Ανάλυση Κοινωνικών Δικτύων (Social Network Analysis)

Εισαγωγή

Καθηγητής Συμεών Παπαβασιλείου

Εθνικό Μετσόβιο Πολυτεχνείο
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Επικοινωνιών, Ηλεκτρονικής & Συστημάτων Πληροφορικής

(E-mail: papavass@mail.ntua.gr Τηλ: 210 772-2550
Γραφείο: Β.3.15 Νέο Κτίριο Ηλεκτρολόγων)

18 Οκτωβρίου, 2019

Γενικά στοιχεία του μαθήματος

Προγραμματισμός Διαλέξεων & Εργαστηρίων

Διαλέξεις: κάθε Παρασκευή, 08:45-10:30, Αίθουσα 013 (Νέο Κτ. Ηλεκτρολόγων)

Έναρξη: Παρασκευή, 18 Οκτωβρίου 2019

Εργαστήριο: Μετά τις 2 πρώτες εβδομάδες μαθήματος, κάθε Τετάρτη, 12:45-14:30 (PC Lab Σχολής)

Μέθοδοι αξιολόγησης:

Ο βαθμός μαθήματος θα προκύψει από το βαθμό του εργαστηρίου (50%) και το βαθμό της εξέτασης στο «θεωρητικό» μέρος του μαθήματος (50%).

Βιβλιογραφία

[1] D. Easley and J. Kleinberg, "Networks, Crowds and Markets: Reasoning about a Highly Connected World", Cambridge University Press, 2010.

[2] V. Karyotis, E. Stai and S. Papavassiliou, "Evolutionary Dynamics of Complex Communications Networks", CRC Press, 2013.

http://mycourses.ntua.gr/course_description/index.php?cidReq=ECE1372

Θεματολογία (1/2)

1. Εισαγωγή στην Επιστήμη Δικτύων (Network Science) – περιεχόμενο και στόχοι
   a. Βασικοί ορισμοί δικτύων, ρόλος δικτύων και παραδείγματα εφαρμογών σε διαφορετικές εφαρμογές
   b. Έλεγχος τοπολογίας και δημιουργία δικτύων
2. Στοιχεία θεωρίας γραφημάτων
   a. Επισκόπηση βασικών ορισμών – ορολογία (μονοπάτι, μέσο μήκος μονοπατιού, διάμετρος, κτλ.)
   b. Δένδρα και ιδιότητες, συνδεδεμένοι γράφοι
   c. Βασικοί αλγόριθμοι δικτύων (π.χ. δρομολόγηση, αναζήτηση, κτλ.)
3. Δομή και χαρακτηριστικά σύνθετων και κοινωνικών δικτύων
   a. Τυχαία μοντέλα δικτύων, δίκτυα μικρού-κόσμου (small-world), δίκτυα ελεύθερης-κλίμακας (scale-free), κανονικά δίκτυα (regular), τυχαία γεωμετρικά δίκτυα (random geometric graphs), κτλ.
4. Στοιχεία ανάλυσης σύνθετων και κοινωνικών δικτύων
   a. Μετρικές ανάλυσης (κατανομή βαθμού κόμβου, συντελεστής συσσωμάτωσης, κεντρικότητα δικτύου, κτλ.)
   b. Επιλεκτική σύνδεση και δημιουργία/εξέλιξη δικτύων
5. Εξελικτικός υπολογισμός
   a. Γενετικοί αλγόριθμοι
   b. Επιγνωστικοί αλγόριθμοι
   c. Δυναμική πληθυσμών

Θεματολογία (2/2)

6. Εφαρμογές στις Τηλεπικοινωνίες και την Επιστήμη των Υπολογιστών
    a. Έλεγχος τοπολογίας, δρομολόγηση και ανάθεση πόρων
    b. Επίδραση δομής δικτύου στη διάδοση πληροφορίας/διαμόρφωσης γνώμης
    c. Επίδραση κοινωνικών δικτύων σε συστήματα σύστασης
7. Εφαρμογές στη Διάδοση Πληροφορίας
    a. Επιδημιολογικά μοντέλα πληροφορίας
    b. Συνεργασία και συγχρονισμός
    c. Επίδραση κοινωνικών δικτύων σε συστήματα διαφήμισης
8. Εργαστηριακές Ασκήσεις
    a. Συλλογή ελεύθερων/ανοιχτών δεδομένων από κοινωνικά δίκτυα
    b. Επεξεργασία δεδομένων και στατιστική ανάλυση
    c. Οπτικοποίηση συλλεγμένων δεδομένων και αποτελεσμάτων σε μορφή δεδομένων δικτύων (networked data)
    d. Παραδείγματα συλλογής, χρήσης και ανάλυσης δεδομένων (π.χ. μελέτη τοπολογιών και χαρακτηριστικών διαφόρων δικτύων, εντοπισμός κόμβων επιρροής δικτύου, ανίχνευση κοινοτήτων με παρόμοια χαρακτηριστικά, μελέτη διάδοσης πληροφορίας/διαμόρφωσης γνώμης, συστήματα και μέθοδοι κοινωνικής σύστασης).
    e. Εξελικτικός υπολογισμός με χρήση γενετικών αλγορίθμων

•

Complex Networks – Social Networks - Network Science

Section I.a
INTRODUCTION

Social Network motivation

- Based on “connectedness” of modern society due to several evolutions such as:
  - Rapid growth of Internet, Web, use of smartphones etc.
  - Ease that global communications take place
  - Ability of news and information, financial crises, epidemics, etc. to spread with high speed and intensity
- Involve networks – incentives – aggregate behavior of objects/people (multi-disciplinarity)
- Disciplines/Points of View
  - Computer Science/Engineering: study technological networks (Internet, telecommunication networks, wireless networks)
  - Economics: study how people’s behavior is affected by incentives and by their expectations about behavior of others
  - Social science: study characteristic structures and interactions that arise within groups and populations

However there is a … Long Dialogue - Multidisciplinary

- Network scientists
  - "Give me data, we‟ll model …"

    ["… and by the way, we are not interested in "details" like how the data was collected or if it can be used for our purpose …"]

- Mathematicians
  - "Give me (network) models, we‟ll do (rigorous) proofs …"

- Internet researchers
  - "Give me network models (with or w/o proofs), we‟ll use them."

    ["… and by the way, we‟ll ignore all the domain knowledge that we have about the Internet …"]

Complex Network definition

- Network:
  - Patterns of interconnections among a set of things, OR
  - any collection of objects in which some pairs of these objects are connected by links. Different forms of relationships can be used to define links
- Social Network:
  - collection of social ties among “friends”. These have grown in complexity due to technological advances facilitating distant travel, global communications, digital interaction.
- Complex Network: Set of interacting entities (actors, nodes, etc.)
  - Each node performs some complex computation
  - Collaborating $\rightarrow$ coalitions
  - Competing
- Emerging trade-off: **gain** vs. **cost** of collaboration
  - Gain obtained by collaboration/selfish behavior
  - Cost incurred by selfishness/collaboration
- How to describe such complex interactions?

Complex Network Taxonomy

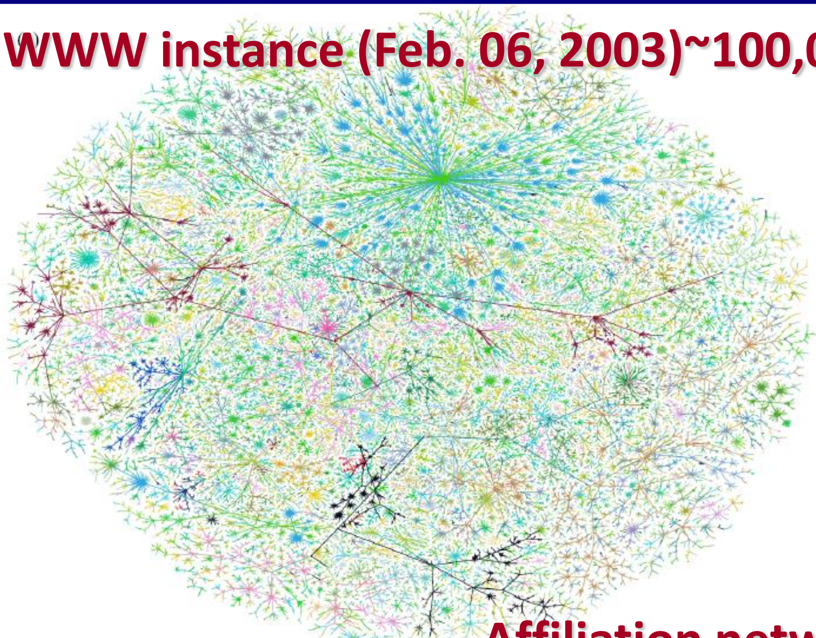
Communication, infrastructure, technological networks	Social and economical networks	Biological networks
Designed and/or engineered	Human initiated, but spontaneous growth	Spontaneous evolution

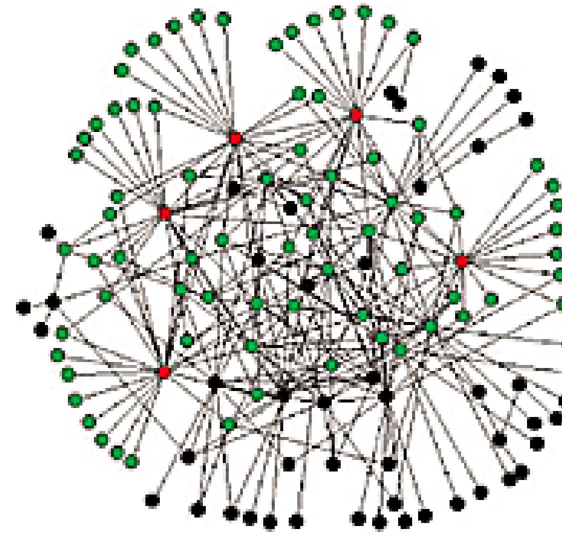
Examples of Socials & Complex Networks

- Network of friendship
- Future Internet – Internet of Things
- E-mail exchanges among employees of a large company
- Network among financial institutions
- Links among web pages revealing (more or less) dense communities and prominent sites
- International trade (country level – countries with powerful positions that derive economic growth)

Complex Networks (CNs) – Examples

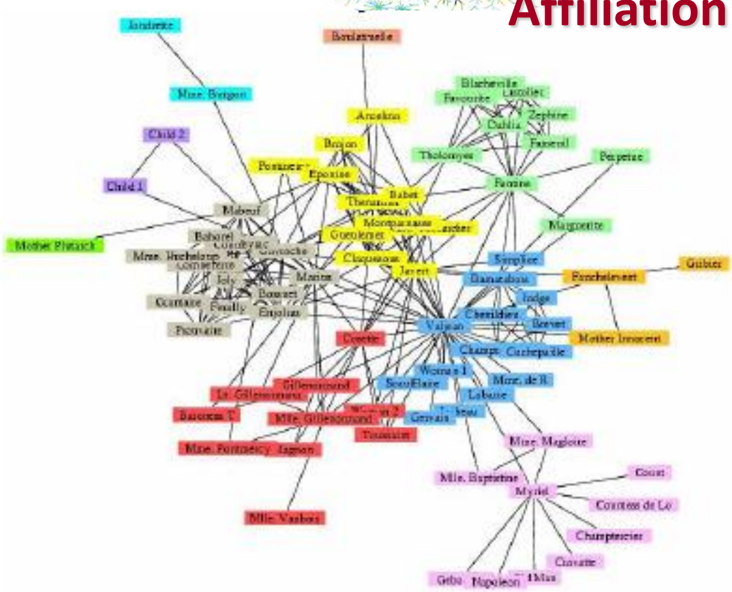
WWW instance (Feb. 06, 2003)~100,000 nodes

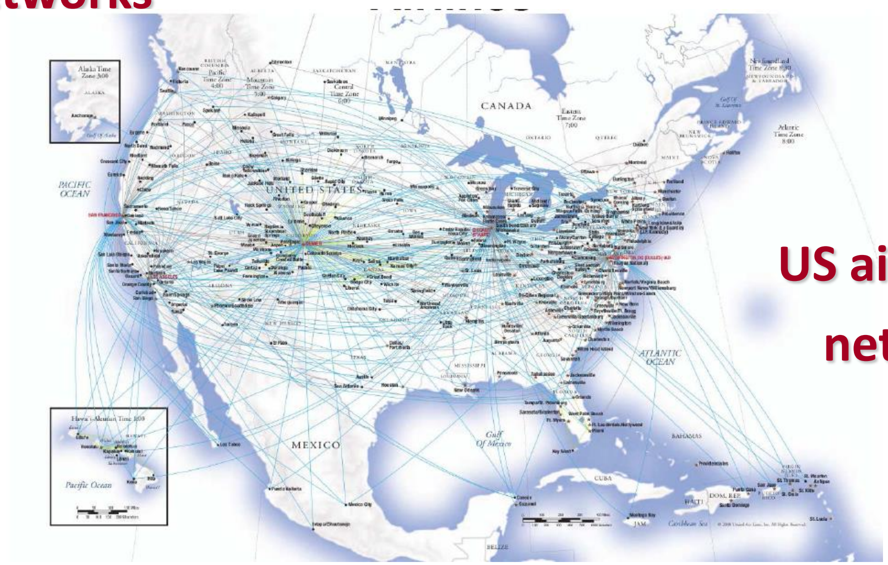




Friendship network

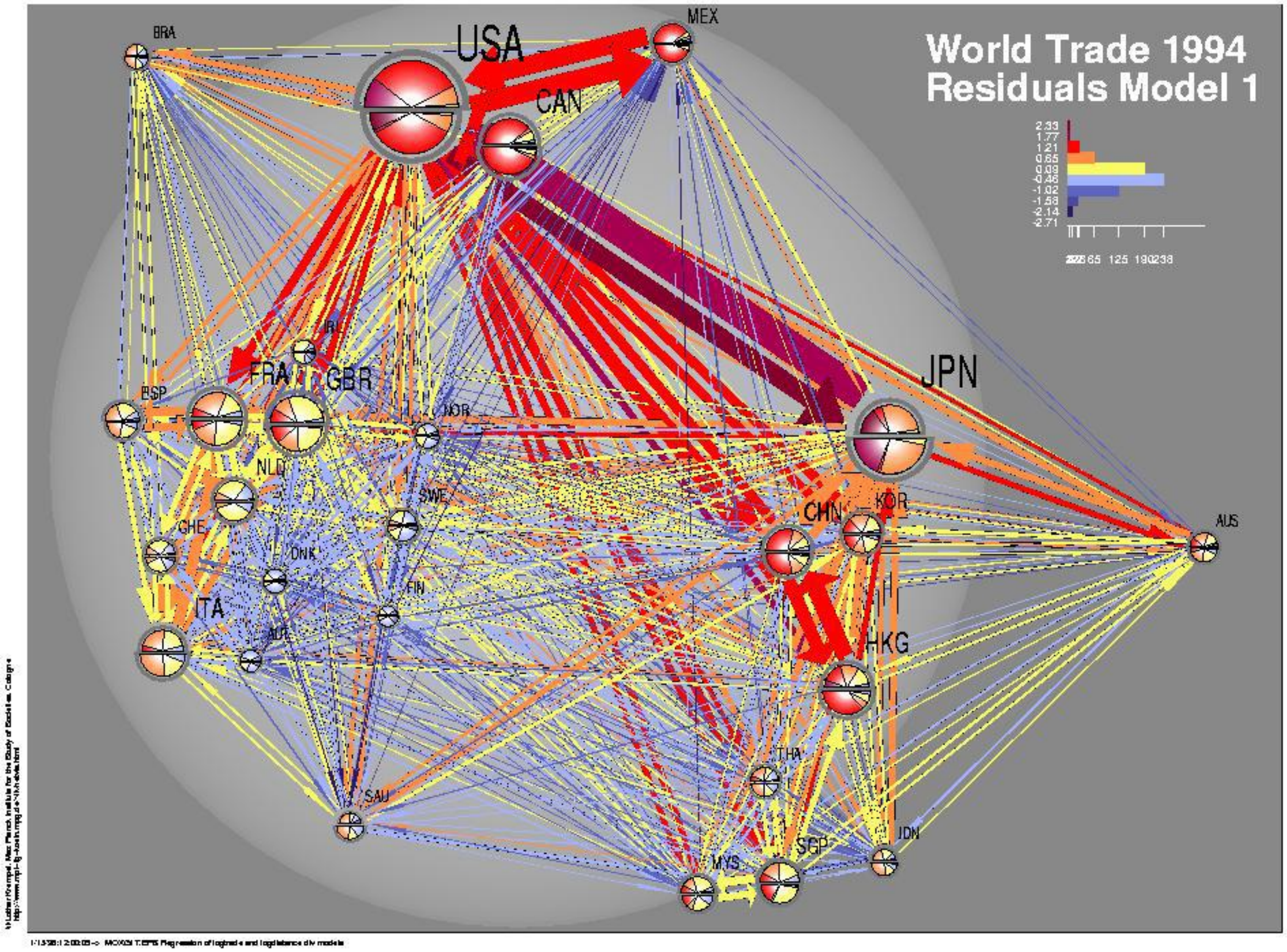
Affiliation networks



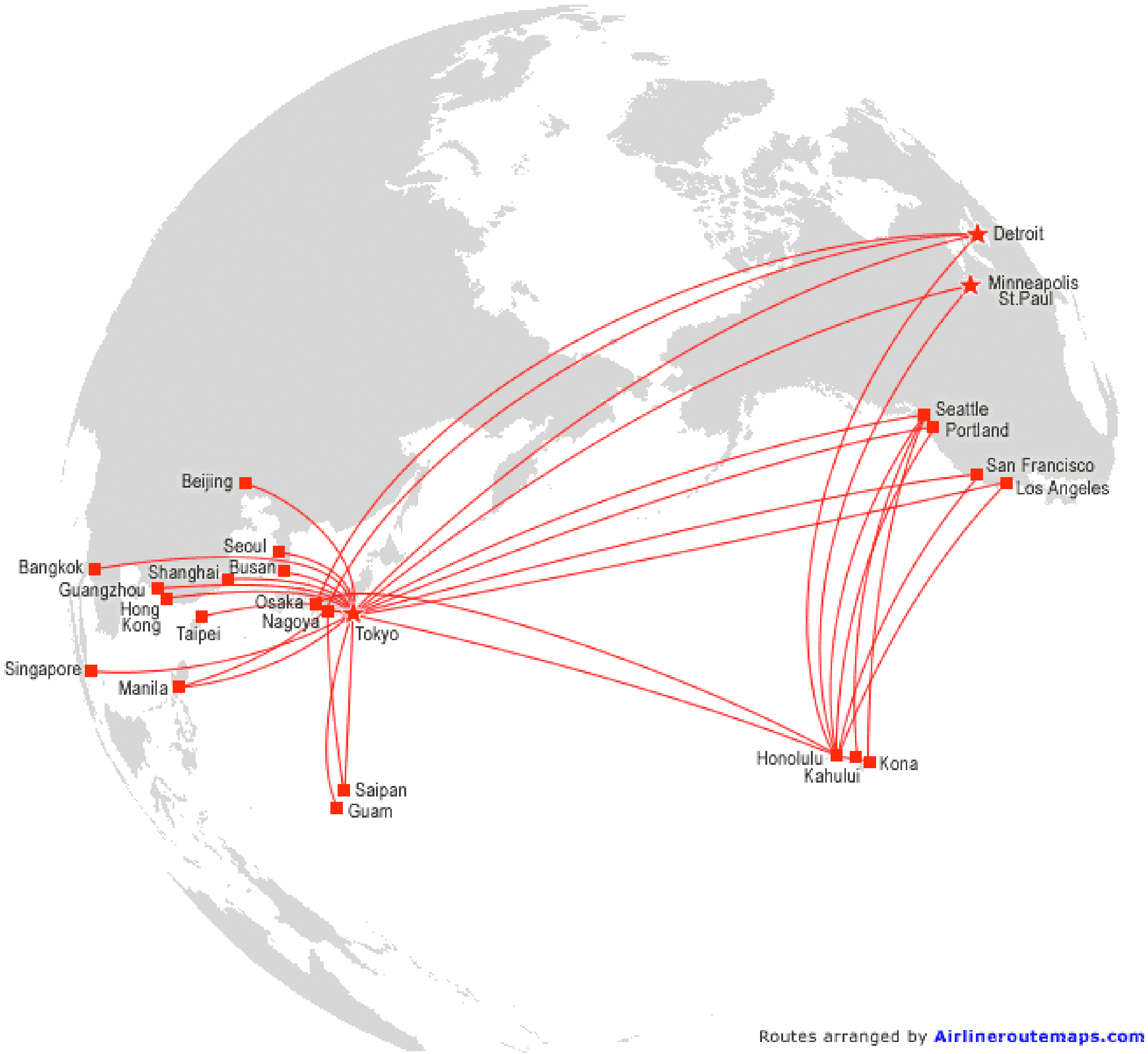


US air-traffic network

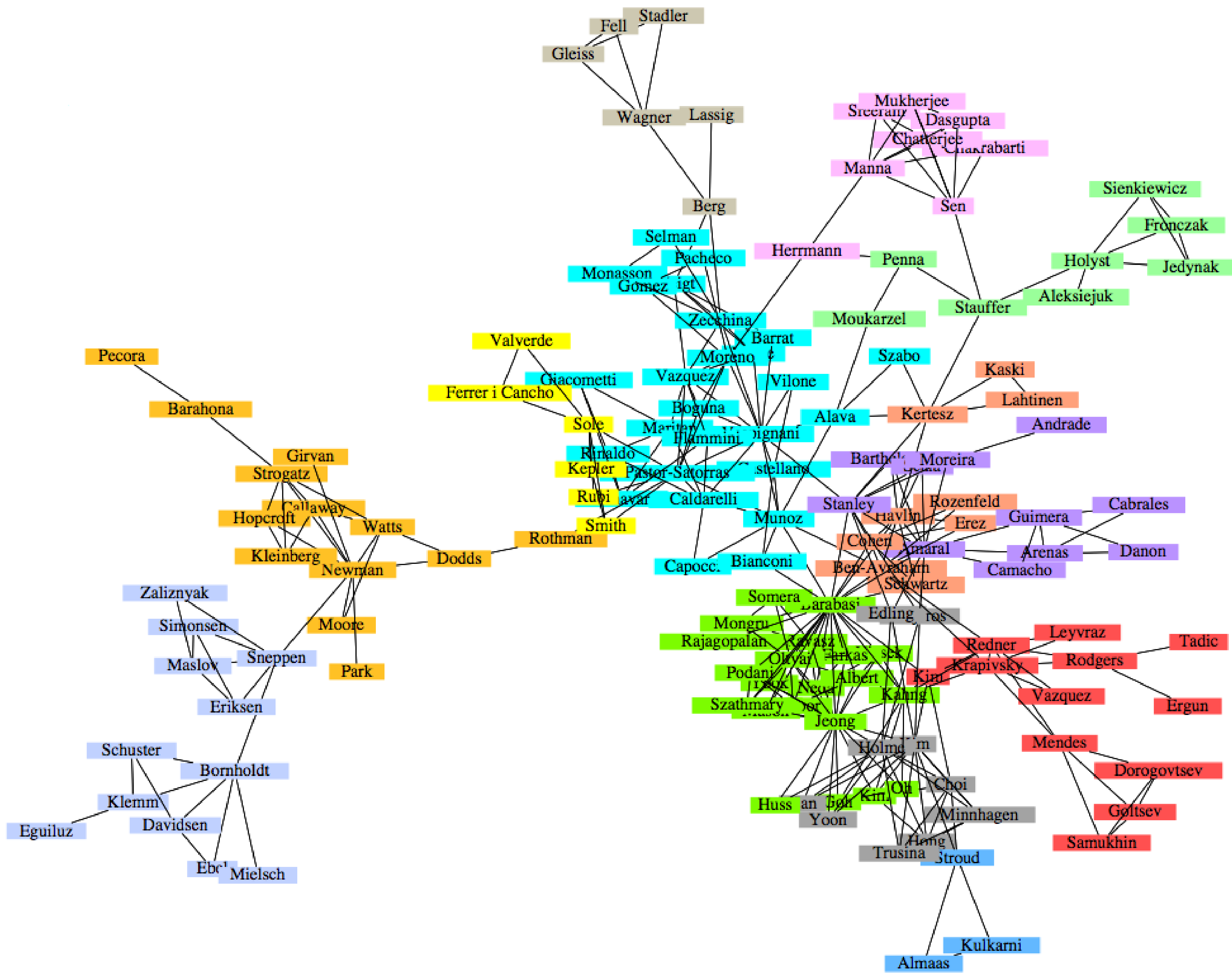
International Trade Network formation



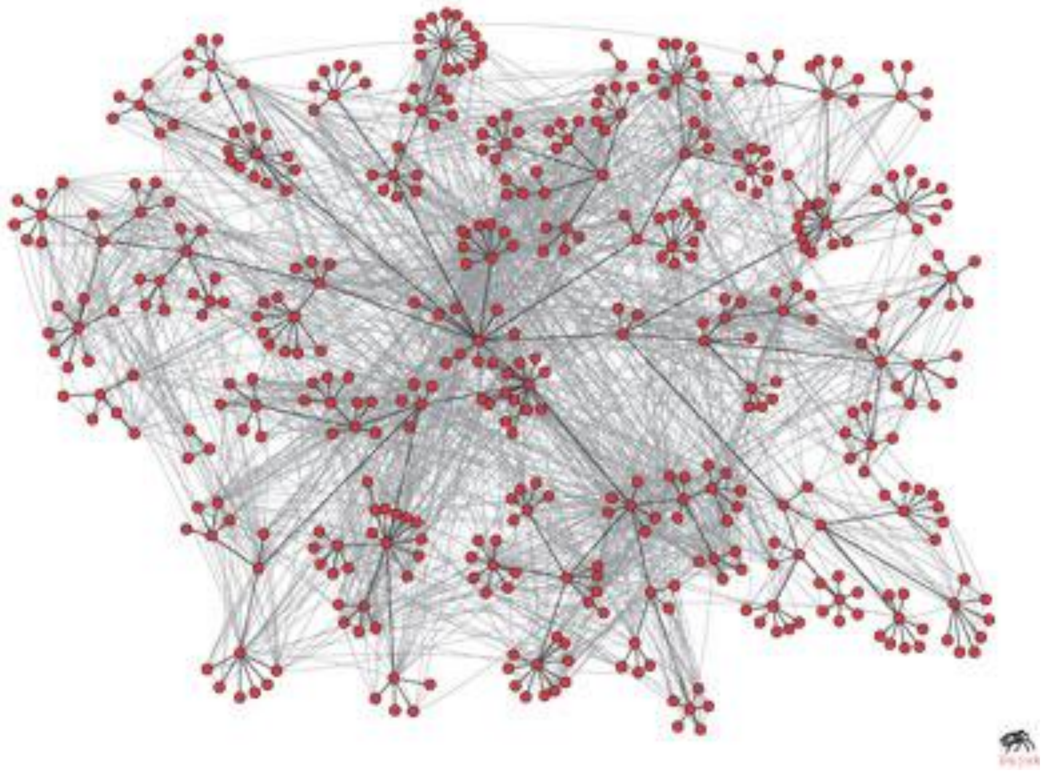
Airline Routes



Co-authorship Network

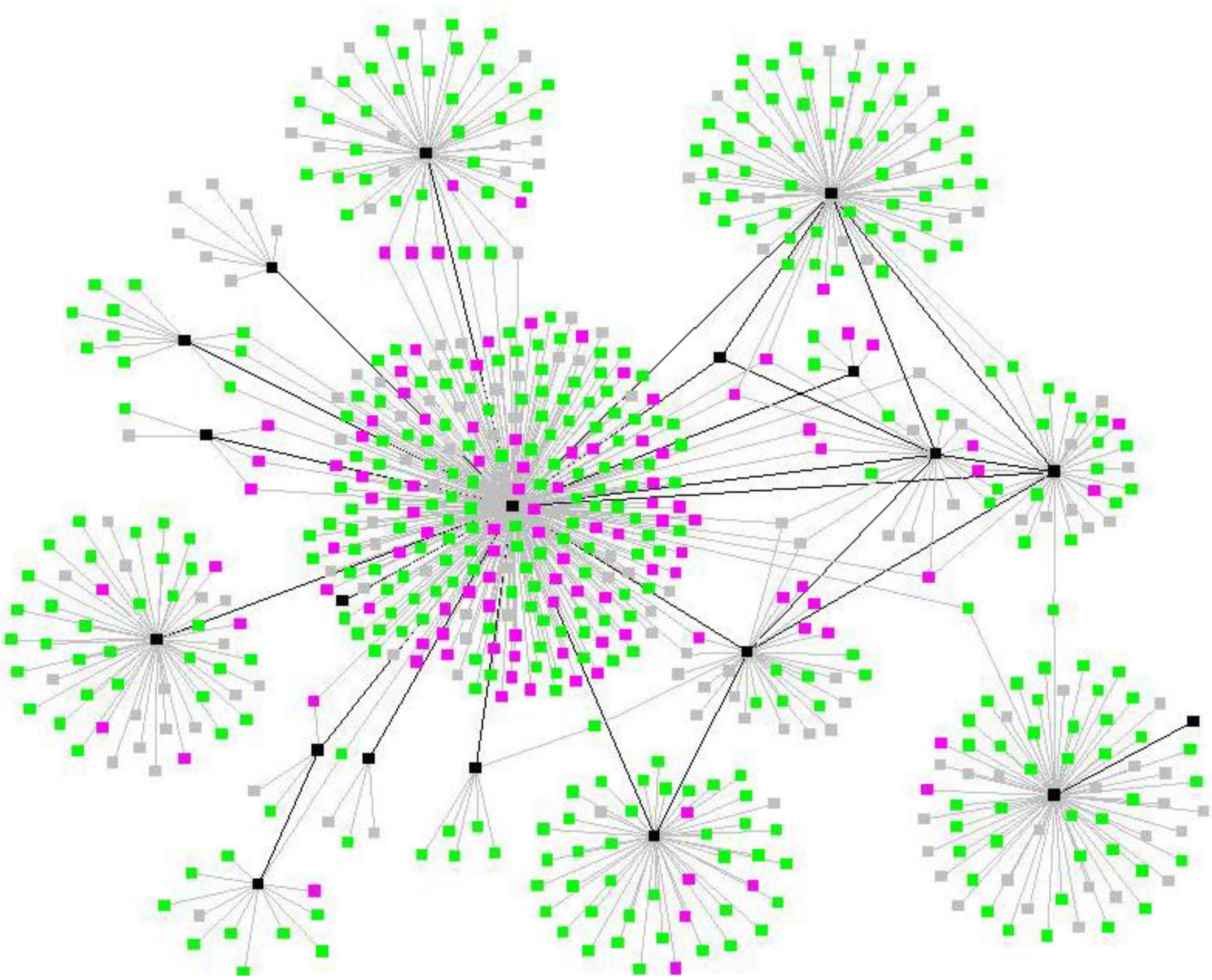


The pattern of e-mail communication among 436 employees



•Image from http://wwwpersonal.umich.edu/ ladamic/img/hplabsemailhierarchy.jpg

Spread of epidemic disease

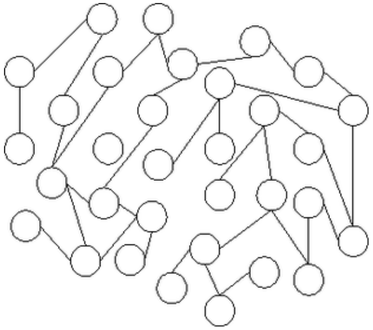


Complex interactions in highly connected systems

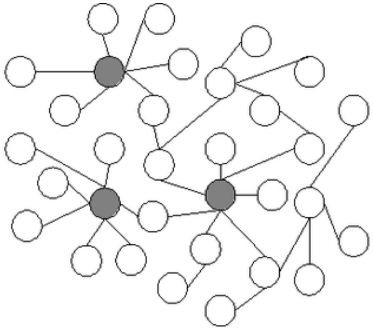
- Understanding highly connected networks and systems requires:
  - Network structure
  - Strategic behavior of actors
  - Feedback effects they produce across large populations

- **Structure:**
  - more or less densely interconnected
  - central core nodes containing most links
  - natural splits in multiple tightly-linked regions
  - Participants can be more central or more peripheral
  - **This requires a graph-theoretical framework to describe network formation and interconnectivity.**

Examples of Complex Networks (small world, scale free, random)

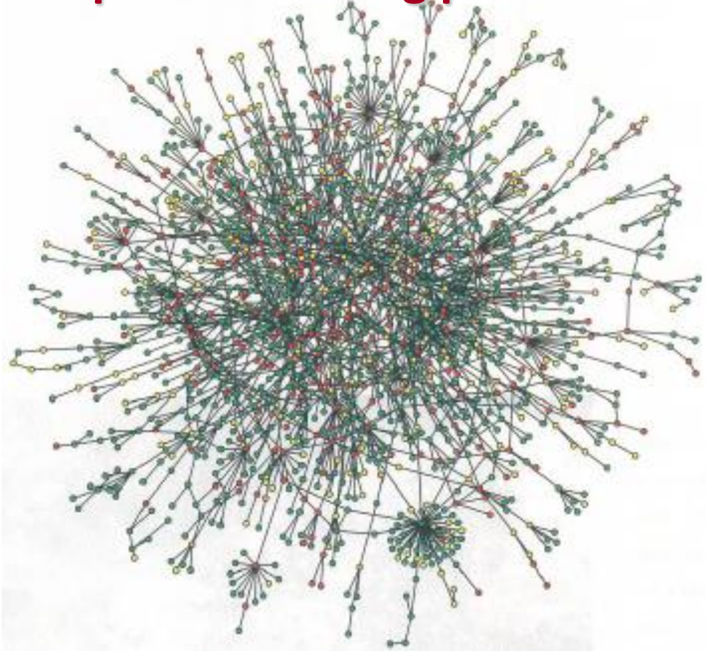


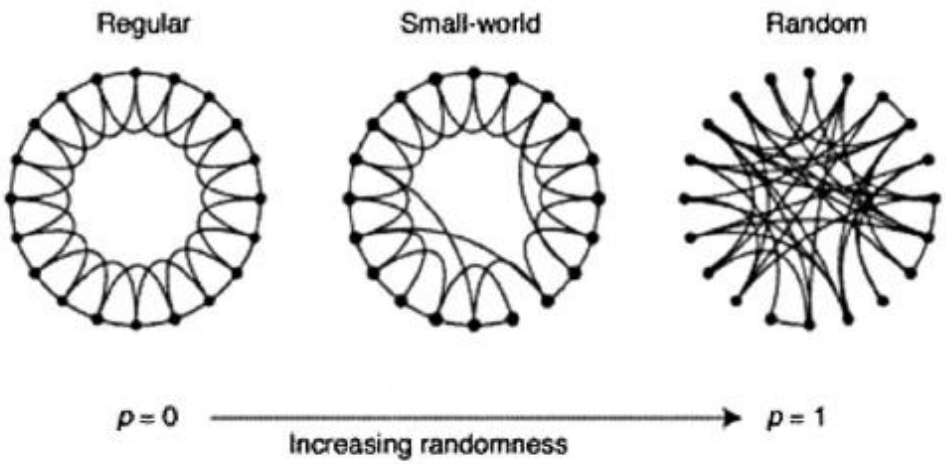
(a) Random network



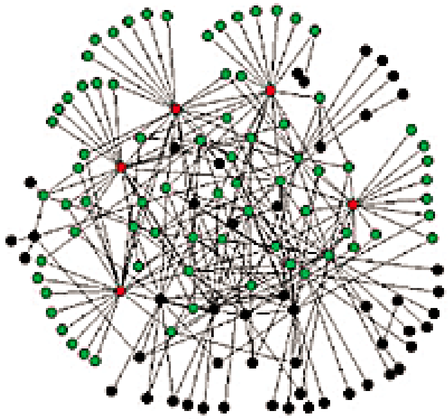
(b) Scale-free network

Map of interacting proteins





Affiliation network



Physical versus Logical Networks

- In many networks the links between nodes are physical:
  - fiber-optic cable connecting computers
  - highways connecting cities
  - wires connecting components in a circuit
- In other networks the links are logical or virtual:
  - friendships on Facebook
  - business relationships among companies
  - protein-protein interactions in a cell
- We make no distinction between physical and logical networks. Our goal is to model any network at an abstract level (graph) that is then amenable to common mathematical concepts and methods of analysis.

Behavior and Dynamics

- Connectedness at level of structure and at level of behavior:
  - each individual’s behavior have impact (explicit/implicit) for the outcome of everyone in the system.
  - This requires reasoning and modeling of behavior and interaction in network context
  - Example: changes in a product, web site, etc. can seem like a good idea when evaluated on the assumption that everything else remains static. But in reality changes can create incentives that shift behavior across the network (initially and usually unintended) .
  - **This requires study of dynamics of aggregate behavior.**

Network dynamics – Population effects

- Observing population over time we see patterns by which, ideas, beliefs, technology, products, social interactions, emerge and evolve
  - The way in which new information/products spread through a population depends in large part on the fact that people influence each other's behavior. If more and more people do something you generally become more likely to do it.
- **In general combine personal opinion and others' behavior (other's behavior conveys information**).
  - Example: in case of a web site (e.g. YouTube) seeing a lot of people using it can suggest that these people know something about it (similarly seeing a restaurant crowded that those people think highly of it).
  - This may lead to cascade effects (abandon your own private information and follow the crowd).
  - Possibly there is a benefit for aligning your behavior with others, regardless of whether they are making best choice

Example: Social networking – media sharing

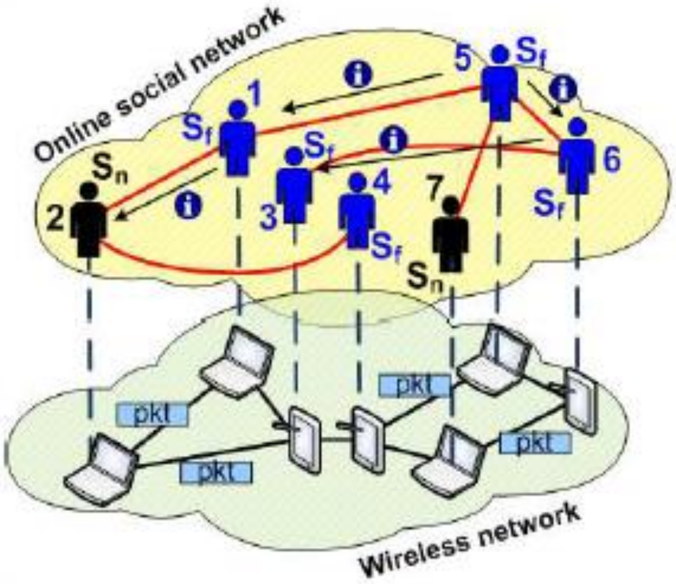
- If the value of such sites is the potential to interact with others, to have access to a wide range of contents, or have a large audience for content you post, then these sites become more popular to join (regardless if they have better features than its competitors).
- Result: Network effects amplify the success of technologies and products that are already doing well (positive externality). Popularity is governed by “rich-get-richer” feedback process
- Outcome: a social network is divided between a small number of prominent items and a long tail of more obscure ones

Network dynamics – Structural effects

- Process of influencing people behavior can gain insight or be affected by network structure
- Align your behavior with the behavior of your neighbors (e.g. in social networks) rather than the population as a whole
- Outcome: a new behavior starts with a small set of initial adapters and then spreads (cascading effects).

Example of Generalized (Cyber-Physical) Systems

- **Generalized Networks (GNs)**: cyber-physical systems of overlay networks
  - Interconnected devices, users, services
  - Nodes connected directly $\rightarrow$ physical layer; nodes socially connected $\rightarrow$ cyber layer
- Focus on information wireless GNs (WGNs)
- Information diffusion over WGNs
  - Useful – malicious info
  - Temporal-topical variations impact information diffusion dynamics
    - E.g., seasonal interest, new hobbies, breaking news, etc.
- Epidemic model for the study of Information Diffusion Dynamics

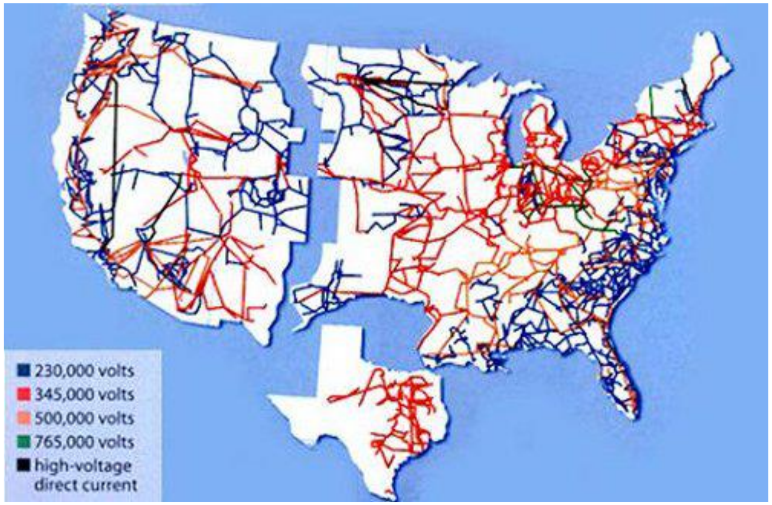


Example: The Power Grid – Dynamic Effects

The Electric Power Grid is an example of a complex network where dynamic effects are important to consider.

The problem of cascading failure in power networks is caused when failure of one node causes successive overload of other nodes, leading to multiple failures.

Cascading failure can also occur in many other systems, including financial networks and computer networks.

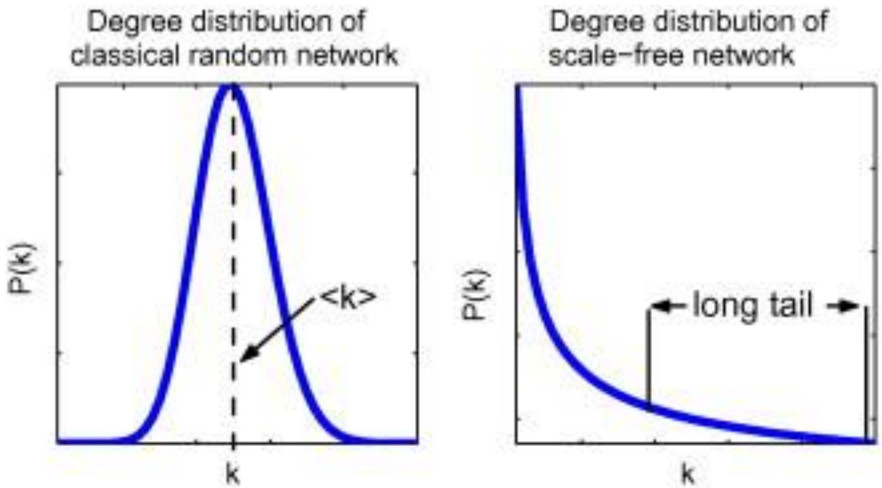


Power Law Distributions & 80/20 Rule

Some commonly held beliefs in business, for example, are

- 80% of a company's profits come from 20% of its customers
- 80% of a company's complaints come from 20% of its customers
- 80% of a company's sales come from 20% of its products
- 80% of a company's sales are made by 20% of its sales staff
- Microsoft noted that by fixing the top 20% most reported bugs, 80% of the errors and crashes would be eliminated
- In software engineering, 20 percent of the code has 80 percent of the errors.

The 80-20 rule, or Pareto principle is a common rule of thumb that, in many endeavors, roughly 80% of the effects come from 20% of the causes.



In a Power Law Distribution the probability that a node u has degree k is:

$$P(k) \sim k^{-\gamma}$$

Network Science

- Emerging {common + generic} problems in CNs
- A complex network theory is required
- Mathematical models for diverse networks

  Study of similar {statistical, social, structural} properties & behaviors

- Working examples
  - Spreading of a disease in a social network
  - Malware diffusion over a telecommunication network
  - Information dissemination in an affiliation network
  - Failure propagation in a large power network
  - Financial crisis spreading in global markets

They all describe the same problem

$\Longrightarrow$ Network Science

Network Science is the organized knowledge of networks based on their study and by using a scientific approach

Network Science – Challenges

- 'umbrella' areas of challenges
  - Complexity
  - Wide range of interacting scales (time, space, size, …)
  - Network-to-network interactions
  - Microscopic and Macroscopic analysis
- More specific challenging problem areas of interest
  - Network formation and modeling
  - Dynamics, spatial location, and information propagation in networks
  - Modeling and analysis of very large networks
  - Design and synthesis of networks
  - Abstracting common concepts across fields
  - Better experiments and measurements of network structure
  - Robustness and security of networks
  - Increasing the level of rigor and mathematical structure
  - Study network dynamics (structural and population effects)