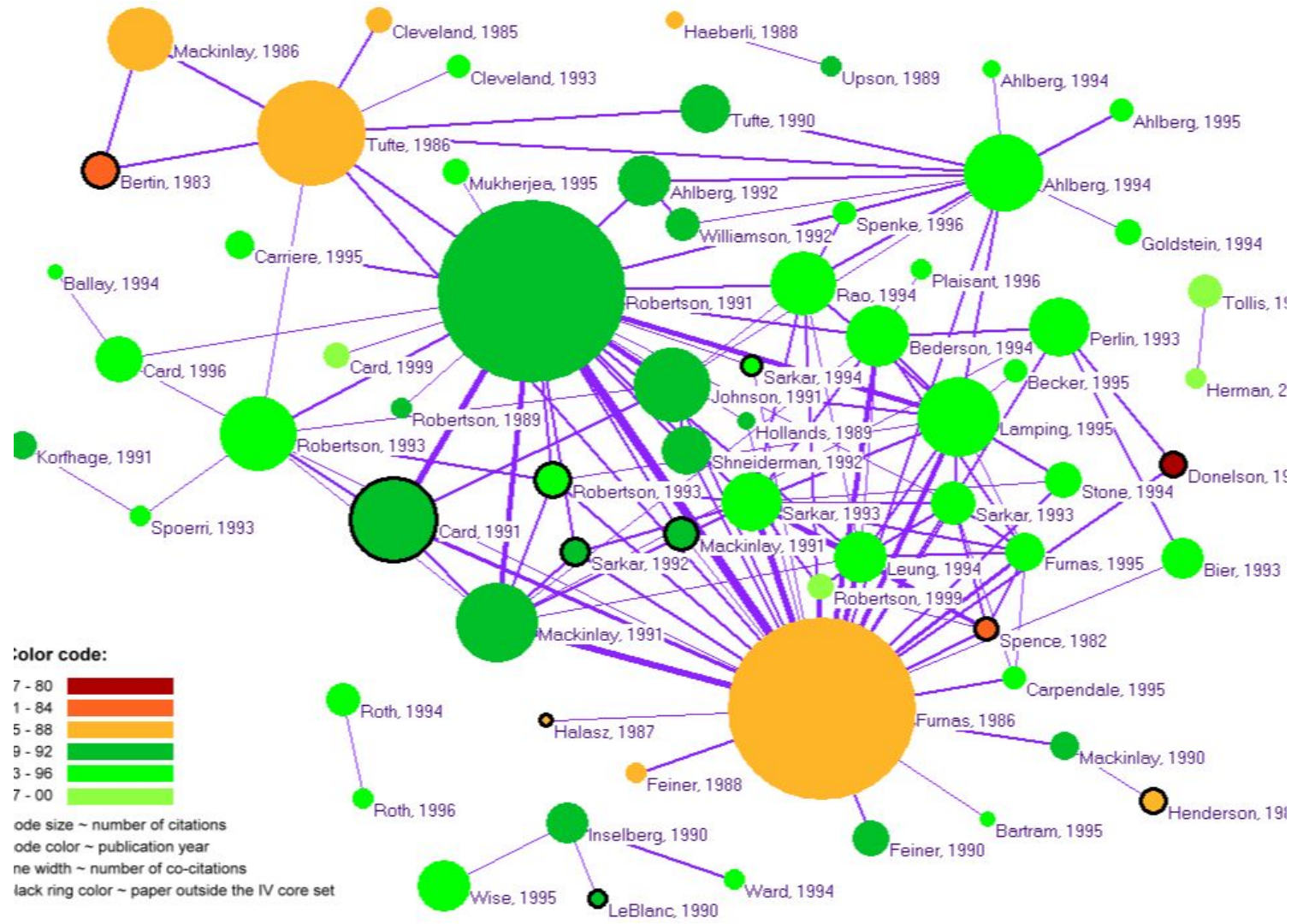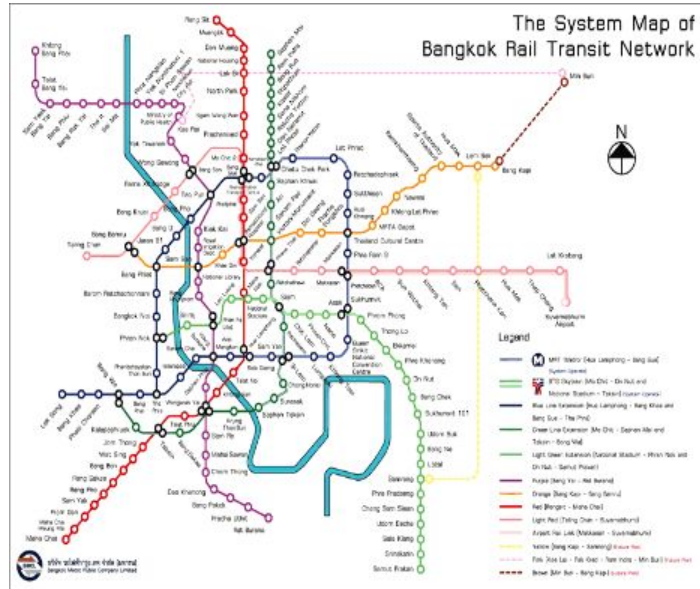|	
▮ 77 - 80	(dark red)
▮ 81 - 84	(orange)
▮ 85 - 88	(amber)
▮ 89 - 92	(dark green)
▮ 93 - 96	(green)
▮ 97 - 00	(light green)

node size ~ number of citations
node color ~ publication year
line width ~ number of co-citations
black ring color ~ paper outside the IV core set

Mackinlay, 1986
Cleveland, 1985
Haeberli, 1988
Upson, 1989
Ahlberg, 1994
Cleveland, 1993
Ahlberg, 1995
Tufte, 1990
Bertin, 1983
Tufte, 1986
Ahlberg, 1994
Mukherjea, 1995
Ahlberg, 1992
Spenke, 1996
Goldstein, 1994
Williamson, 1992
Carriere, 1995
Plaisant, 1996
Ballay, 1994
Robertson, 1991
Rao, 1994
Tollis, 19
Card, 1996
Card, 1999
Bederson, 1994
Perlin, 1993
Herman, 2
Sarkar, 1994
Johnson, 1991
Becker, 1995
Robertson, 1989
Hollands, 1989
Lamping, 1995
Korfhage, 1991
Robertson, 1993
Shneiderman, 1992
Donelson, 19
Robertson, 1993
Stone, 1994
Spoerri, 1993
Sarkar, 1993
Sarkar, 1993
Card, 1991
Mackinlay, 1991
Leung, 1994
Furnas, 1995
Bier, 1993
Sarkar, 1992
Robertson, 1999
Mackinlay, 1991
Spence, 1982
Carpendale, 1995
Roth, 1994
Furnas, 1986
Halasz, 1987
Mackinlay, 1990
Roth, 1996
Feiner, 1988
Henderson, 19
Inselberg, 1990
Bartram, 1995
Feiner, 1990
Wise, 1995
LeBlanc, 1990
Ward, 1994

Set-III: Technological Networks

- Man made networks
- Distribution of commodity or resources

The System Map of Bangkok Rail Transit Network

EXAMPLES:

Railways

Electric Power grid

Telephone network

Internet

230,000 volts
345,000 volts
500,000 volts
765,000 volts
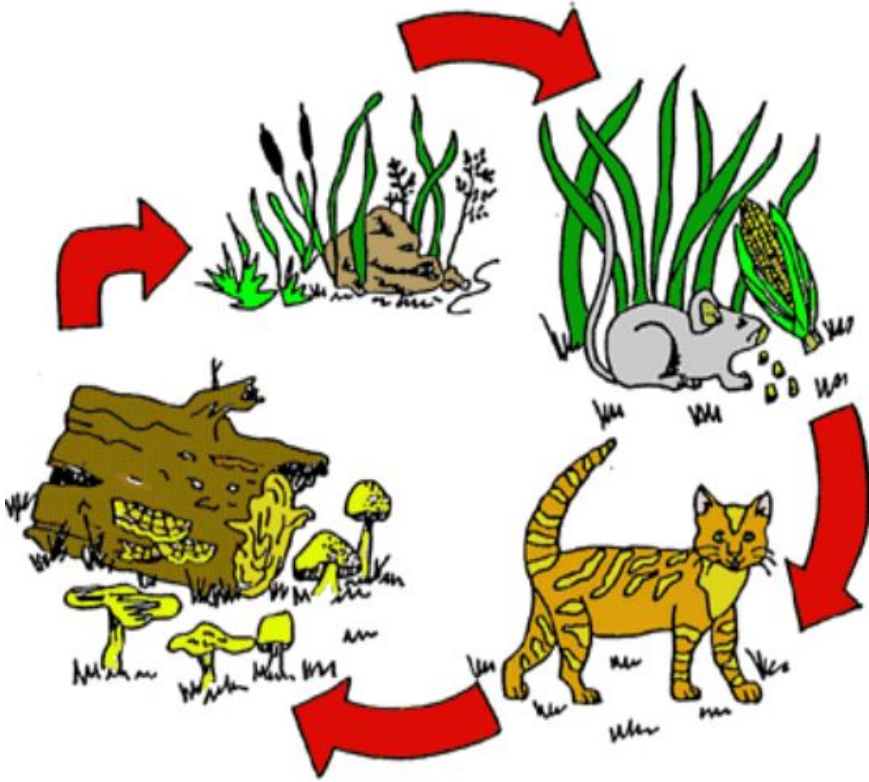high-voltage direct current

Set-IV: Biological Networks

EXAMPLES:

Blood vessels

Genetic Regulatory

Metabolic pathways

Neural networks

Food Web (Who eats who?)

Networks are Everywhere!

- **Social networks** : online social networks, edges represent interactions between people
- **Networks with ground-truth communities** : ground-truth network communities in social and information networks
- **Communication networks** : email communication networks with edges representing communication
- **Citation networks** : nodes represent papers, edges represent citations
- **Collaboration networks** : nodes represent scientists, edges represent collaborations (co-authoring a paper)
- **Web graphs** : nodes represent webpages and edges are hyperlinks
- **Amazon networks** : nodes represent products and edges link commonly co-purchased products
- **Internet networks** : nodes represent computers and edges communication
- **Road networks** : nodes represent intersections and edges roads connecting the intersections
- **Autonomous systems** : graphs of the internet
- **Signed networks** : networks with positive and negative edges (friend/foe, trust/distrust)
- **Location-based online social networks** : social networks with geographic check-ins
- **Wikipedia networks, articles, and metadata** : talk, editing, voting, and article data from Wikipedia
- **Temporal networks** : networks where edges have timestamps
- **Twitter and Memetracker** : memetracker phrases, links and 467 million Tweets
- **Online communities** : data from online communities such as Reddit and Flickr
- **Online reviews** : data from online review systems such as BeerAdvocate and Amazon
- **User actions** : actions of users on social platforms.
- **Face-to-face communication networks** : networks of face-to-face (non-online) interactions
- **Graph classification datasets** : disjoint graphs from different classes

How do we understand what is going on with networks

- How do we understand what is going on?
- How do we analyze their properties?

- Shift of focus:
  - We want to analyze large-scale graphs
  - We will need sophisticated software tools (Next topic!)

- Three aims of research:
  - Find statistical properties
  - Create network models
  - Predict their behavior

Introduction to Social Network Analysis Software

Social Network Analysis Software

- Many tools are available online, mostly free for students (Links available in the [syllabus](#))
  - You can choose/suggest another tool that is not in the syllabus.
- You are free to choose whichever tool works with your system.
- Make sure to try them out and choose one or two you like, you will need it starting Homework-1 all the way to your final project!

Social Network Analysis Software

- Some are totally visual, some are script based.
- Some work with Windows only, some need programming environment installed etc.
- Most tools have sample datasets available
  - They usually have their own file format.
  - Almost all of them accept csv format and convert from that.
- Some of them are open source and free, some are paid
  - Paid software usually gives long enough trial for students to complete a semester.

Social Network Analysis Software

- Wikipedia has a comprehensive list, but even that is not complete.
- https://en.wikipedia.org/wiki/Social_network_analysis_software

ORA

- 180-day trial version (enough for a semester)
- From Carnegie Mellon University, CASOS Group
- Runs on Windows
- Has very detailed metrics and reports
  - Stronger than Gephi
- Many datasets available: CASOS Tools: Network Analysis Data | CASOS

Gephi

- Open source and free
- Runs on Windows, Mac OS X and Linux (You can download from https://gephi.org/)
- Be careful with the Undo button!
- Many plugins available ([Gephi Plugins](https://gephi.org/plugins))
  - They have well published APIs, you can write your own plugin ([Gephi developers](https://gephi.org))
- So many ways of getting started with Gephi
  - http://literaturegeek.com/2013/09/09/dataintogephi
  - [Gephi Tutorial Quick Start](https://gephi.org)
  - https://www.youtube.com/watch?v=FpOIbhOmGUs&ab_channel=%7BYangSong%7D

SocNetV

- Runs on Windows, Mac OS X and Linux (You can download from [Downloads - Social Network Analysis and Visualization Software](#))
- Easy to create some of the simulated network models.

Specialization: SocioViz

- Very basic tool that allows queries on Twitter/Facebook

- Free access is limited

  - More sophisticated queries can be run in paid form.

- SocioViz

Specialization: VOSviewer

- [Download VOSviewer](#)

- Focused on scientific networks

- Works for Windows, Mac, Linux

Next Lecture:

Tool Demos
Representing Social Network Data
(Graphs and Matrices)