SC435
Introduction to Complex Networks

Mukesh Tiwari

Course Grading & Policy

Course policy

Evaluation Scheme: 70% exam (Insem + Final) + 30% project.

- **Exam:** One mid-semester (30%) and one final exam (40%).
- **Project:** The students would implement a relevant computational paper from the literature. The outcome of the project will be (i) documented and working code for the project (existing libraries can be used) (ii) report on the work carried out with proper citation to the references used to develop the understanding (iii) viva/presentation.
- **Attendance Policy:** Any student with less than **75%** of attendance in lectures will not be allowed to appear for the final exam.
- **Plagiarism and cheating policy:** Plagiarism and/or cheating will result in disqualification from the course.

Project: Important Dates

- Abstract Submission: September 30
- Project Report Submission: November 7

Grading (tentative)

$> 85\%$	:	AA
$70\% - 85\%$	:	$BB - AB$
$55\% - 70\%$	:	$CC - BC$
$35\% - 50\%$	:	$DD - CD$
$< 35\%$	:	F

Basic Idea

What is the course about?

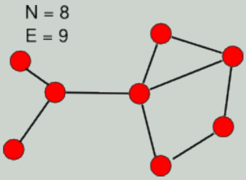
Complex

A system composed of many interacting components.

- Difficult to model: dependencies (hierarchical), nonlinear/stochastic interactions, competitions etc.
- Properties: Emergence, spontaneous order (bifurcation), adaptation etc.

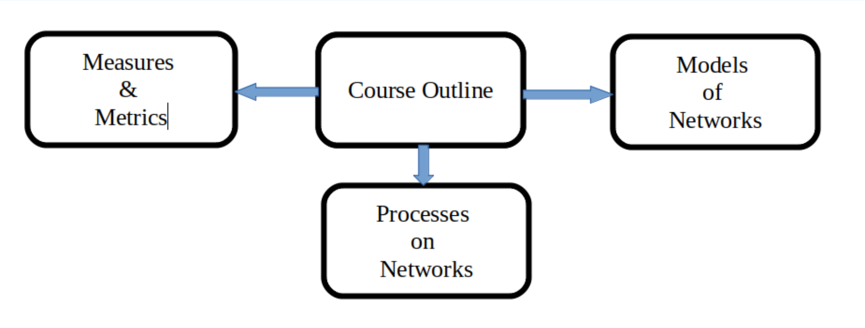
Network

Graphs:

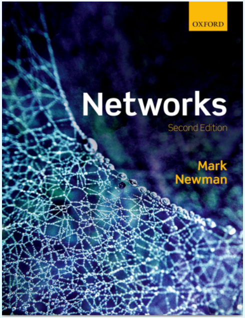


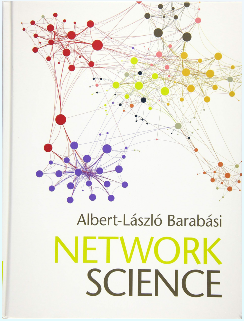
- Nodes: Entities
- Edge: interactions

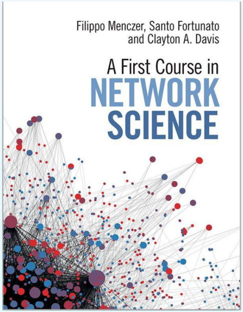
Road map


Measures
&
Metrics
Course Outline
Models
of
Networks
Processes
on
Networks

Books







Plethora of online material

A few online resources

Data

- Stanford Large Network Dataset Collection
- NetWiki
- KONECT
- Complex Networks Data Sets
- Uri Alon Lab (Weizmann Institute)
- Mark Newman's data set

Software

- Gephi
- NetworkX
- Cytoscape
- BCT
- Social network analysis software (Wikipedia)
- Pymnet(Multilayer networks)

People

- Albert-László Barabási
- Mark Newman
- Steven Strogatz
- Alessandro Vespignani
- Alain Barrat
- Guido Caldarelli
- Ernesto Estrada
- Brian Castellani

Should I consider this course as an elective?

Positives

Long term

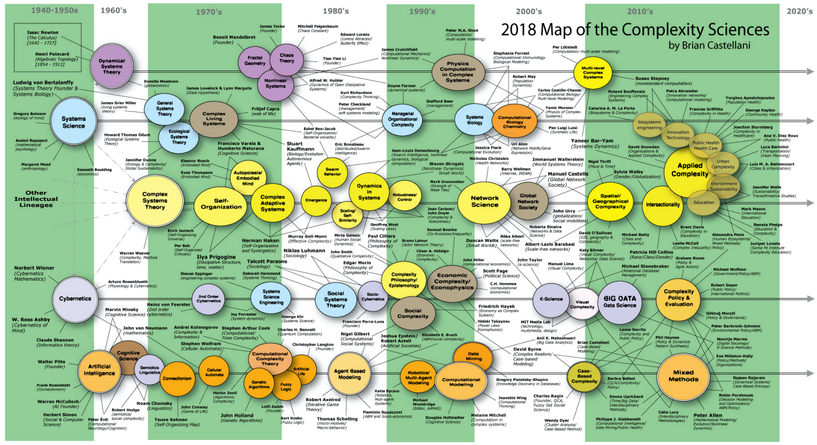
- Economic: companies base their technology on networks (Google, Facebook, Twitter)
- Health: Drug Design
- Epidemic: forecasting to control
- Mapping the Brain (Neuroscience)
- Policy making: rumor spread, propaganda, fake news

Short term

- Interest in Interdisciplinary sciences.
- Final year project

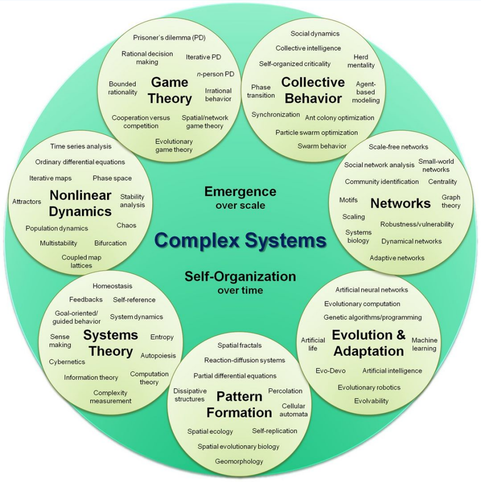
Negatives??

- Abundant online material/courses.
- Course may be theoretical at times.
- Course Intructor



1

[1] Source:Brian Castellani

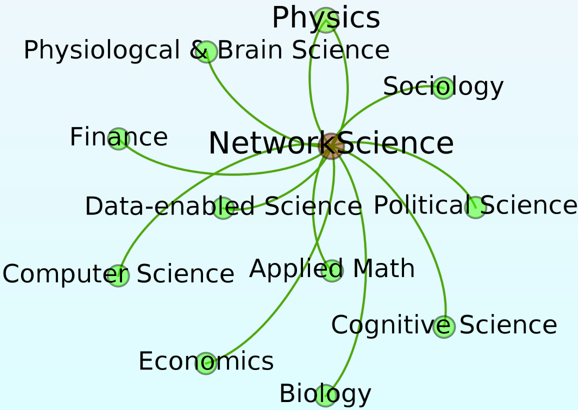


2

[2] Source:Wikipedia

Network Science

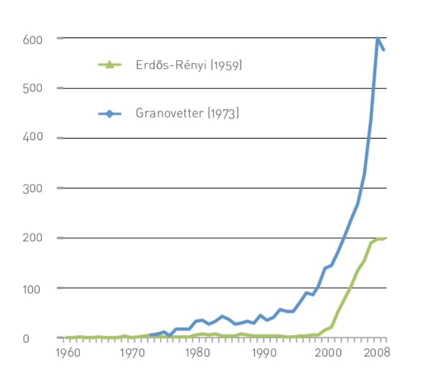
- Interdisciplinary
- Empirical, data driven
- Mathematical formalism from Graph Theory, Statistical Physics, Engineering and Statistics
- Large scale computation



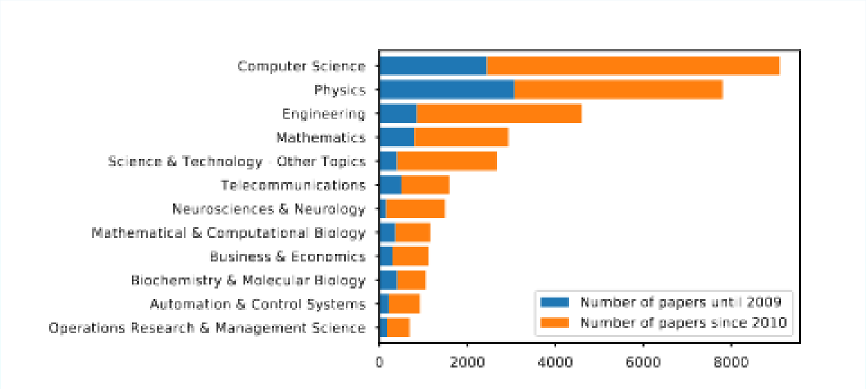
Not just Graph Theory

Emergence of Complex Network Research

- large electronically recorded sets of data
- advances in computer based simulation–*search for regularities and patterns in different complex systems*.
- Universality: the architecture of networks emerging in various domains of science, nature, and technology are similar–governed by the same organizing principles.
- common set of mathematical tools to explore these systems.



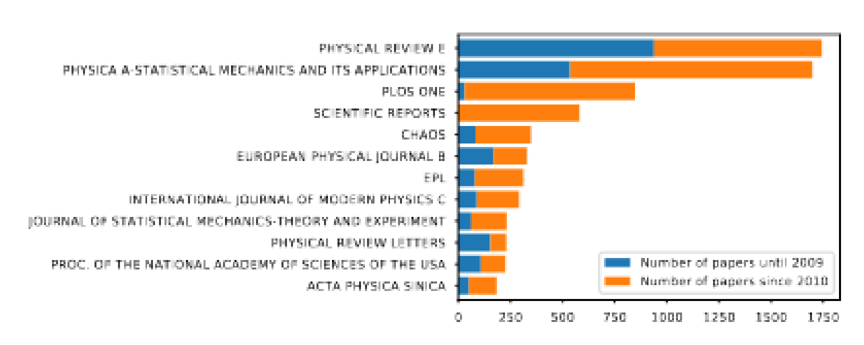
Emergence of Complex Network Research



3

[3] Roland et. al. Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2019

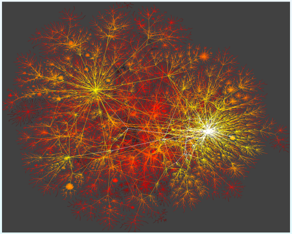
Emergence of Complex Network Research



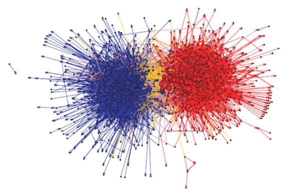
4

[4] Roland et. al Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2019

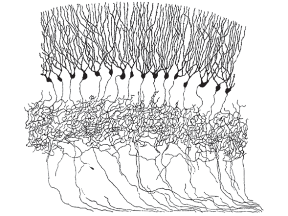
Internet



Political Blogs



Camillo Golgi (1906))



Protein

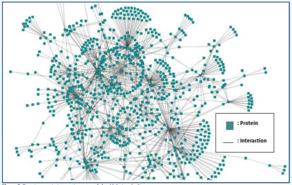
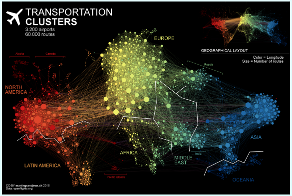


Figure 2: Protein-protein interaction network for Alzheimer's disease

World Airport

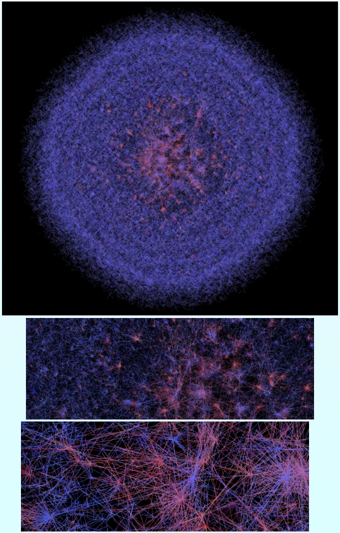


Example: Social Network

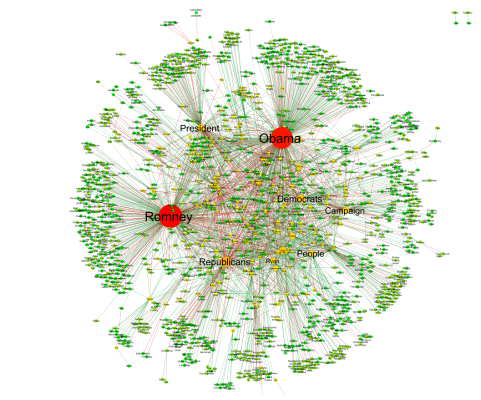
[DOUBT]
What the discussion is about?

Questions

- What is the type of the network?
- Are other social networks possible?
- What we expect to achieve by analyzing them?
- How is it useful?



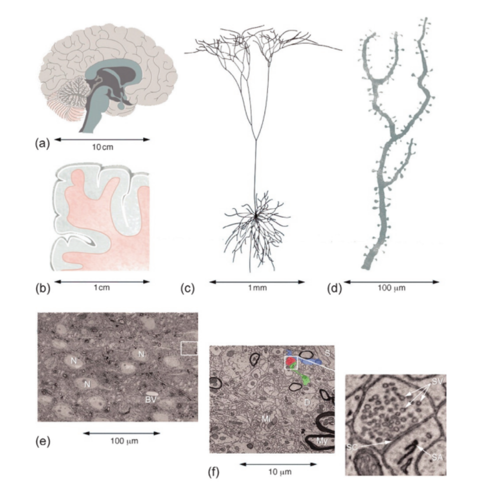
Example-Political Blog



Quantitative narrative analysis of US elections in International Media.

Sudhahar et. al. (2012)

Example-Brain Network



5

[5] Fornito et. al. Fundamental of Brain network Analysis

Questions

Natural questions to ask?

- What can we do with network data?
- What is its structure and topology?
- Is it large or small?
- Why does it have the features it has?
- How did it emerge and develop?
- What can we do with it?
- Are they robust against failure?
- Do they help in information flow?
- How to build these networks? (engineering complex networks)

Are all these of any use ?

Questions

Some deeper contextual questions

- Is social media the reason behind political polarization?
- Does fake(false) news lead to political polarization?
- Why does Government schemes not reach everyone?
- Can we understand conciousness? (local segregation vs global integration)
- Is to possible to identify precursors?

Analyzing networks

Concepts of practical importance

- hubs: A small number of vertices with extremely high degree.
- small world effect: on average geodesic distances are much smaller compared to the size of the network.
- communities or clusters: way a network breaks into communities.

What do we hope to achieve?

- Patterns and statistical properties of network data.
- design principles and models
- Understand how networks are organized the way they are
- control

Understanding of Networks

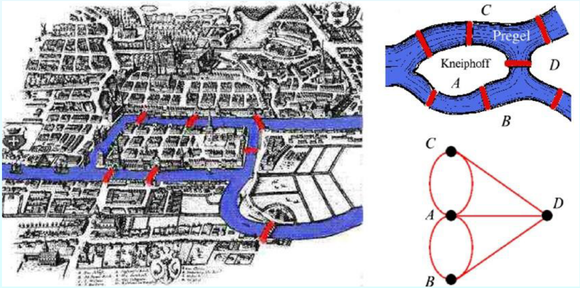
How to reason about networks?

- **Empirical:** Study network data $\Longrightarrow$ Statistical tools, Data Mining, Linear Algebra, Data Science, AI (Machine/Deep Learning)
- **Mathematical/Computational Modeling** $\Longrightarrow$ Probabilistic, Graph theory, Calculus, Statistical Physics.
- **Algorithmic** $\Longrightarrow$ Graph Theory (probability theory), Data Structures, Discrete Mathematics.

Some Historical Perspective

Königsberg Bridge problem
Euler,1736

- Königsberg was built on the river Pregel and on two islands that lie midstream, seven bridges connected the land masses.
- Does there exist any walking route that crosses all seven bridges exactly once?
- Leonhard Euler mapped the problem to a graph.



Leonard Euler, 1736

In addition to that branch of geometry which is concerned with magnitudes, and which has always received the greatest attention, there is another branch, previously almost unknown, which Leibniz first mentioned, calling it the geometry of position. This branch is concerned only with the determination of position and its properties; it does not involve measurements, nor calculations made with them. It has not yet been satisfactorily determined what kind of problems are relevant to this geometry of position, or what methods should be used in solving them.

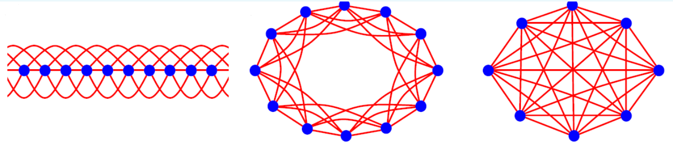
Geometria situs: Geometry of position(Graph Theory)

Historical Overview

History

1. 1736 - 1966 Mathematics of graphs.
2. 1967 - 1988 Application of network appeared from research literature.
3. 1998 - present fundamentals set -forth and shown that they have meaning in real world, burst in interdisciplinary applications of network science.

Much of graph theory focused on the study of regular graphs.



- Regular graph: All nodes have the same degree.
- Lattice: A regular network where each node is connected to its nearest neighbor.

Networks in Evolution

Yule (1925)

- preferential attachment in evolution.

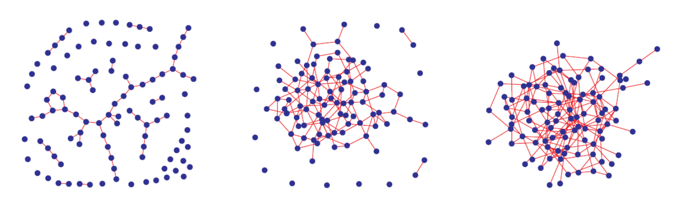
Kermack-McKendrick epidemic model (1927)

- Non-network but set the stage for two important innovations.
  - explained the spread of a contagion along (social) links connecting (individuals) nodes.
    virus spreading in network like internet
  - *coincidentally* described new -product adoption, diffusion of technology.
    how product information spreads in a network

Random graphs
mid twentieth century

Gilbert (1959): showed how to build a random graph by first constructing a complete graph and then deleting randomly a desired number of links.

Erdös-Rényi model (1960)



Increasing p...

- Erdös realized that if networks develop randomly, they are highly efficient.
- Even with a few connections on average per link, the network can be connected with small paths.
- Laid the foundation of modern graph theory.

Small World Experiment
Stanley Milgram (1967)

The Experiment:
A group of people from Omaha (Nebraska) and Wichita (Kansas) were asked to send a letter to an unknown person in Boston (Massachusetts)

Rules

- People were asked to forward a folder to a target person.
  - If the target person is known directly (if the two have met and know each other on a first name basis) then mail the folder directly.
  - Else the person should forward the folder to a personal acquaintance who is more likely to know the target person
- Result: Average number of intermediaries 5.2

Six-degrees of separation

Wattz-Strogatz small world networks

Collective dynamics of 'small-world' networks

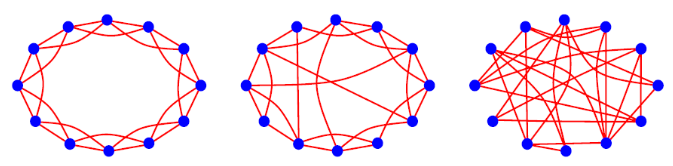
Duncan J. Watts* **& Steven H. Strogatz**

Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA

Networks of coupled dynamical systems have been used to model biological oscillators[1–4], Josephson junction arrays[5,6], excitable media[7], neural networks[8–10], spatial games[11], genetic control networks[12] and many other self-organizing systems. Ordinarily, the connection topology is assumed to be either completely regular or completely random. But many biological, technological and social networks lie somewhere between these two extremes. Here we explore simple models of networks that can be tuned through this middle ground: regular networks 'rewired' to introduce increasing amounts of disorder. We find that these systems can be highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs. We call them 'small-world' networks, by analogy with the small-world phenomenon[13,14] (popularly known as six degrees of separation[15]). The neural network of the worm *Caenorhabditis elegans*, the power grid of the western United States, and the collaboration graph of film actors are shown to be small-world networks. Models of dynamical systems with small-world coupling display enhanced signal-propagation speed, computational power, and synchronizability. In particular, infectious diseases spread more easily in small-world networks than in regular lattices.

Wattz-Strogatz small world networks

- How information travels across thousands even millions of entities. (Synchronization in Malaysian fireflies). Is there a leader?
- Is there a strong connection with six degrees of freedom (Information travels very fast?)



- conciliates high clustering and short average path length of real networks.
- Starting from a regular lattice with probability p two edges are rewired.

Barabasi - Albert Model

Emergence of Scaling in Random Networks

Albert-László Barabási* and Réka Albert

Systems as diverse as genetic networks or the World Wide Web are best described as networks with complex topology. A common property of many large networks is that the vertex connectivities follow a scale-free power-law distribution. This feature was found to be a consequence of two generic mechanisms: (i) networks expand continuously by the addition of new vertices, and (ii) new vertices attach preferentially to sites that are already well connected. A model based on these two ingredients reproduces the observed stationary scale-free distributions, which indicates that the development of large networks is governed by robust self-organizing phenomena that go beyond the particulars of the individual systems.

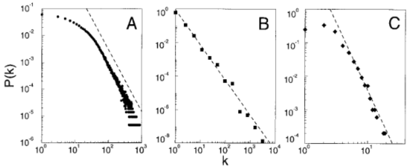


Fig. 1. The distribution function of connectivities for various large networks. (**A**) Actor collaboration graph with N = 212,250 vertices and average connectivity $\langle k \rangle$ = 28.78. (**B**) WWW, N = 325,729, $\langle k \rangle$ = 5.46 (6). (**C**) Power grid data, N = 4941, $\langle k \rangle$ = 2.67. The dashed lines have slopes (A) γ_{actor} = 2.3, (B) γ_{www} = 2.1 and (C) γ_{power} = 4.

What are Complex Networks?

Networks with non-trivial topological features, with patterns of connections between their elements that are neither purely regular nor purely random.