Complex Network Theory

Lecture 1:

Networks: what they are and what they are good for

Instructor: S. Mehdi Vahidipour
(Vahidipour@kashanu.ac.ir)

Spring 2018

Thanks A. Rezvanian
L. Adamic , A. Barabasi, J. Leskovec

Feb., 2018

Outline

- Overview of class topics
  - Course overview
  - Syllabus
  - References
  - Evaluation

- First topic
  - Types of networks
  - Network models
  - Implications of network structure

- Next class:
  - Basic definitions & metrics

Course overview

- **Ref:**
  - Newman, Mark. ***Networks: an introduction***. Oxford University Press, 2010.
- **Goals**
  - Introduce the basic mathematical tools to understand the fundamentals of complex networks.
  - Provide the skills that are needed to perform basic analyses of such networks.
- **Means**
  - Study fundamental concepts from graph theory and random networks.
  - Lots of exercises in proving properties of various well-known networks
  - Practice the use of network analysis tools

Topics

- Complex network theory
  - Complex network characteristics
  - Complex network models (Erdős–Rényi, watts-strogatz, Barabási–Albert)
  - Network metrics, centralities, measures & algorithms
  - Structural and behavioral analysis (static & dynamic)
  - Subgraphs, cliques and Motifs analysis & algorithms
  - Social network analysis
  - Spectral and random walks techniques on networks
  - Community structure discovery and analysis in complex networks
  - Diffusion models on networks
  - Cascading failures in real networks
  - Network visualizations
  - Supplementary materials
    - Research methods, report and presentation
    - Advanced methods for designing experiments
    - MATLAB and R programming

Evaluation

- Homeworks and Quizzes (no notice)
  - About 6 prearranged homework.
  - **Important!** no large delay is accepted.
  - 20% of the final grade (tentative)
- Research presentation, discussion, report and simulation
  - **Map and Reduce**
  - **Community Detection scope**
    - ISI journal papers published after 2015 having IF>1.5 or SNIP>1.5 or Q1
  - 40% of the final grade (tentative)
- Final Exam
  - Conceptual questions about topics
  - 40% of the final grade (tentative)
- Evaluation on your works (not any other)!
- 30% penalty per day for late submission of homework
- Email policy (subject: DCN_ID), Vahidipour@kashanu.ac.ir)

Policies & Assumptions

- Familiar with fundamental computer sciences, mathematics and computer problems
- Love (like) mathematic and algorithms(at least don't hate)
- Interested in graph problems and network applications
- Everything is abstract model of real problem
- We don't discuss about details of problem, protocols and technologies.
- You have no problem with class time: Monday and Tuesday, 8:00-9:40.
- You as an engineer should behave like an engineer !
  - Accuracy, Time, Ethics

Complex networks analysis perquisites

- Algorithms
- Graphs theory
- Machine learning
- Probability
- Statistics
- Linear algebra
- Network applications

Networks and complex systems

- Complex systems are around us:
  - Society is a collection of six billion individuals
  - Communication systems link electronic devices
  - Information and knowledge is organized and linked
  - Interactions between thousands of genes regulate life
  - Our thoughts are hidden in the connections between billions of neurons in our brain
- What do these systems have in common?
- How can we represent them?

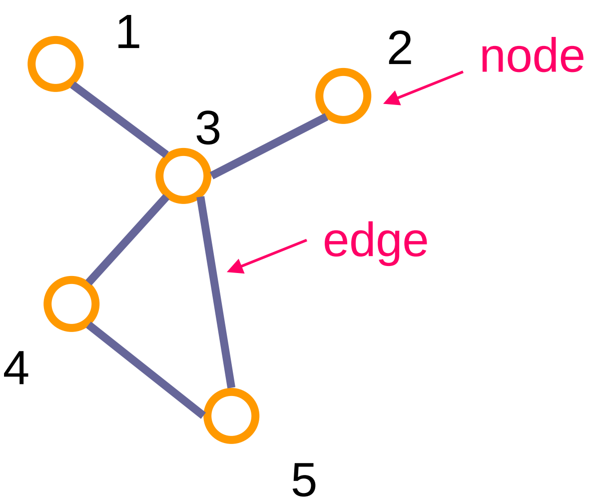
What are networks?

- Networks are collections of points joined by lines.

G(V, E)

"Network" ≡ "Graph"

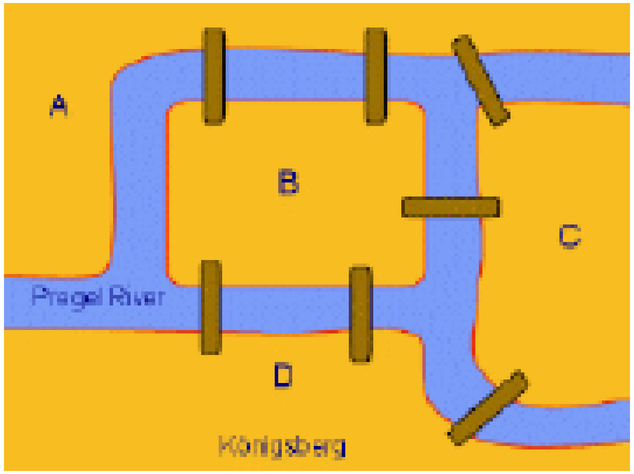
- A network model treats all nodes and links the same
- In a picture of a network, the spatial location of nodes is arbitrary
- Networks are abstractions of connection and relation
- Networks have been used to model a vast array of phenomena
- Graph is a mathematical representation of a network

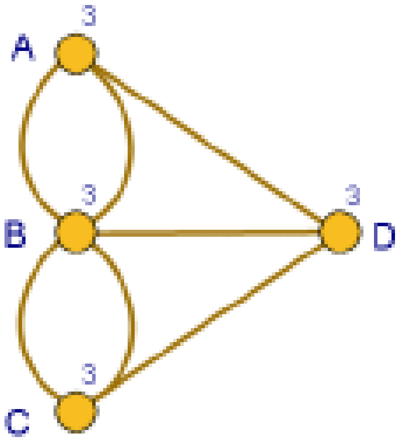


Object	interaction	
points	lines	system
vertices	edges, arcs	math
nodes	links	computer science
sites	bonds	physics
actors	ties, relations	sociology

Graphs

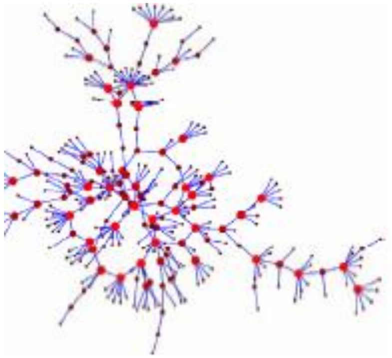
- In mathematics, networks are called graphs, the entities are nodes, and the links are Edges
- Graph theory starts in the 18th century, with Leonhard Euler
  - The problem of Königsberg bridges
  - Since then graphs have been studied extensively





Networks in the past and now

- Past
  - Graphs have been used in the past to model existing networks (e.g., networks of highways, social networks)
    - usually these networks were small
    - network can be studied and visual inspection can reveal a lot of information



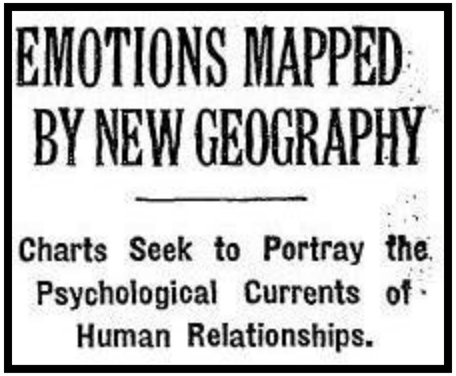
- Now
  - More and larger networks appear
    - Products of technological advancement
      - e.g., Internet, Web, online social networks
    - Result of our ability to collect more, better, and more complex data
      - e.g., gene regulatory networks
  - Networks of thousands, millions, or billions of nodes
    - impossible to visualize
- We will never understand complex system unless we map out and understand the networks behind them.

The History of Network Analysis

- **Graph theory:** 1735, Euler
- **Social Network Research:** 1930s, Moreno
- **Communication networks/internet:** 1960s
- **Ecological networks:** May, 1979.
- **Technological and Information networks**: 1990s
- **Complex social networks:** 2000s

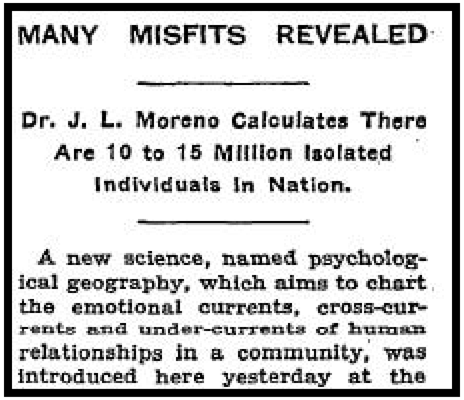
Social network analysis - history

- 1933 Moreno displays first **sociogram** at meeting of the Medical Society of the state of New York
  - article in NYT
  - interests: effect of networks on e.g. disease propagation



EMOTIONS MAPPED BY NEW GEOGRAPHY

Charts Seek to Portray the Psychological Currents of Human Relationships.



MANY MISFITS REVEALED

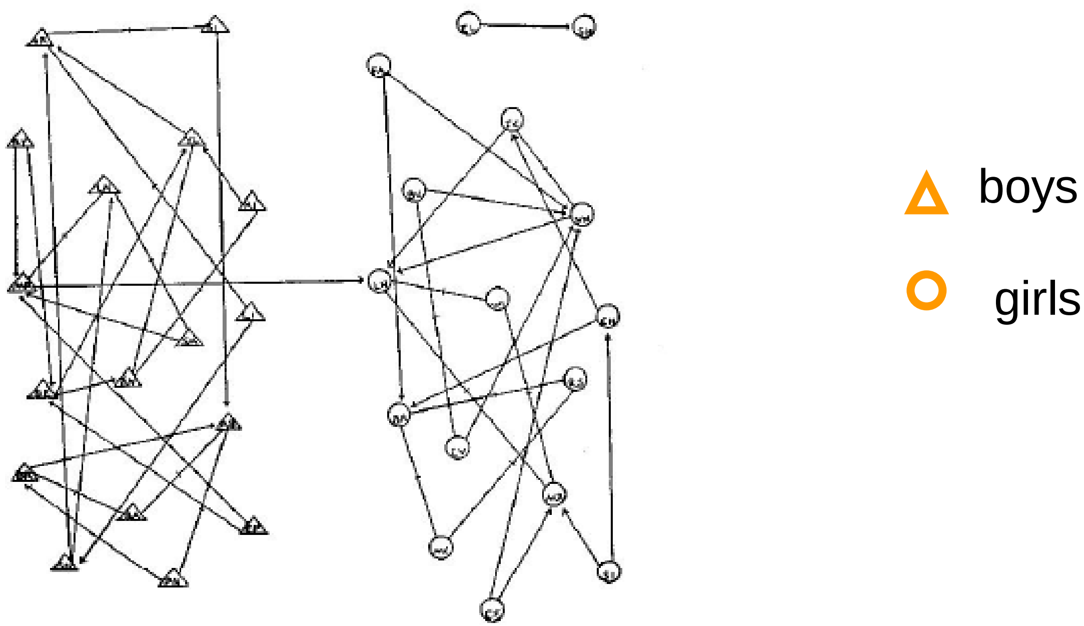
Dr. J. L. Moreno Calculates There Are 10 to 15 Million Isolated Individuals in Nation.

A new science, named psychological geography, which aims to chart the emotional currents, cross-currents and under-currents of human relationships in a community, was introduced here yesterday at the

- Preceded by studies of (pre) school children in the 1920's

Social network analysis - history

- School kids – favorite (and captive) subjects of study
- These days much more difficult because need parental consent to gather social network data



An Attraction Network in a Fourth Grade Class (Moreno, ‘Who shall survive?’, 1934).

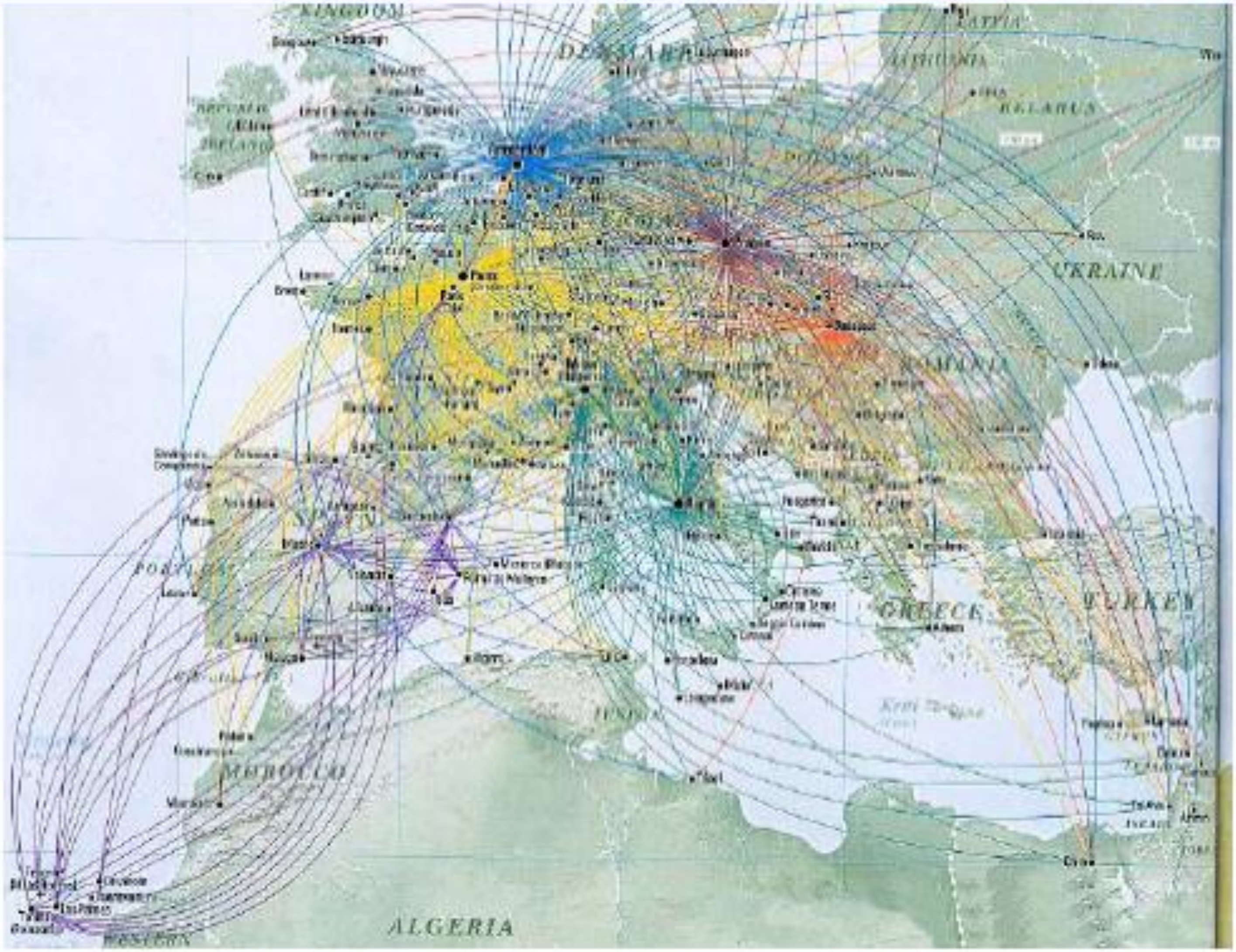
Examples of Networks

- Brain interconnectivity map
- Ecosystem and food chains
- Gene, protein, molecule, and virus networks
- Political relations and Lobbies
- Companies and their financial relations
- Bank transactions
- Power lines, phone lines
- Roads, highways, airlines, railways
- Paper citations
- Emails, Web, Internet
- Mobile, SMS and Phone networks
- Online friendship and social networks

Examples of networks: social

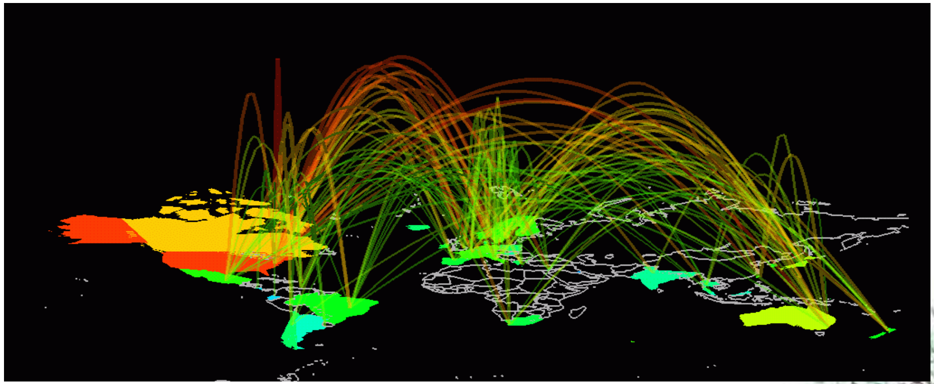
- Evaluation of one person by another (for example expressed friendship, liking, or respect)
- Transfers of material resources (for example business transactions, lending, or borrowing things)
- Behavioral interaction (talking together, sending messages, giving knowledge)
- Formal relations (for example authority; who gives you orders)
- Biological relationship (kinship or descent)
- Sending and receiving social and emotional support
- Association or affiliation (for example jointly attending a social event, or belonging to the same social club)

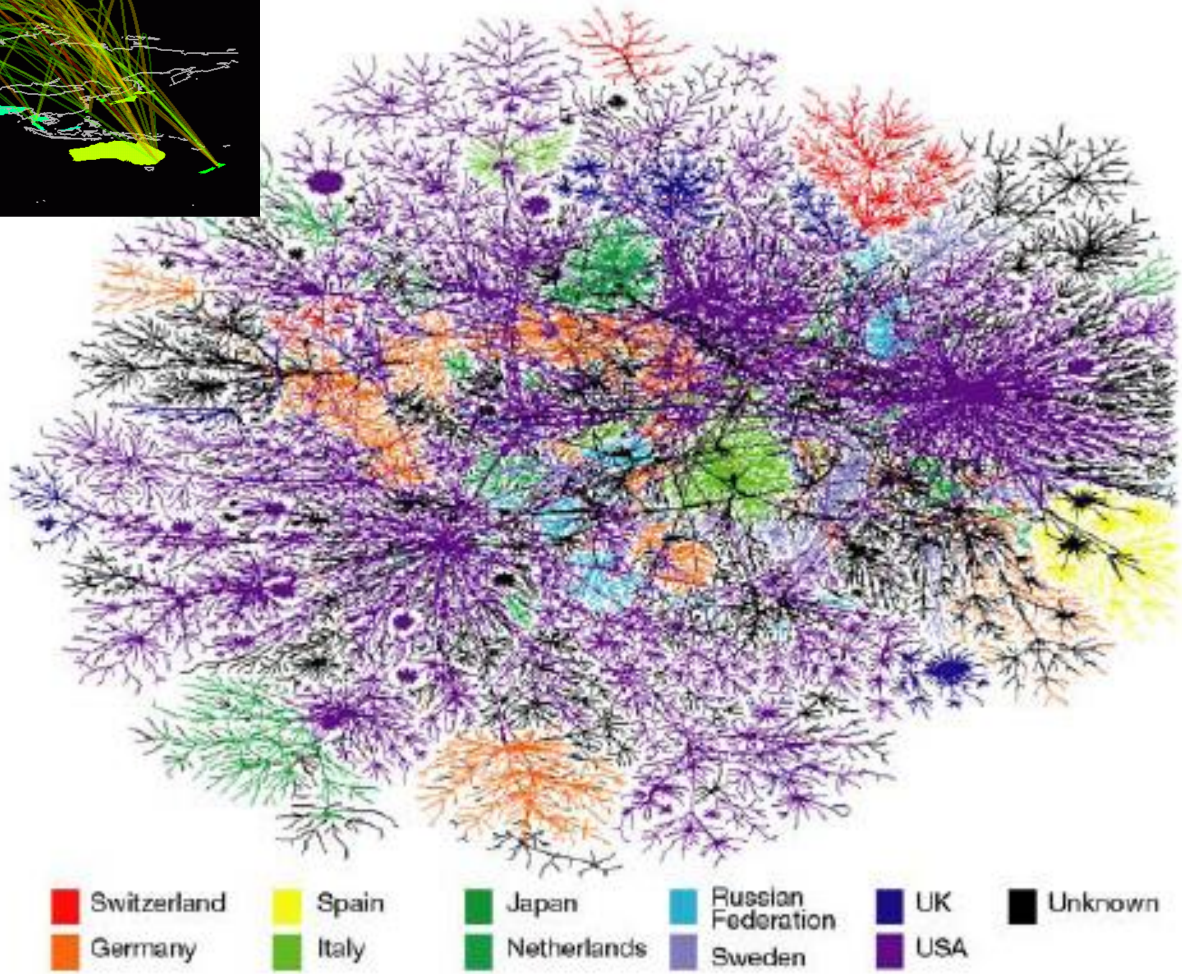
The airline networks



Source: Northwest Airlines WorldTraveler Magazine

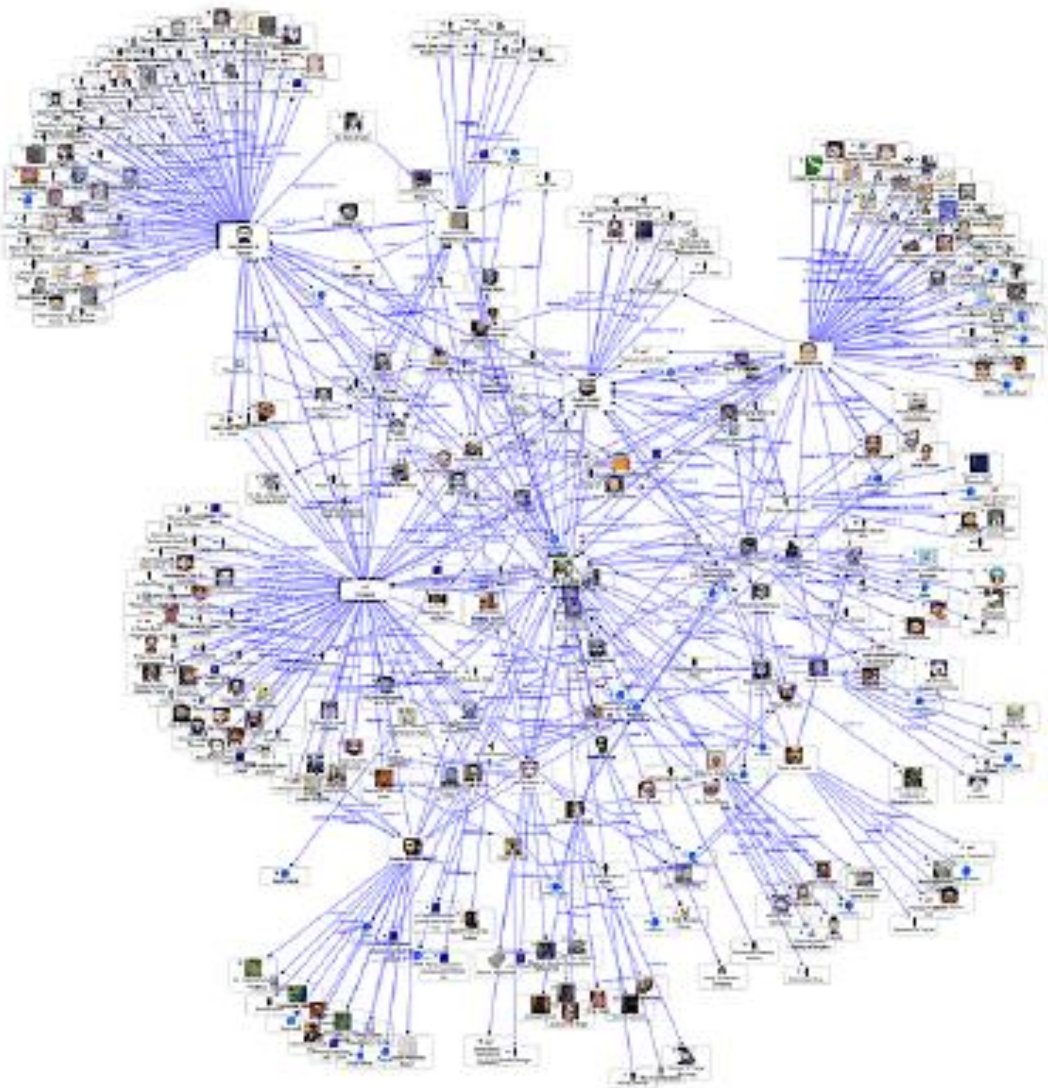
The internet map



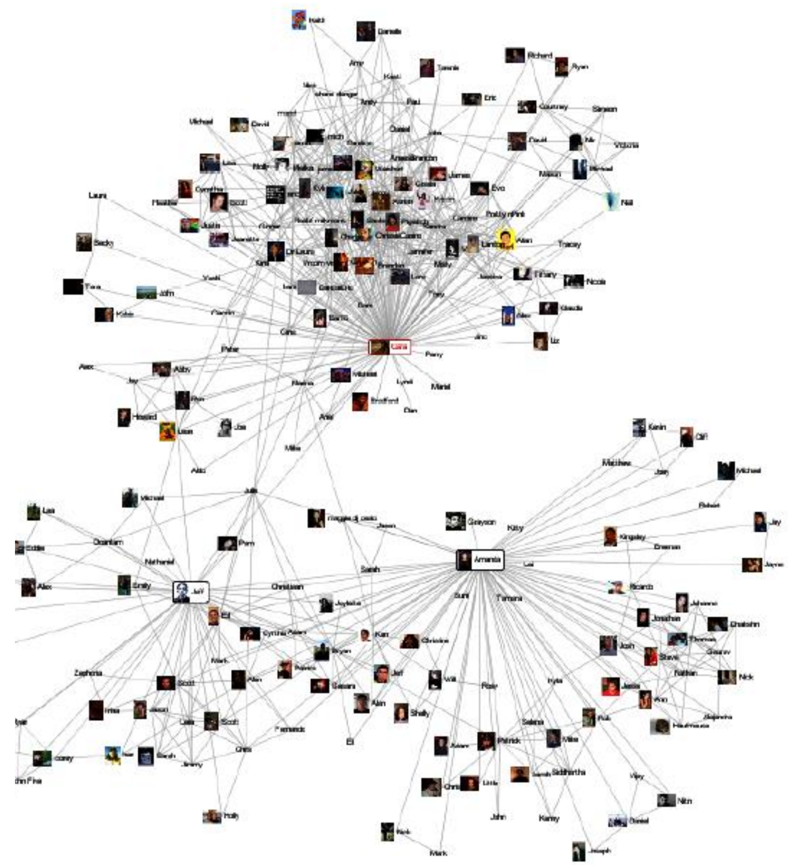


Online social networks

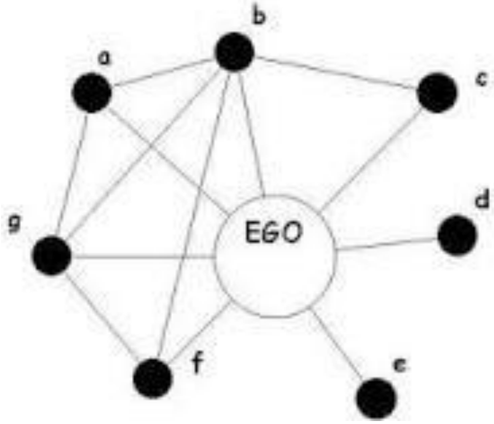
- Friendship in Facebook



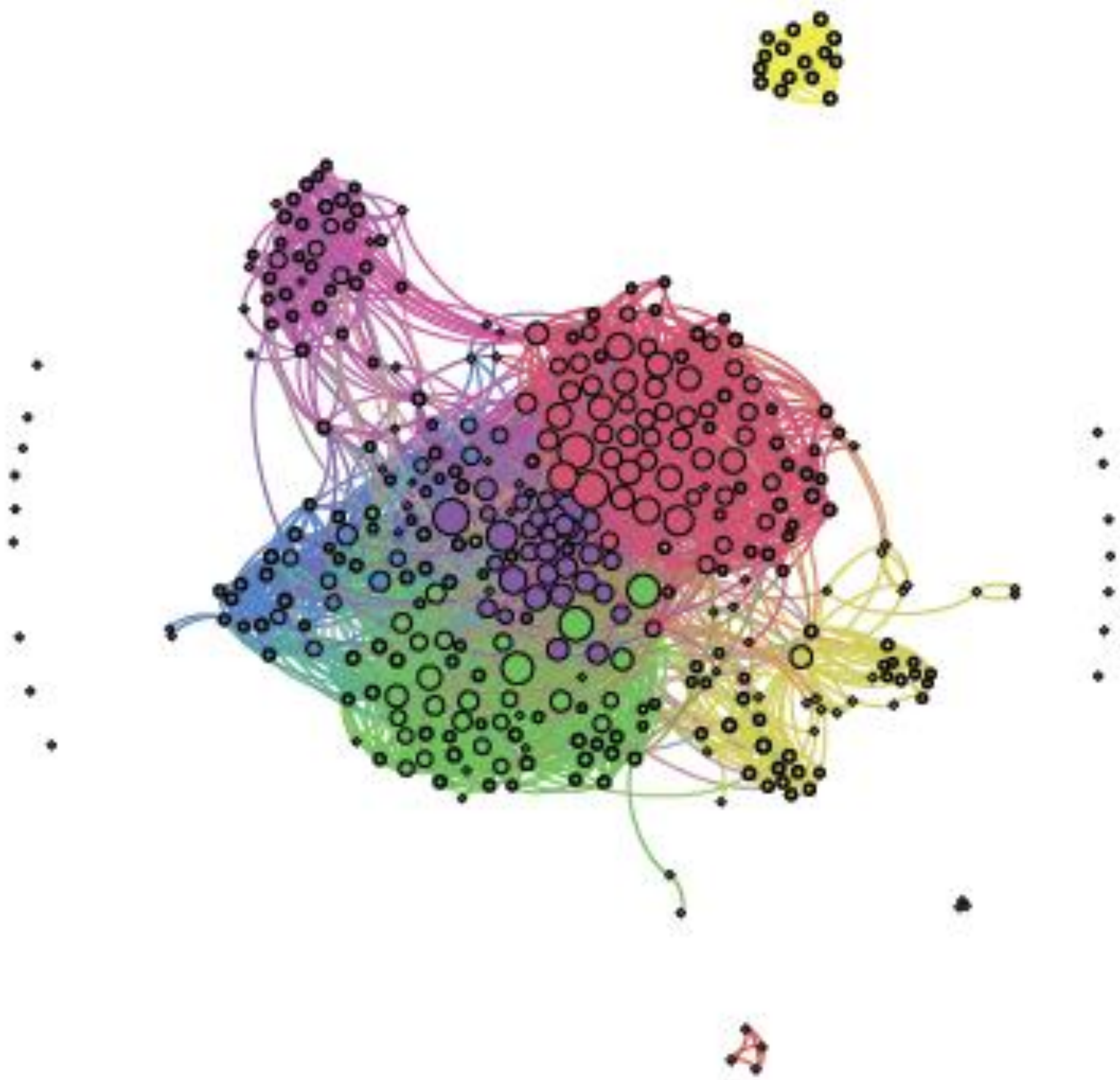
- Friendster



"Vizster: Visualizing Online Social Networks." Jeffrey Heer and danah boyd. *IEEE Symposium on Information Visualization (InfoViz 2005).*



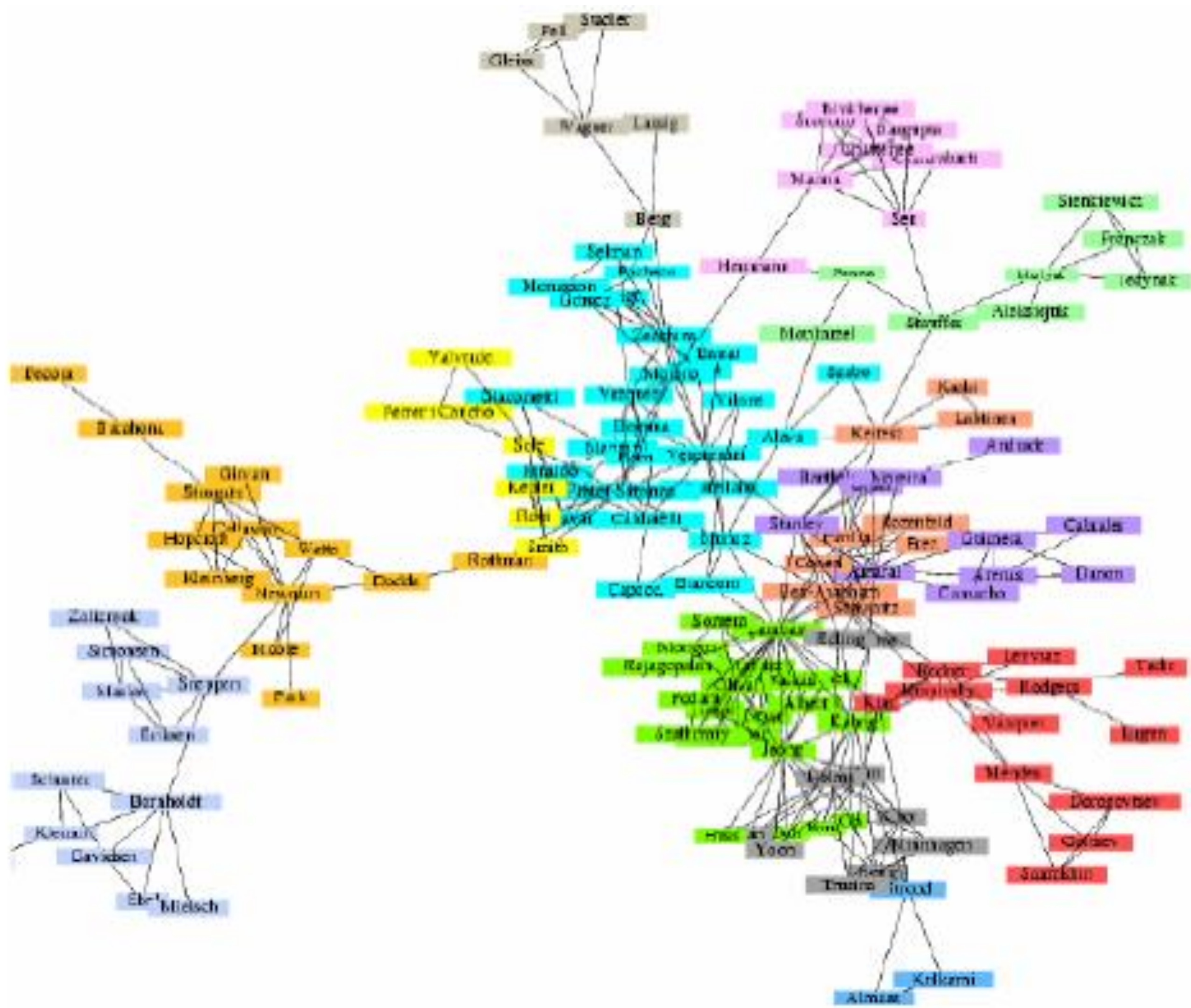
Ego networks



The "Social Graph" behind Facebook

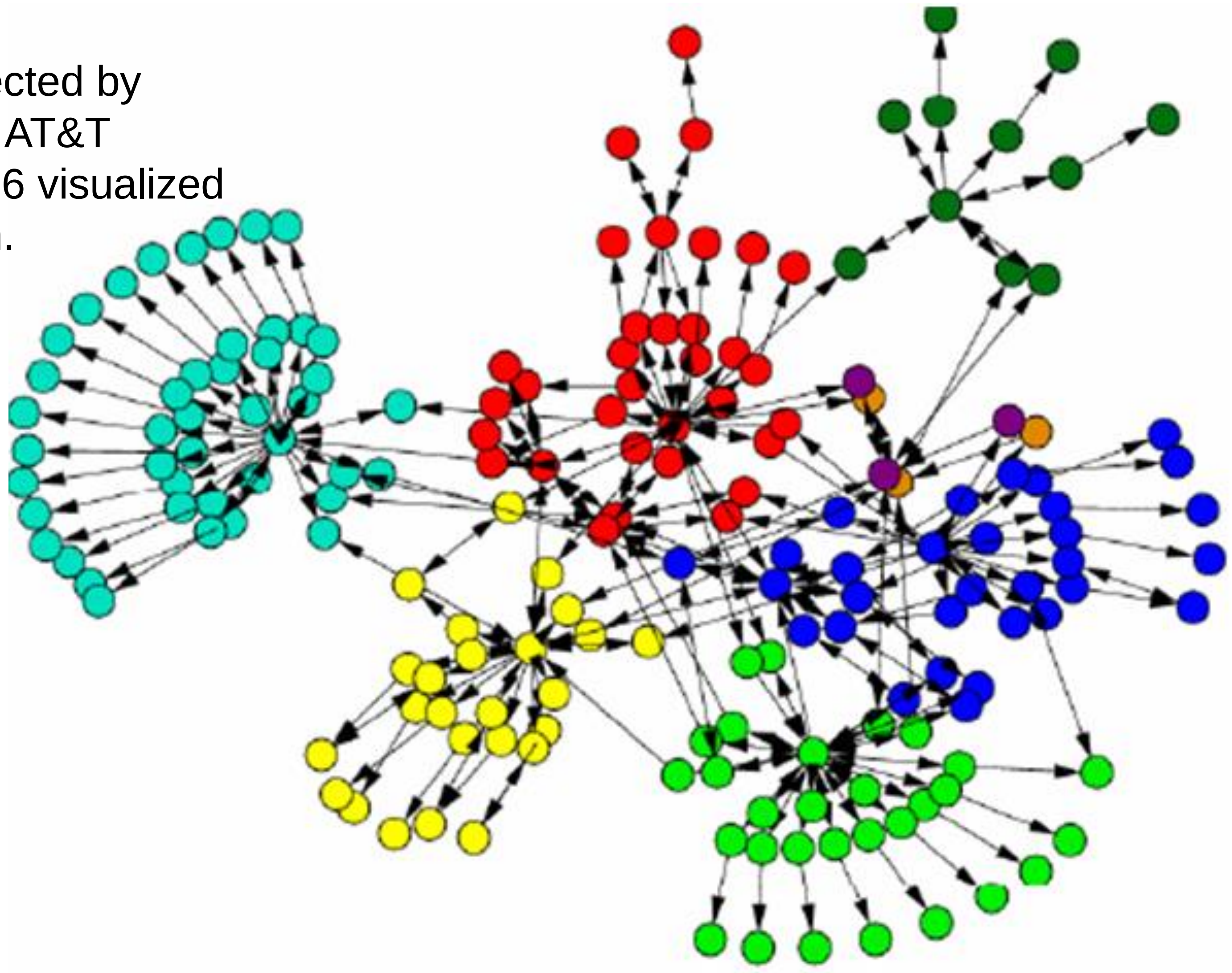
Keith Shepherd's "Sunday Best”. http://baseballart.com/2010/07/shades-of-greatness-a-story-that-needed-to-be-told/

Co-authorship network

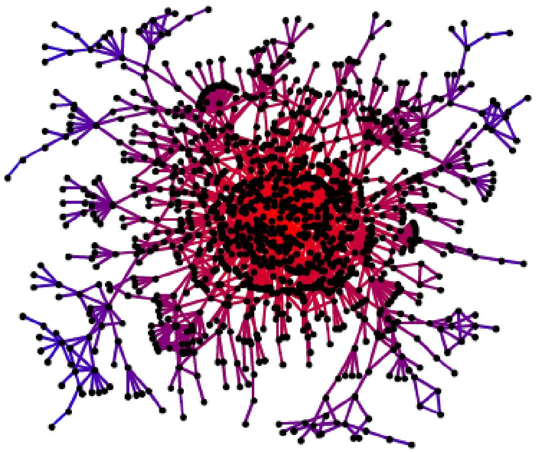


Webpages

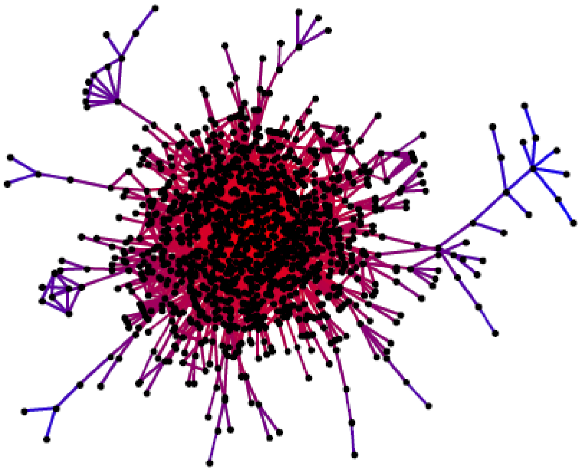
Webpages connected by hyperlinks on the AT&T website circa 1996 visualized by Mark Newman.



Networks of personal homepages



Stanford



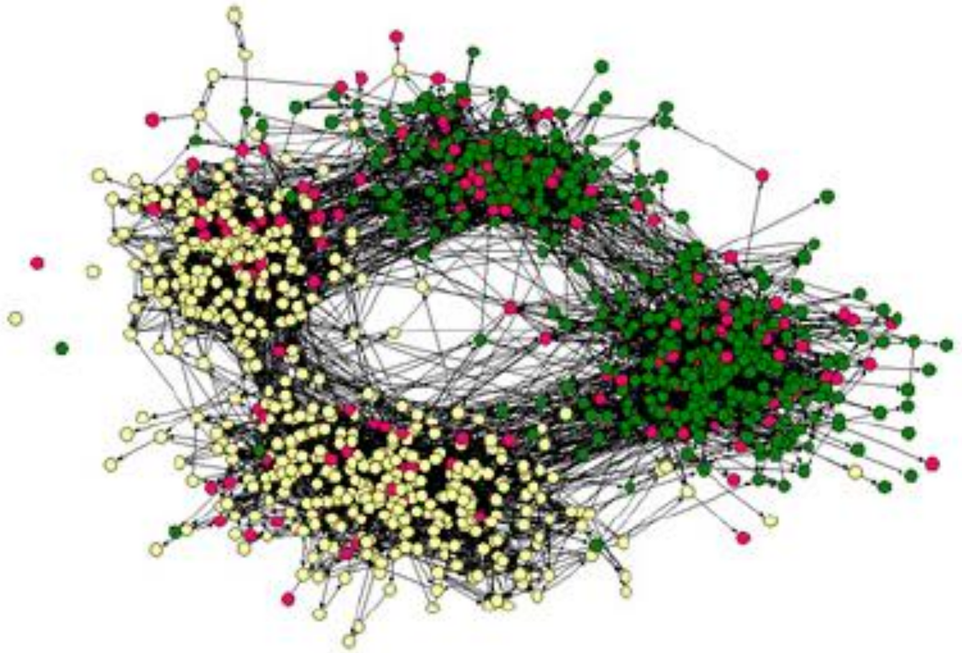
MIT

homophily: what attributes are predictive of friendship?

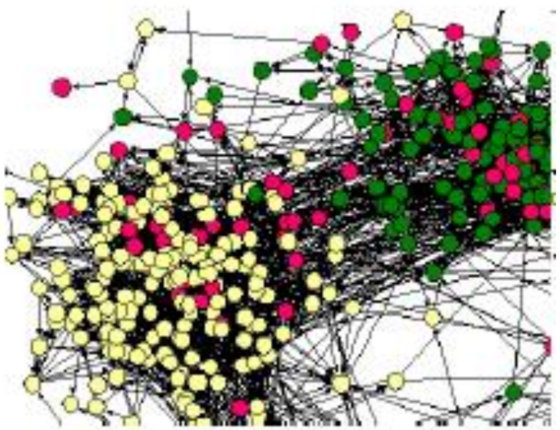
group cohesion

Lada A. Adamic and Eytan Adar, 'Friends and neighbors on the web', *Social Networks*, 25(3):211-230, 2003.

Friendship network of children in a school

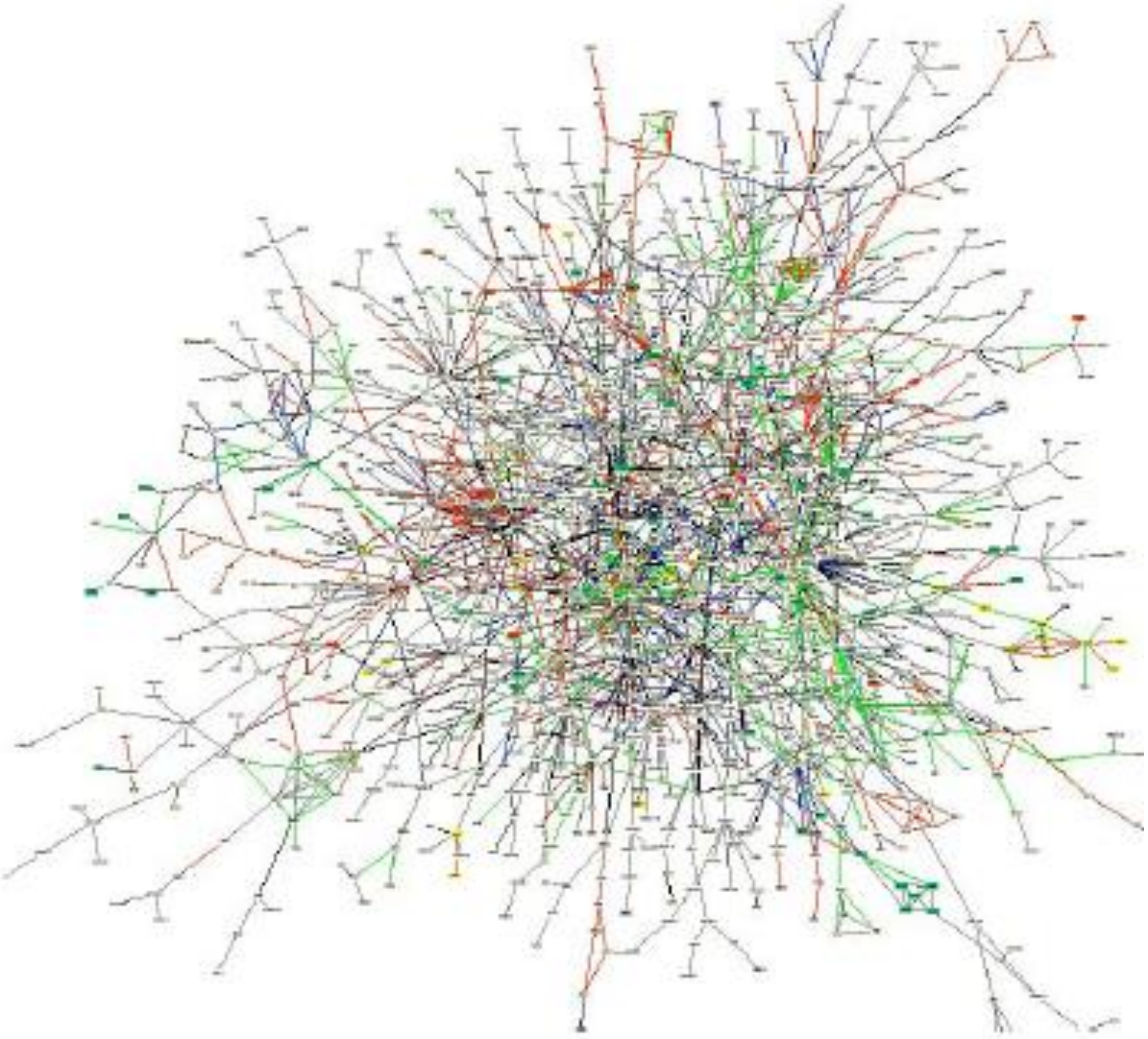


Yellow - White Race
Green - Black Race
Pink - Other

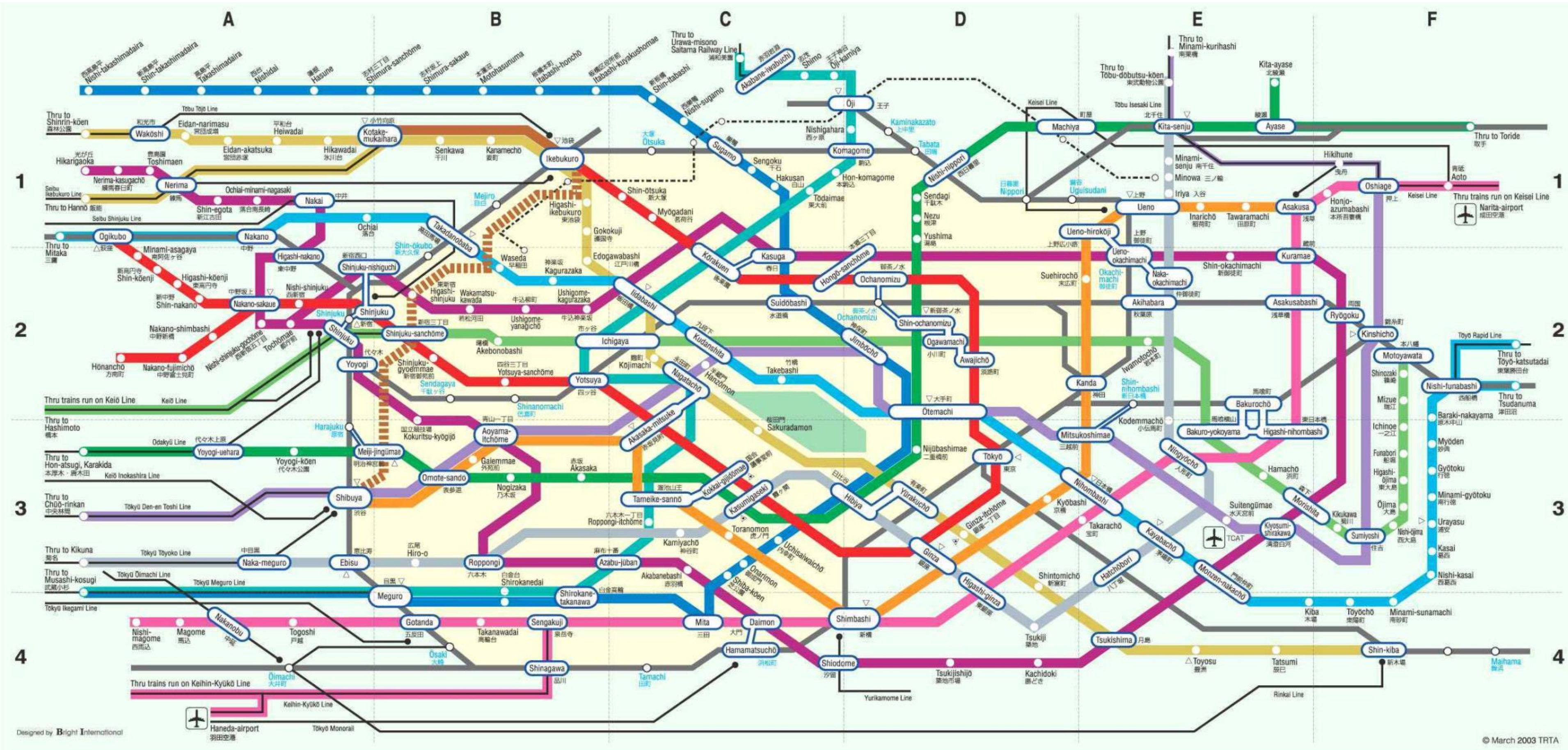


http://www.visualcomplexity.com/vc/

Yeast protein-protein interaction network



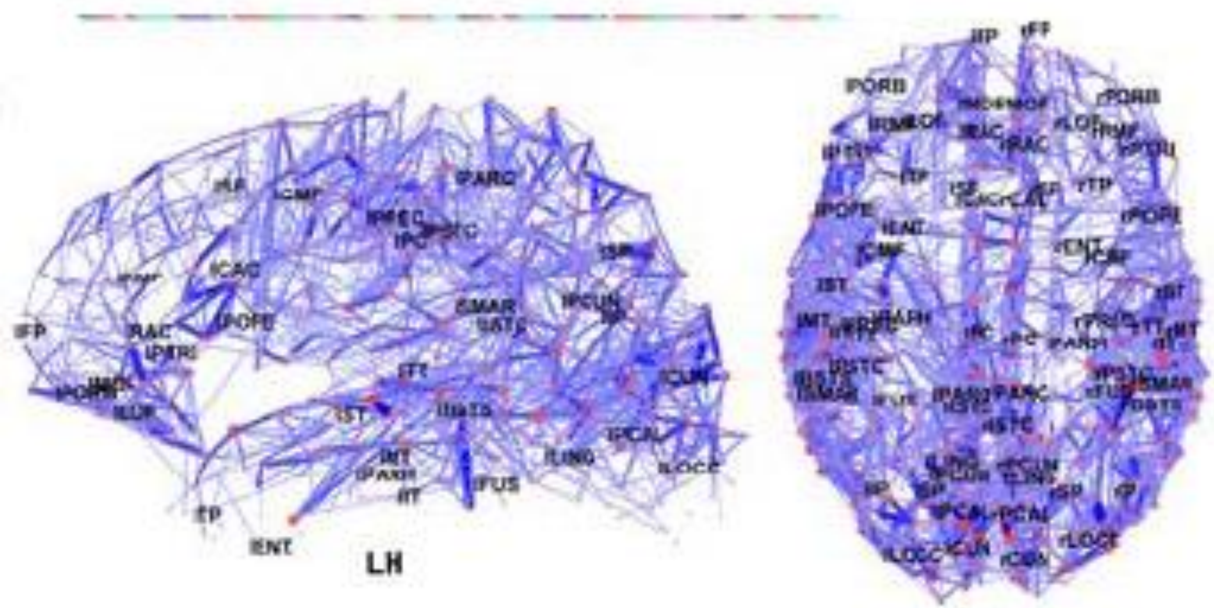
Railway networks

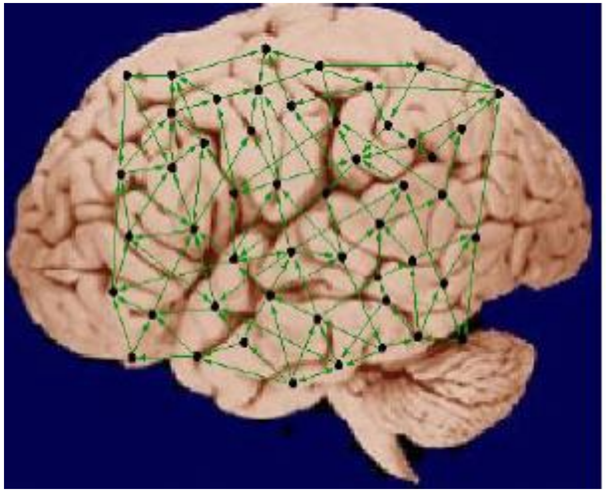


Source: TRTA, March 2003 - Tokyo rail map

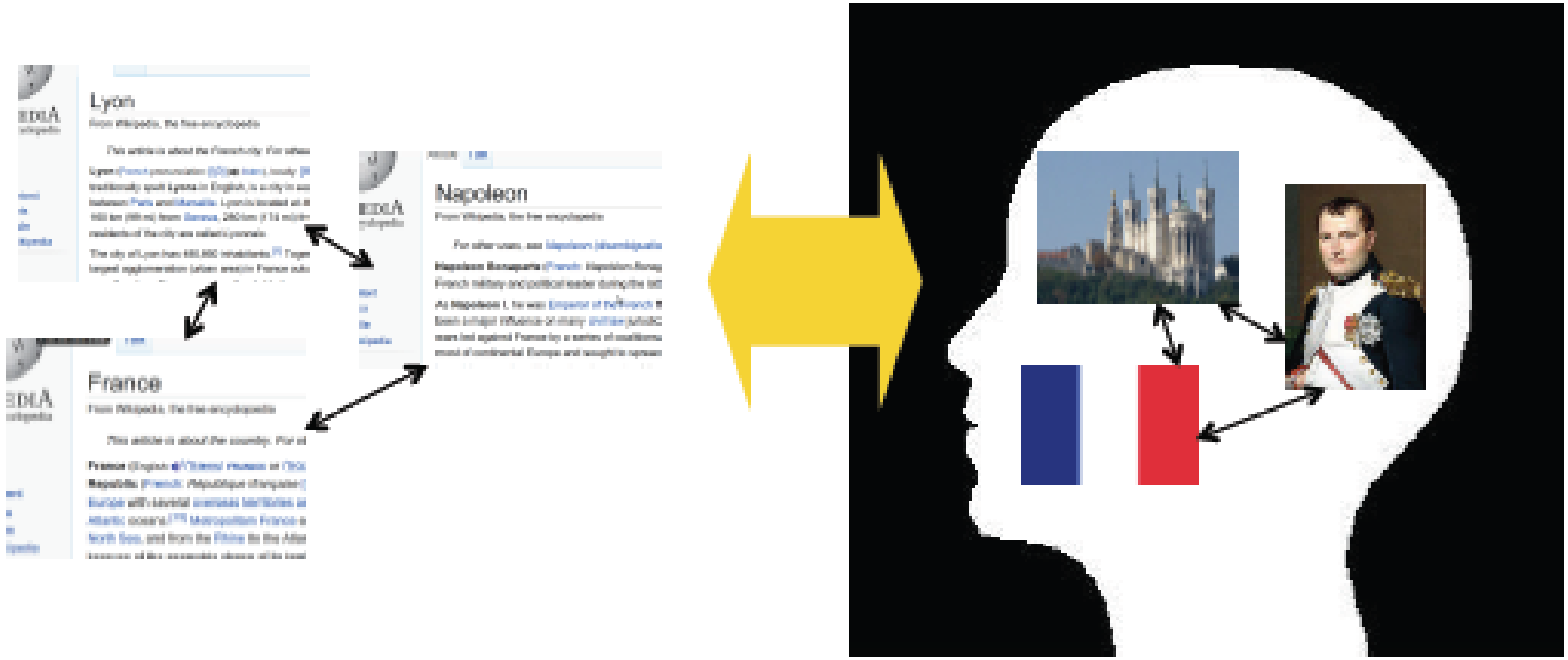
Brain networks

- **Human Brain has between 10-100 billion neurons.**





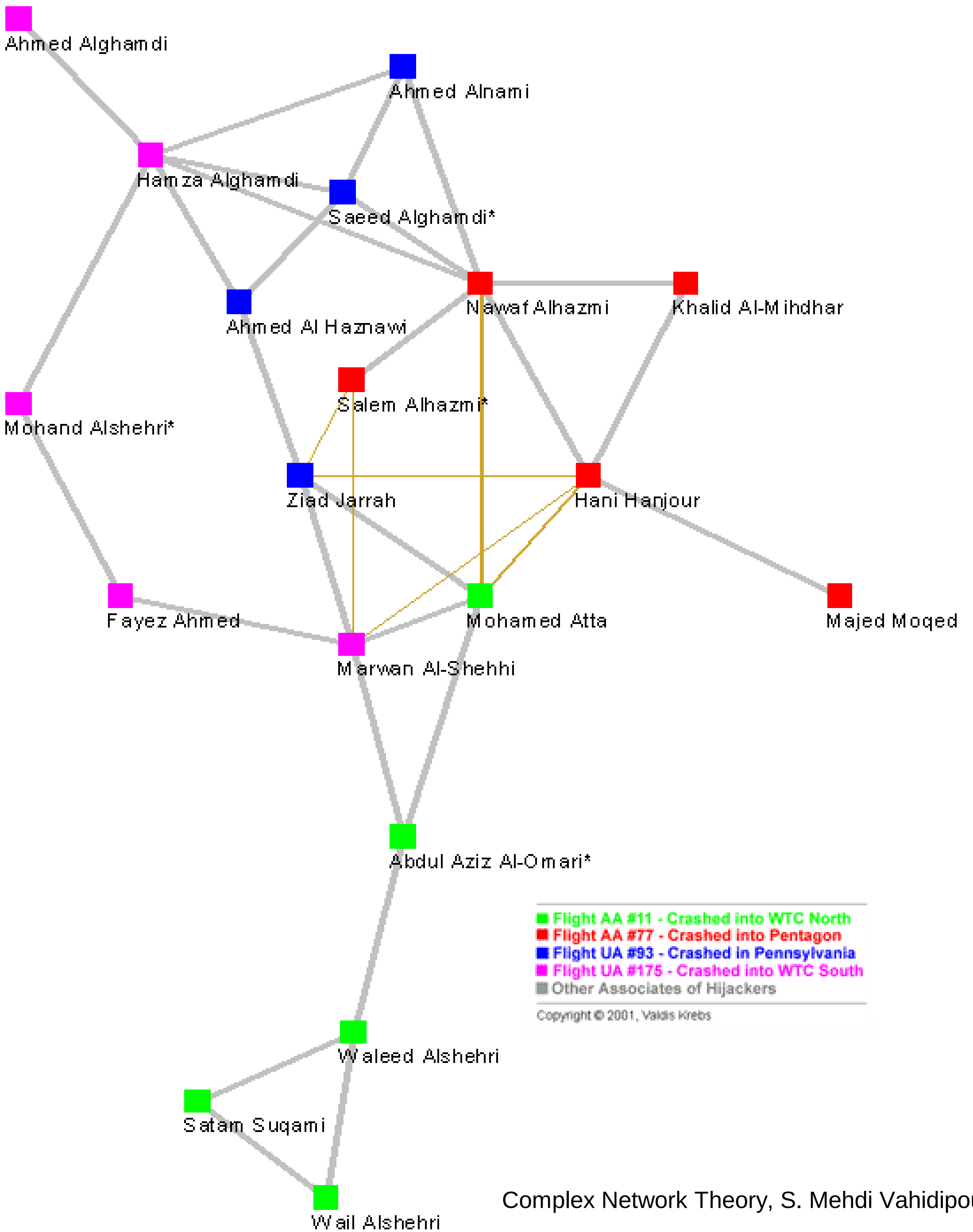
Knowledge networks



Understand how humans navigate Wikipedia

Get an idea of how people connect concepts

Source: [West-Leskovec, 2012], Jure Leskovec, Stanford CS224W: Social and Information Network Analysis

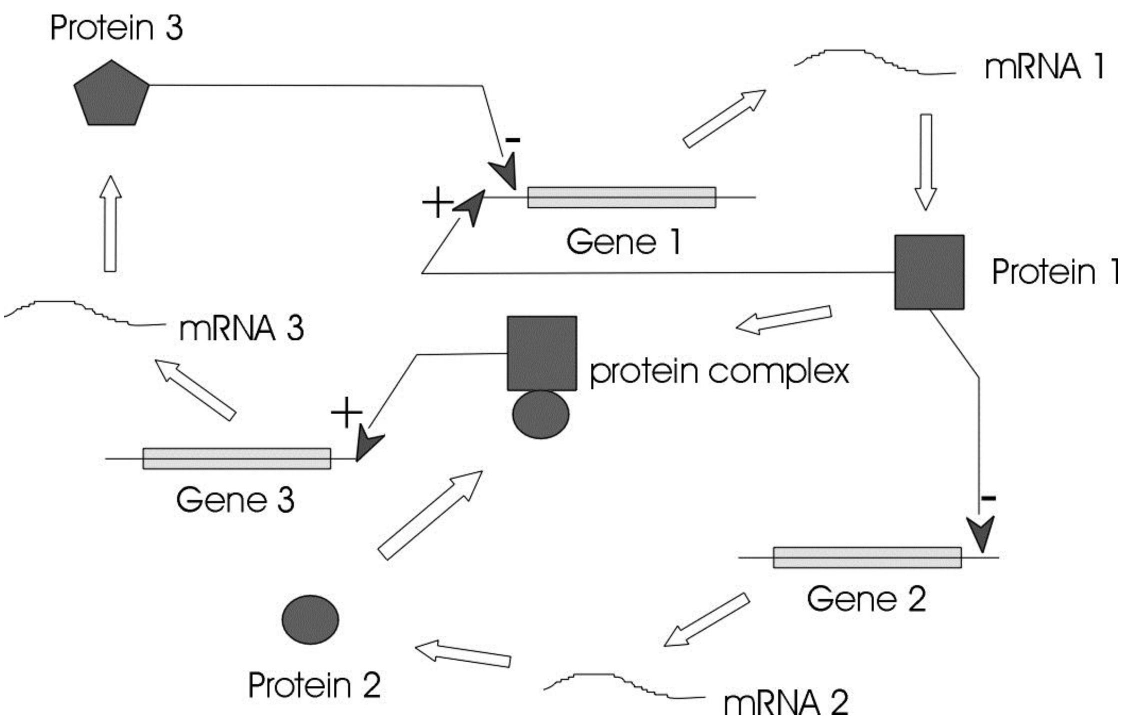


"Six degrees of Mohammed Atta"

Uncloaking Terrorist Networks, by Valids Krebs

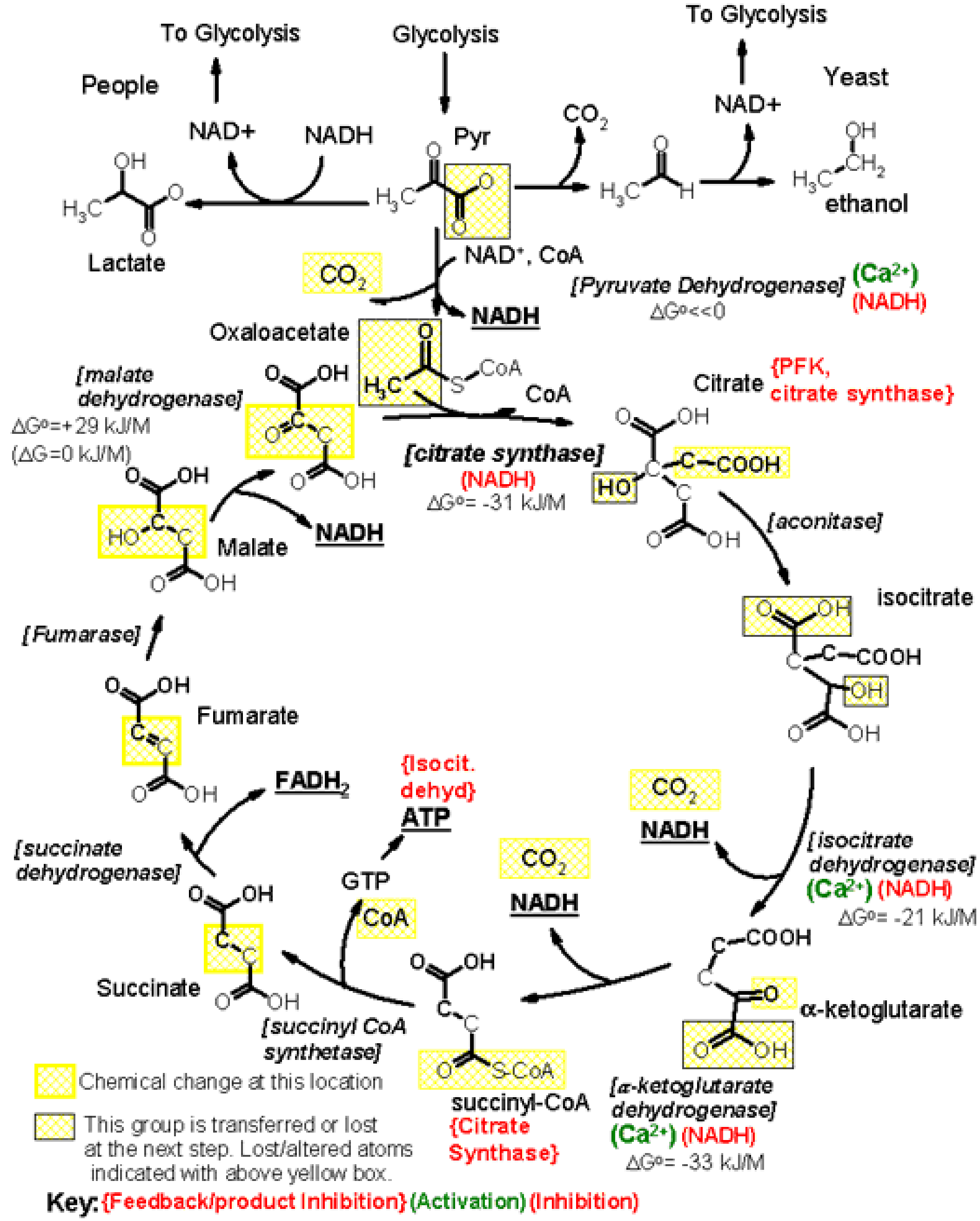
research in biological networks

- gene regulatory networks
  - humans have only 30,000 genes, 98% shared with chimps
  - the complexity is in the interaction of genes
  - can we predict what result of the inhibition of one gene will be?

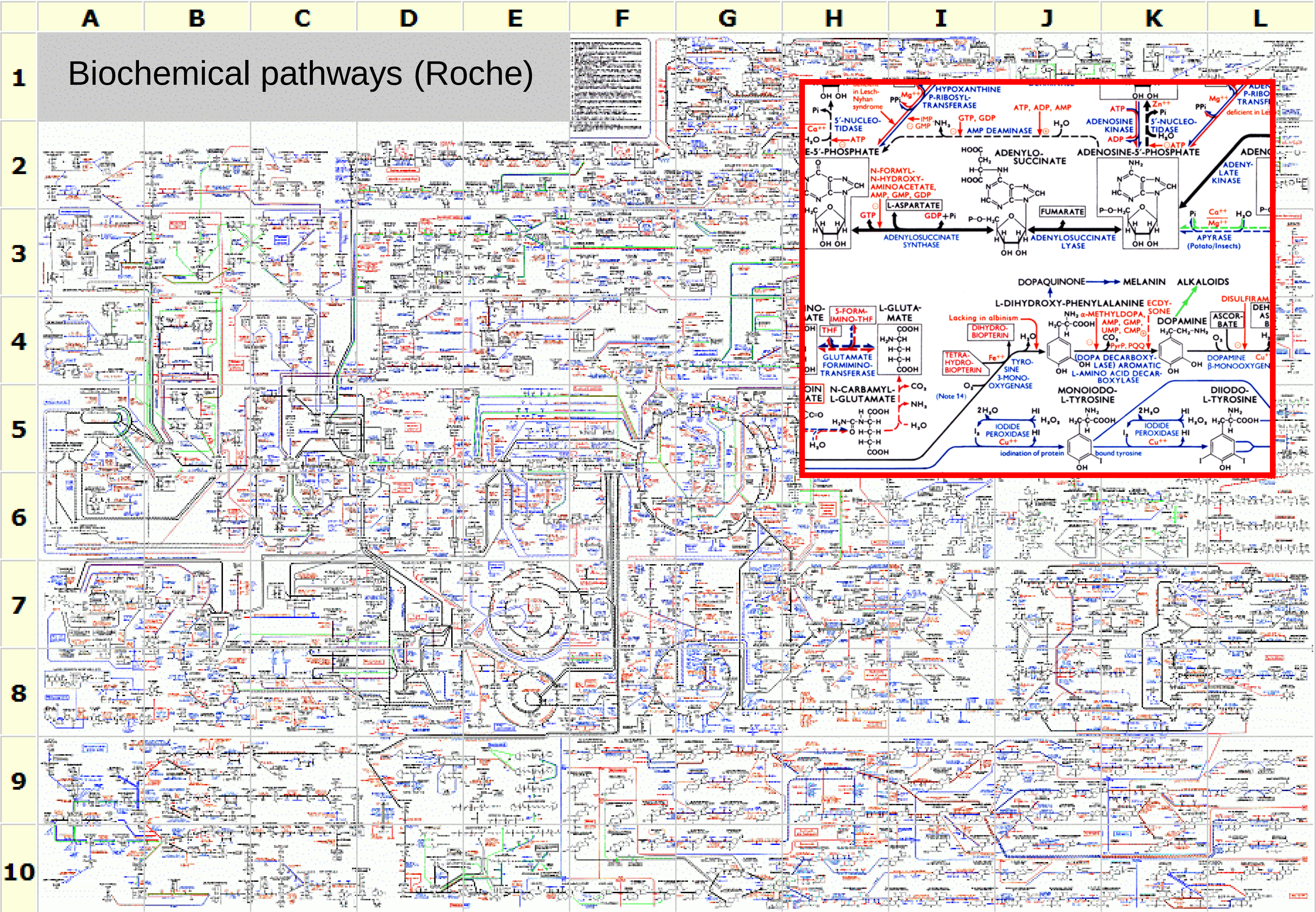


Other biological networks

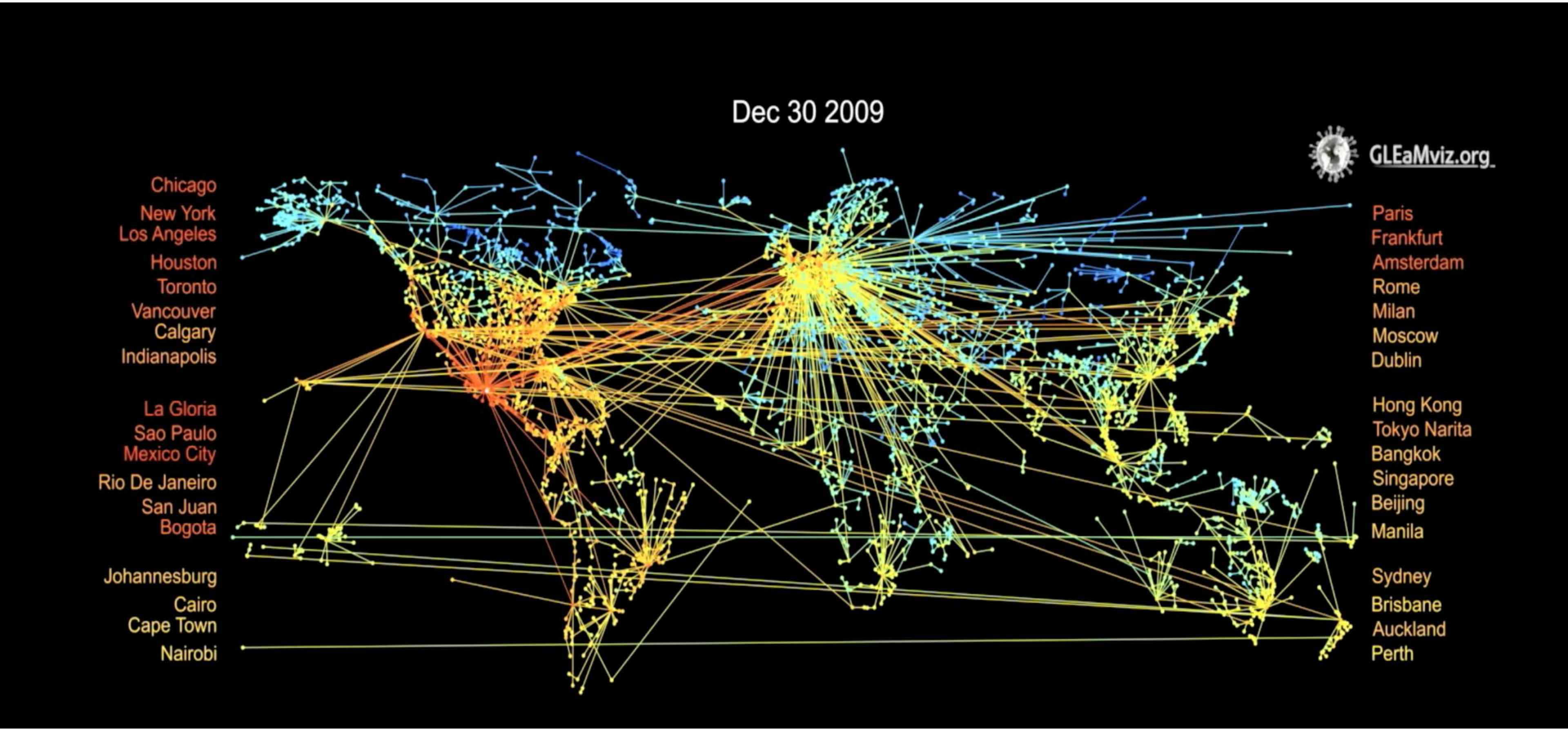
- Citric acid cycle
- Metabolites participate in chemical reactions



Biochemical pathways (Roche)

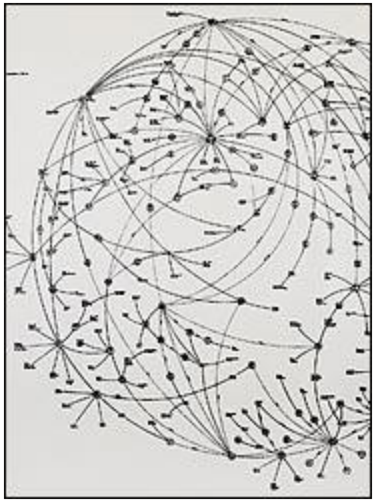


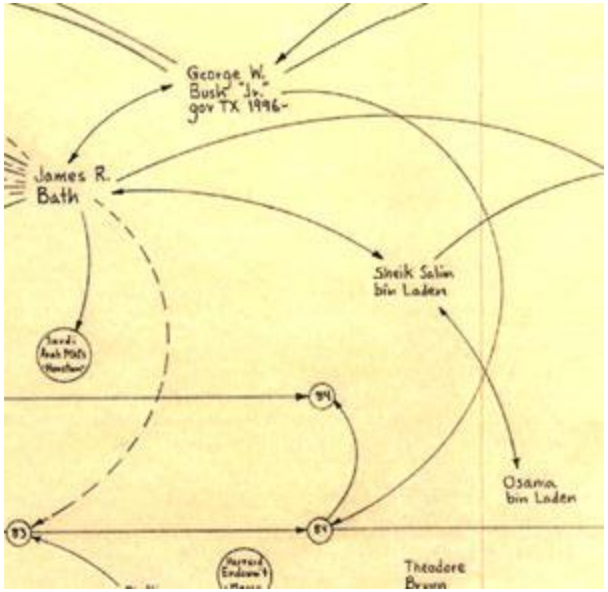
Predicting the H1N1 pandemic



Political/Financial Networks

- Mark Lombardi: tracked and mapped global financial fiascos in the 1980s and 1990s (committed suicide 2000)
- searched public sources such as news articles
- drew networks by hand (some drawings as wide as 10ft)
- Book: Hobbs, Robert.: Mark Lombardi :global networks /Robert Hobbs.. New York : Independent Curators International, c2003..





Networks

Behind many systems there is an intricate wiring diagram, **a network**, that defines the **interactions** between the components

We will never understand these systems unless we understand the networks behind them!

Complex systems

Complex

[adj., v. kuh m-pleks, kom-pleks; n. kom-pleks]
–adjective

1.
composed of many interconnected parts; compound; composite: a complex highway system.

2.
characterized by a very complicated or involved arrangement of parts, units, etc.: complex machinery.

3.
so complicated or intricate as to be hard to understand or deal with: a complex problem.

Source: Dictionary.com

Complexity, a **scientific theory** which asserts that some systems display behavioral phenomena that are completely inexplicable by any conventional analysis of the systems' constituent parts. These phenomena, commonly referred to as emergent behaviour, seem to occur in many complex systems involving living organisms, such as a stock market or the human brain.

Source: John L. Casti, Encyclopædia Britannica

Complexity

- No blueprint, No "master-mind", **Self-organization, Evolution, Adaptation**
- I think the next century will be the century of complexity (Stephen Hawking, January 23, 2000)

Why networks? Why now?

- Universal language for describing complex data
  - Networks from science, nature, and technology are more similar than one would expect
- Shared vocabulary between fields
  - Computer Science, Social science, Physics, Economics, Statistics, Biology
- Data availability (/computational challenges)
  - Web/mobile, bio, health, and medical
- Impact!
  - Social networking, Social media, Drug design

Why Network Science is important?

- Reductionism: to disassemble nature in order to understand it.
- After spending a lot to disassemble nature, now we almost know the pieces. But we are still far from understanding nature as a whole.
- Today we recognize that nothing happens in isolation. Most events are interconnected
- A string of recent discoveries has forced us to acknowledge that amazingly simple and far-reaching natural laws govern the structure and evolution of all the complex networks that surround us (social networks)
- Nowadays, power lines within the hands of the ones who know the network: Brokering, Public relation, Advertisement, Communication, & Etc.

Network questions: structural

- Given a network, there are a number of structural questions we may ask:
  1. What are the patterns and statistical properties of network data?
  2. How many connections does the average node have?
  3. Are some nodes more connected than others? (organization?)
  4. Is the entire network connected? (Models and principles?)
  5. On average, how many links are there between nodes?
  6. Are there clusters or groupings within which the connections are particularly strong?
  7. Is there any hierarchal structure?
  8. What is the best way to characterize a complex network?
  9. How can we tell if two networks are “different”?
  10. Are there useful ways of classifying or categorizing networks?
  11. What are the important nodes and links?

Network questions: communities

1. Are there clusters or groupings within which the connections are particularly strong?
2. What is the best way to discover communities, especially in large networks?
3. How can we tell if these communities are statistically significant?
4. What do these clusters tell us in specific applications?
5. How we can optimize the number of communities?
6. Is there any method applicable to large networks composed on millions of nodes and edges?
7. Is there a way to discover overlapping communities?

Network questions: dynamics

1. How can we model the growth of networks?
2. What are the important features of networks that our models should capture?
3. Are there "universal" models of network growth? What details matter and what details don't?
4. To what extent are these models appropriate null models for statistical inference?
5. What's the deal with power laws, anyway?
6. How is the time-evolution of a network?
7. How the network properties affected by its dynamical evolution?
8. How can we predict behavior of network systems?

Network questions: dynamics

1. How do diseases, computer viruses, innovations, rumors, revolutions, and opinions propagate on networks?

2. What properties of networks are relevant to the answer of the above question?

3. If you wanted to prevent (or encourage) spread of something on a network, what should you do?

4. What types of networks are robust to random attack or failure?

5. What types of networks are robust to intentional and cascading attack?

6. How collective behavior such as synchronization emerges from interaction of dynamical systems over networks?

7. Does a social game such as Prisoner’s Dilemma survive on a network?

Network questions: algorithms

1. What types of networks are searchable or navigable?
2. What are good ways to visualize complex networks?
3. What are the optimal algorithms for computing network metrics and measures?
4. How does the random models work?
5. How does Google PageRank work?
6. If the internet were to double in size, would it still work?
7. There are also many domain-specific questions:

   1. Are networks a sensible way to think about gene regulation or protein interactions or food webs?

   2. What can social networks tell us about how people interact and form communities and make friends and enemies?

   3. Lots and lots of other theoretical and methodological questions ...

   4. What else can be viewed as a network? Many applications await ...

Analysis and reasoning about networks

- How do we reason about networks?
  - **Empirical:** Study network data to find organizational principles
    - How do we measure and quantify networks?
  - **Mathematical models:** Graph theory and statistical models
    - Models allow us to understand behaviors and distinguish surprising from expected phenomena
  - **Algorithms** for analyzing graphs
    - Hard computational challenges

How it all together

Properties	Models	Algorithms
Small diameter, Edge clustering	Small-world model, Erdös-Renyi model	Decentralized search
Scale-free	Preferential attachment, Copying model	PageRank, Hubs and authorities
Strength of weak ties, Core-periphery	Kronecker Graphs	Community detection: Girvan-Newman, Modularity
Densification power law, Shrinking diameters	Microscopic model of evolving networks	Link prediction, Supervised random walks
Patterns of signed edge creation	Structural balance, Theory of status	Models for predicting edge signs
Information virality, Memetracking	Independent cascade model, Game theoretic model	Influence maximization, Outbreak detection

Software and Tools

- Network analysis tools
  - Gephi
  - Cytoscape
  - NodeXl
  - Yed
  - Pajek
  - UCI Net
  - Netdraw
  - Network Workbench
- Graph and network libraries
  - Matlab_BGL (MATLAB)
  - igraph (R)
  - SNAP library (C++)
  - NetworkX (python)

Famous researchers

Mark Newman, 146, 55

- University of Michigan

Albert Laszlo Barabasi, 198, 72

- Harvard University

Jon Kleinberg, 150, 40

- Cornell University

Réka Albert, 78, 31

- Pennsylvania State University

Santo Fortunato, 65, 23

- Institute for Scientific Interchange Foundation

Steven H. Strogatz, 107, 29

- Cornell University

Duncan Watts, 43, 25

- Yahoo Research Labs

Alex Arenas, 92, 27

- University of Zaragoza

Jure Leskovec, 54, 19

- Stanford University

David Kempe, 53, 17

- University of Southern California

Research centers

- Stanford Network Analysis, Stanford University (Leskovec)
- Social Media Lab, Cornell University (Kleinberg, Strogatz)
- Center for the Study of Complex Systems, University of Michigan (Newman)
- Center for Infectious Disease Dynamics, Pennsylvania State University (Albert)
- Complex Systems and Networks Lab, University of Zaragoza (Arenas)
- Center for Complex Network Research, Harvard University (Barabasi)

Resources

Journals

- **Social Network Analysis and Mining**, Springer
- **Journal of Complex Networks**, Oxford University Press
- **Network Science**, Cambridge University Press
- **Knowledge-Based Systems**, Elsevier
- **IEEE Transactions on Computational Social Systems,** IEEE
- **IEEE Transactions on Knowledge and Data Engineering**, IEEE
- **IEEE Transactions on Cybernetics**, IEEE
- **Computer Communications,** Elsevier
- **Computer Networks,** Elsevier,
- **IEEE Journal on Selected Areas in Communications**, IEEE
- **Physica A: Statistical Mechanics and its Applications**, Elsevier
- **Physical Review E- Statistical, Nonlinear, and Soft Matter Physics**, American Physical Society

Conferences

- IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)
- International World Wide Web Conference (WWW)
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)
- IEEE Conference on Applications of Social Network Analysis (ASNA)
- International Conference on Computational Aspects of Social Networks (CASoN)
- International Conference on Complex Networks (COMPLEX)

References (Books)

- Newman, Mark. **Networks: an introduction**. Oxford University Press, 2010.
- Easley, David, and Jon Kleinberg. **Networks, crowds, and markets: Reasoning about a highly connected world**. Cambridge University Press, 2010.
- Barabási, Albert-László, and Jennifer Frangos. **Linked: the new science of networks science of networks**. Basic Books, 2014.
- Silva, Thiago Christiano, and Liang Zhao. **Machine Learning in Complex Networks**. Springer, 2016.
- Wasserman, Stanley, and Katherine Faust. **Social network analysis: Methods and applications**. Cambridge university press, 1994.
- Aggarwal, C.C., "**Social Network Data Analytics**", Springer-Verlag, 2011
- Alhajj, Reda, and Jon Rokne, eds. **Encyclopedia of social network analysis and mining**. Springer, 2014.
- Mieghem, P. V., "**Graph spectra for Complex Networks**", Cambridge University Press, 2011.
- Gros, C., "**Complex and adaptive dynamical systems: A primer**", Springer Verlag, 2010.

References (Papers)

- Watts, Duncan J., and Steven H. Strogatz. "**Collective dynamics of 'small-world' networks**." *nature* 393.6684 (1998): 440-442.
- Barabási, Albert-László, and Réka Albert. "**Emergence of scaling in random networks**." *science* 286.5439 (1999): 509-512.
- Albert, Réka, Hawoong Jeong, and Albert-László Barabási. "**Error and attack tolerance of complex networks**." *nature* 406.6794 (2000): 378-382.
- Albert, Réka, and Albert-László Barabási. "**Statistical mechanics of complex networks**." *Reviews of modern physics* 74.1 (2002): 47.
- Newman, Mark EJ. "**The structure and function of complex networks**." *SIAM review* 45.2 (2003): 167-256.

References (Papers)

- Milo, Ron, et al. "**Network motifs: simple building blocks of complex networks." Science** 298.5594 (2002): 824-827.
- Costa, L. da F., et al. "**Characterization of complex networks: A survey of measurements**." *Advances in Physics* 56.1 (2007): 167-242.
- Costa, Luciano da Fontoura, et al. "**Analyzing and modeling real-world phenomena with complex networks: a survey of applications**." *Advances in Physics* 60.3 (2011): 329-412.
- Fortunato, Santo. "**Community detection in graphs**." *Physics Reports* 486.3 (2010): 75-174.
- Lü, Linyuan, and Tao Zhou. "**Link prediction in complex networks: A survey**."*Physica A: Statistical Mechanics and its Applications* 390.6 (2011): 1150-1170.

Tentative List of Topics

- Random Networks, random graphs
- Small world networks, Scale-Free networks
- Signed networks
- Network centrality measures
- Models of network formation: Barabasi-Albert model, Copying network model
- Measuring Network Properties (Centrality metrics, Similarity, Homophily,
- Assortative mixing, Reciprocity, Motifs, and Signed links)
- Cascading Behavior - Decision Based Models of Cascades
- Epidemics Models, Influence Maximization, information diffusion, Outbreak Detection, linear threshold model, independent cascade model
- Crowd Sourcing
- Network Sampling
- Spectral Analysis of Networks
- Community Detection
- Link Prediction, Network Inference, recommender systems
- Heterogeneous networks
- Mining Information Networks
- Game Theory & Network Analysis

Tentative List of Topics

- Tie Strength and Network
- Homophily in social networks
- Influence maximization
- Signed networks
- Matching Markets
- PageRank
- Power Laws networks
- Cascading Behavior in Networks
- Cascading failure in Networks
- Diffusion in Networks
- Crowd sourcing
- Small-World networks
- Epidemics
- Community Discovery (detection, identification)
- Node classification
- Network Evolution
- Expert finding
- Link Prediction
- Privacy in Social Networks

- Visualizing Social Networks
- Data Mining in Social Media
- Text Mining in Social Networks
- Information networks
- Technological networks
- Biological networks
- Power networks
- Electrical networks
- Protein-protein interaction networks
- Social networks
- Financial networks
- Opportunistic networks
- Chemical networks
- Ecological networks
- Citation networks
- Geographical networks
- Multi layered networks
- Temporal networks

Reading

- Newman, Mark. ***Networks: an introduction***. Oxford University Press, 2010. (chapters 1-4)

- Easley, David, and Jon Kleinberg. ***Networks, crowds, and markets: Reasoning about a highly connected world***. Cambridge University Press, 2010. (chapter 1)