Lecture 1

Network Science

Overview & Tools

Overview & Tools

Agenda

- General Information
- Instructors
- Class Schedule
- Poll (via Google Form)
- Textbooks
- Introduction to Networks
- Tools

General Information

- Computer Science (6 and 9 credits)
- SDS (6 credits)
- PhD "Modeling and Data Science" (exam: a seminar or a project + oral exam.)
- Moodle:

https://informatica.i-learn.unito.it/course/view.php?id=1700

Instructors

Giancarlo RUFFO

giancarlo.ruffo@unito.it

office hours:

Wed. 1:30pm - 2:30pm

Alessandro FLAMMINI

INDIANA UNIVERSITY

Class Schedule

	Monday	Tuesday	Wednesday	Thursday	Friday
25/02-01/03	25-Feb	26-Feb	27-Feb aula f 11-13 NetSci Ruffo	28-Feb aula f 11-13 DataViz Ruffo	1-Mar
05/03-08/03	4-Mar aula e 11-13 NetSci Ruffo	5-Mar	6-Mar aula f 11-13 NetSci Ruffo	7-Mar aula f 11-13 DataViz Ruffo	8-Mar
11/03-15/03	11-Mar aula e 11-13 NetSci Ruffo	12-Mar	13-Mar aula f 11-13 NetSci Ruffo	14-Mar aula f 11-13 DataViz Ruffo	15-Mar
18/03-16/03	18-Mar aula e 11-13 NetSci Ruffo	19-Mar	20-Mar aula f 11-13 NetSci Ruffo	21-Mar aula f 11-13 DataViz Ruffo	22-Mar
25/03-29/03	26-Mar aula e 11-13 NetSci Ruffo	27-Mar	28-Mar aula f 11-13 NetSci Ruffo	29-Mar aula f 11-13 DataViz Ruffo	30-Mar
01/04-05/04	1-Apr aula e 11-13 NetSci Ruffo	2-Apr	3-Apr aula f 11-13 NetSci Ruffo	4-Apr aula f 11-13 DataViz Ruffo	5-Apr
09/04-13/04	9-Apr aula e 11-13 NetSci Ruffo	10-Apr	11-Apr aula f 11-13 NetSci Ruffo	12-Apr aula f 11-13 DataViz Ruffo	13-Apr
15/04-19/04	15-Apr aula e 11-13 NetSci Ruffo	16-Apr	17-Apr aula f 11-13 DataViz Ruffo	18-Apr aula f 11-13 DataViz Ruffo	19-Apr
22/04-26/04	22-Apr	23-Apr	24-Apr aula f 11-13 DataViz Ruffo	25-Apr	26-Apr
29/04-03/05	29-Apr aula e 11-13 NetSci Ruffo	30-Apr	1-May	2-May aula f 11-13 DataViz Ruffo	3-May
06/05-10/05	6-May aula e 11-13 NetSci Ruffo	7-May	8-May aula f 11-13 NetSci Flammini	9-May aula f 11-13 NetSci Flammini	10-May
13/05-17/05	13-May aula e 11-13 NetSci Flammini	14-May	15-May aula f 11-13 NetSci Flammini	16-May aula f 11-13 NetSci Flammini	17-May
20/05-24/05	20-May aula e 11-13 NetSci Flammini	21-May	22-May aula f 11-13 NetSci Flammini	23-May aula f 11-13 NetSci Flammini	24-May
27/05-30/05	27-May	28-May	29-May	30-May aula f 11-13 DataViz Ruffo	31-May

6 credits: NetSci classes

9 credits: NetSci + Data Viz classes

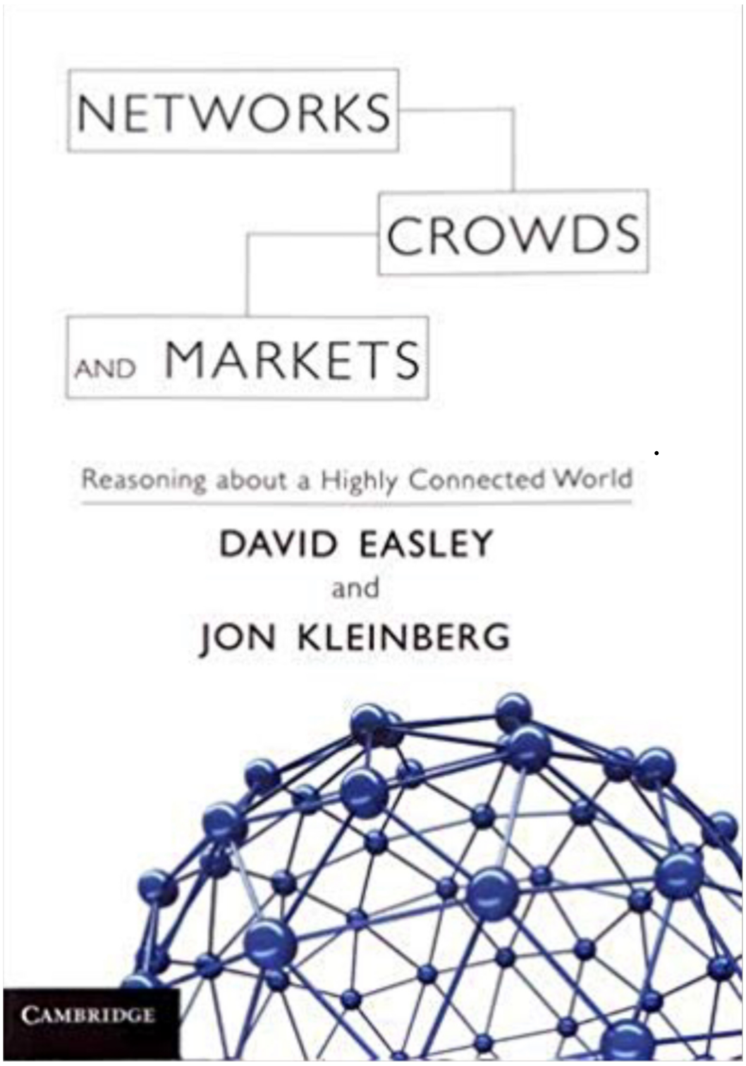
Poll

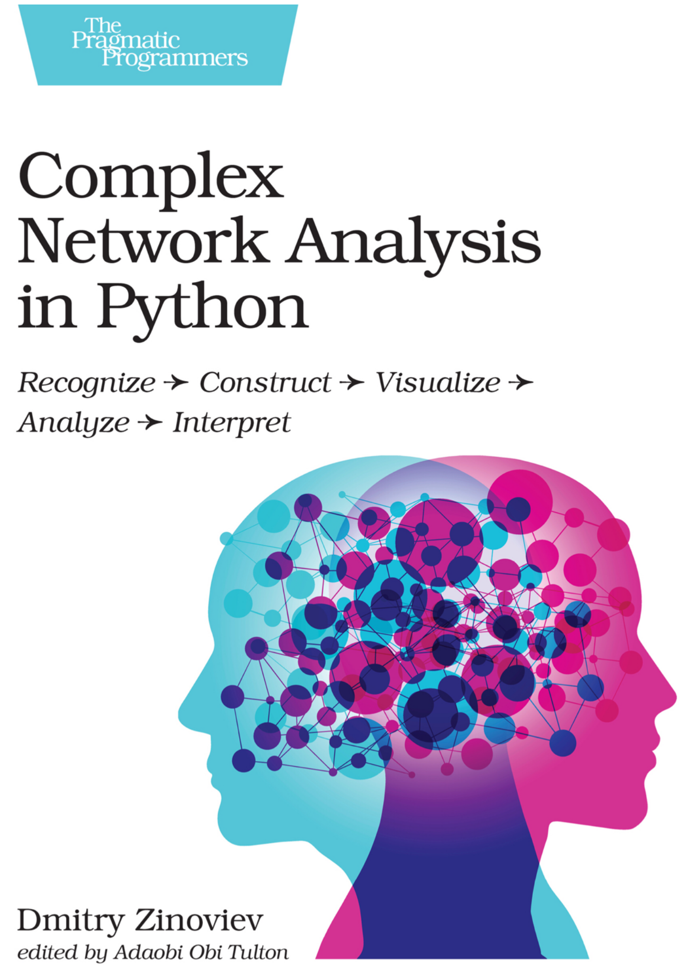
https://goo.gl/forms/rXXZBKTDq4NUm5m22

fill out the form
because I need
to know:

- general statistics
- your programming skills

Textbooks





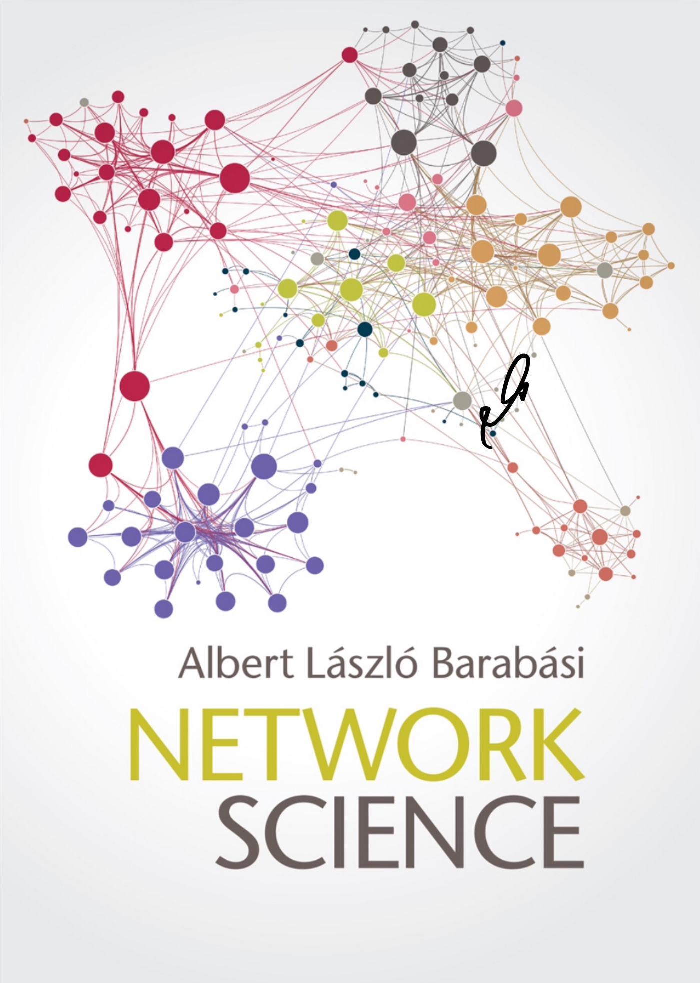
ch. 1 – 6

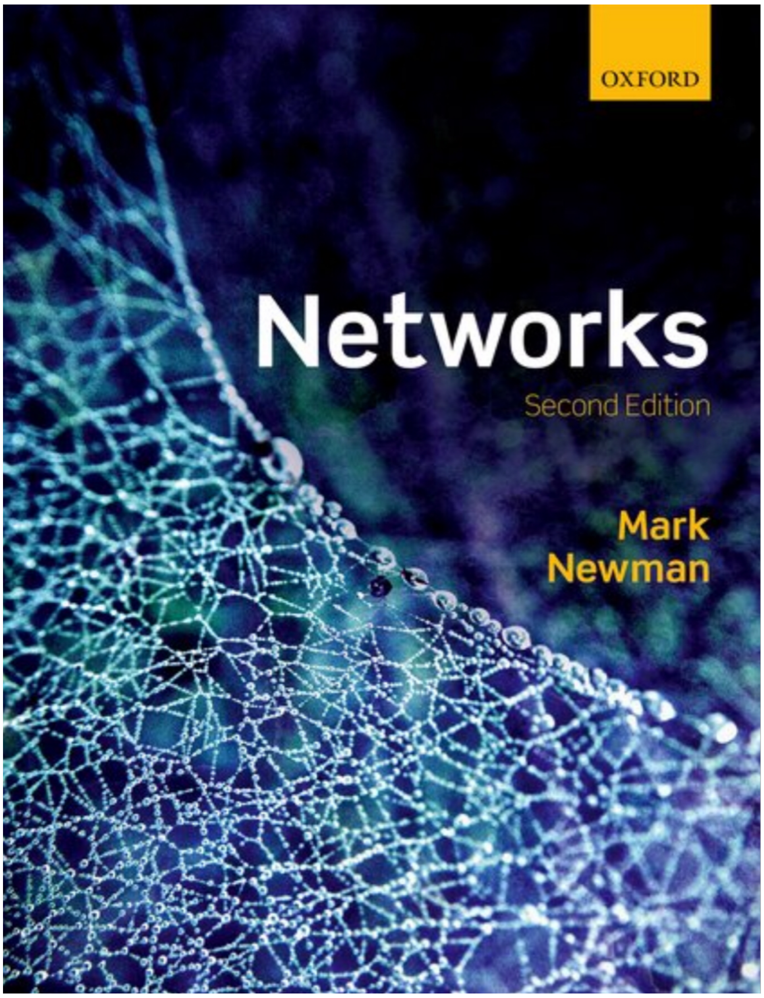
ch. 13 – 14

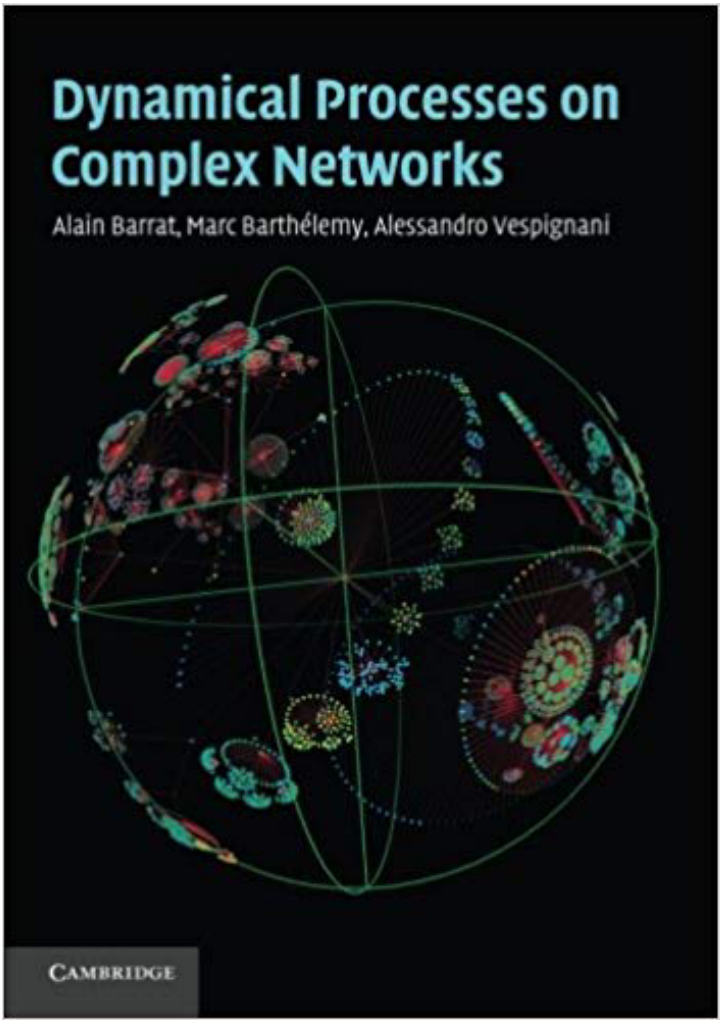
ch 16 – 21

on-line

practical


Albert László Barabási
NETWORK
SCIENCE


OXFORD
Networks
Second Edition
Mark
Newman


Dynamical Processes on
Complex Networks
Alain Barrat, Marc Barthélemy, Alessandro Vespignani
CAMBRIDGE

+ papers

Exam

- essay: solution of an exercise
- "real data" analysis: report with interpretation of results

---

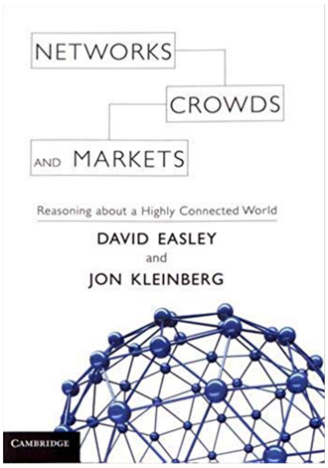
- oral examination

Data Visualization +3 credits

- project

Introduction to Networks

Chapter 1



Networks are apparently "everywhere"

- Social Networks
  - actors
  - social ties
- Information Systems
  - book, web page
  - citation
- techn. and economic systems
  - can structural crisis be predicted?

Complex Systems

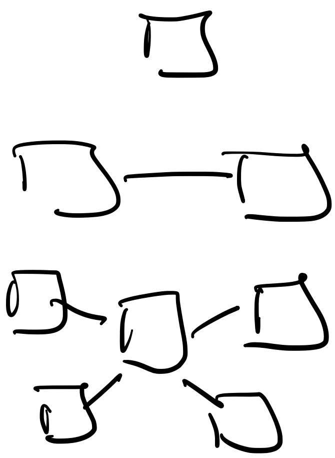
https://youtu.be/5v5eBf2KwF8

Complex ≠ Complicated

- individuals
- connected to each other

---

local level



---

global level

phenomena

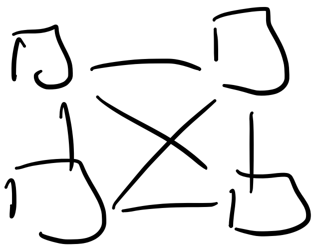
Synchronize.
Congestion
Epidemics
...

Networks aspects

"trivial" representation of a complex system

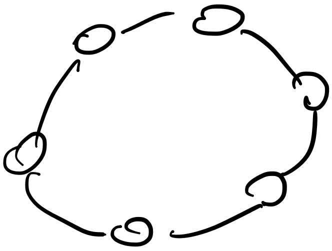
Simple Networks

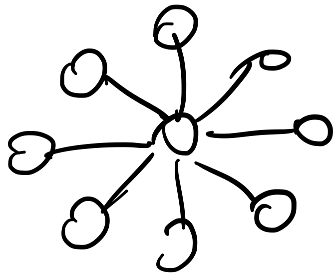
Exemples



grid

ring





star

here few characteristics describe the network

We need a framework and a language to describe and understand complex Networks.

Zachary's Karate Club

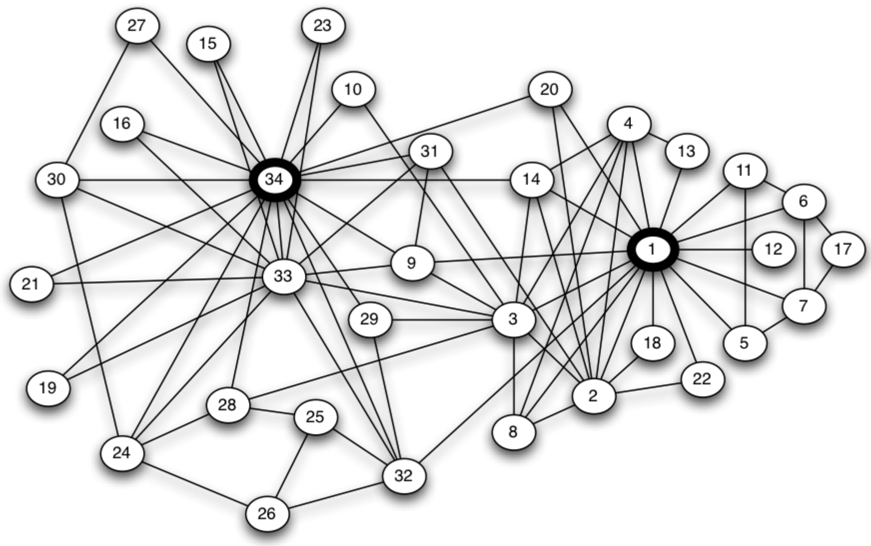


Figure 1.1: The social network of friendships within a 34-person karate club [421].

Apparently no "simple" description is possible.

However there is some "regularities"

Some individuals are more "central" than others

Does it resemble some kind of randomness?

Communication networks

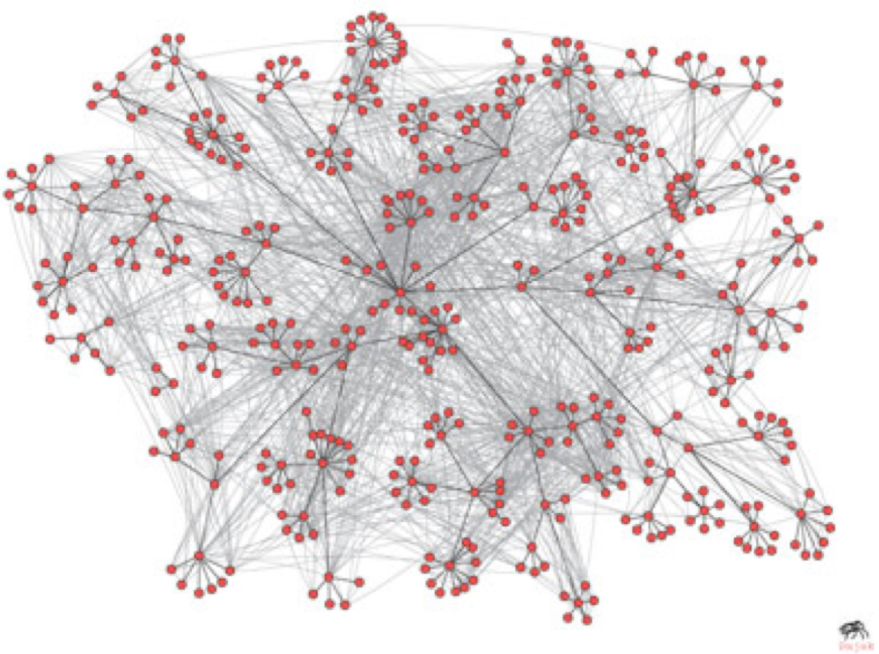


Figure 1.2: Social networks based on communication and interaction can also be constructed from the traces left by on-line data. In this case, the pattern of e-mail communication among 436 employees of Hewlett Packard Research Lab is superimposed on the official organizational hierarchy [6]. (Image from http://www-personal.umich.edu/ ladamic/img/hplabsemailhierarchy.jpg)

Some "order" emerges from chaos.

hierarchy?

We need a language to better describe these regularities

Financial Networks

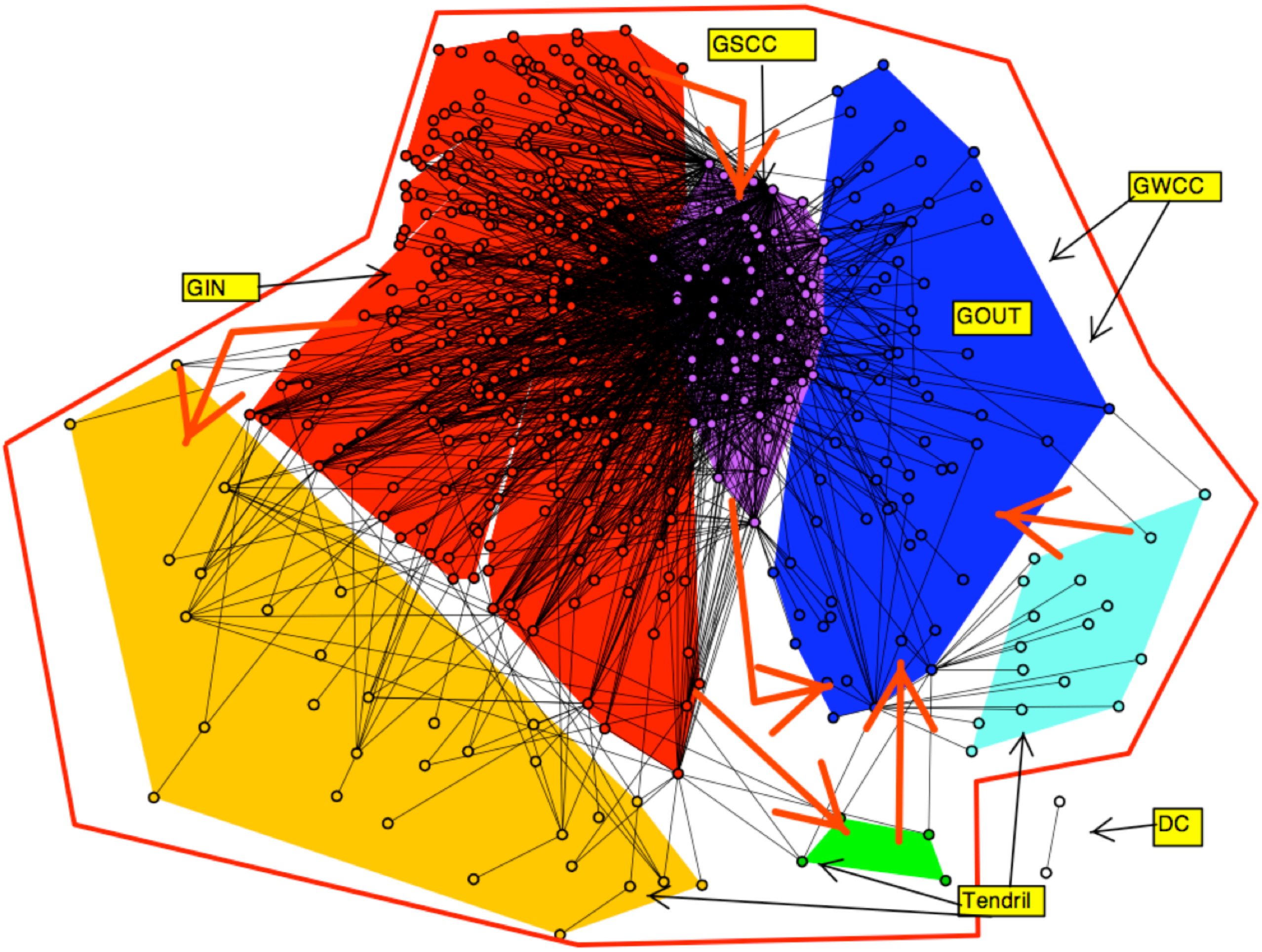


Figure 1.3: The network of loans among financial institutions can be used to analyze the roles that different participants play in the financial system, and how the interactions among these roles affect the health of individual participants and the system as a whole The network here is annotated in a way that reveals its dense core, according to a scheme we will encounter in Chapter 13. (Image from Bech and Atalay [50].)

the emergence of a core vs periphery structure.

inverse homophily

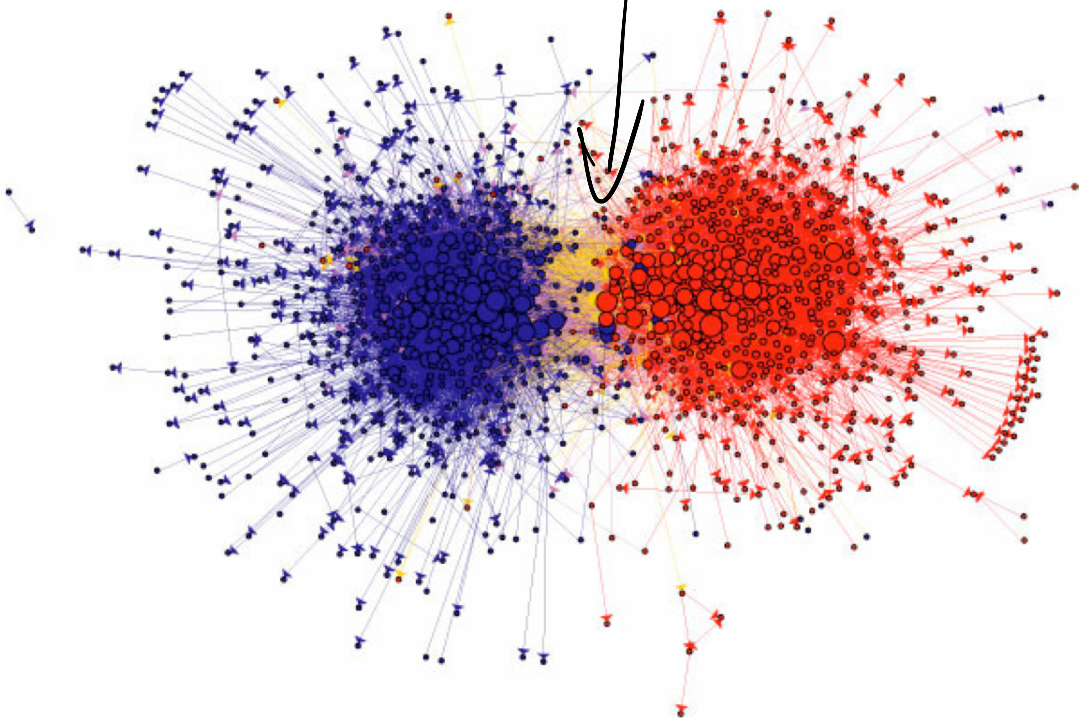


Figure 1.4: The links among Web pages can reveal densely-knit communities and prominent sites. In this case, the network structure of political blogs prior to the 2004 U.S. Presidential election reveals two natural and well-separated clusters [5]. (Image from http://www-personal.umich.edu/ ladamic/img/politicalblogs.jpg)

communities   echo chambers

clusters   partitions

homophily → tendency of individuals to link with similar ones

WARNING: no trivial linear relationships

But INTERPLAY

Visualization ?

visualizing a network suggests some inherent complexity

> this tells a lot about the structure,
> but almost nothing about the reasons

it is difficult to summarize succintly the whole network

- core vs periphery structure
- hierarchical structure
- tightly-linked regions
- ...

we need a language

the structure is only a starting point

Behavior and Dynamics

Connectedness — structure

— behavior

each individual's actions have consequences for the outcomes of everyone in the system.

We need a framework for reasoning about behavior and interaction in networks

We need to take into account a rational behavior that sometimes lead to

strategies

Actions are not evaluated at individual level (in isolation), but at a network level.

When a large group of people is tightly interconnected, they often respond in complex ways that can be observable only at the population level:

rich get richer

Winner takes it all

...

viral ideas

At individual level:

Sometimes an information goes viral or one person becomes very popular, some other times this does not happen!

Why? No prediction here yet.

Recap (1)

1. network structure
2. strategic behavior
3. feedback effects

at a population level

language and frameworks from

computer science, physics, mathematics, statistics, operational research, sociology, economics ...

From Data to Networks

We have
massive network
datasets

---

data-driven models
allow you to
make predictions

BUT

sometimes these predictions
cannot be
generalized:
networks can be "different"

Topics

Graph theory

Game theory

Information Networks

Network Dynamics

Graph theory

paths / distances

clustering coefficient

degree / distribution of degrees

centralities

groups / clusters / partitions
communities

structural balance

homophily / spatial segregation

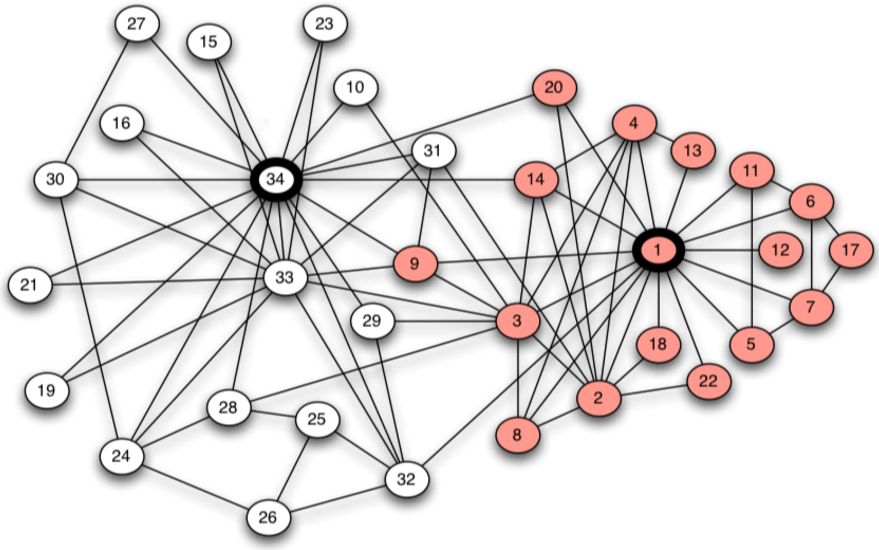


Figure 1.7: From the social network of friendships in the karate club from Figure 1.1, we can find clues to the latent schism that eventually split the group into two separate clubs (indicated by the two different shadings of individuals in the picture).

individuals vs groups

we can have different roles ...

Game theory

framework where our decisions depend on others' decisions

strategy → pay-off

Ex. Transportation Networks
choice of a route can bring to congestion

Braess' Paradox

Equilibrium : a state that is self-reinforcing: no individuals have an incentive to unilaterally change their strategy, even if those individuals know how others will behave.

Information Networks

the Web

search engines

rank

web pages authors

react to maximize

Interplay

If you define a measure (e.g., Page Rank), the application of such measure may change the system.

Network Dynamics: Population Effects

In large populations we have the emergence of new ideas, beliefs, opinions, innovations, technologies, products, and social conventions.

When and how they establish as <u>social practices</u>?

individuals can <u>influence</u> or being influenced by others

At the surface: humans have the tendency to <u>conform</u>

WHY?

Network Dynamics: Structural Effects

population vs local behaviors

Individuals may have incentive to adopt the behavior of their neighbors:

Cascading effects

Social Contagion

epidemics

the structure of the network has a role

spreaders

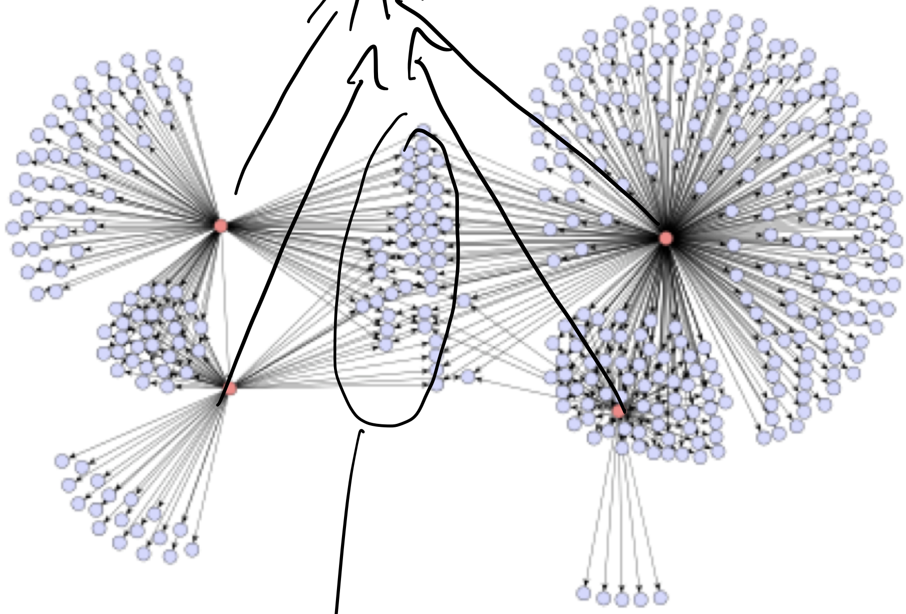


Figure 1.11: When people are influenced by the behaviors their neighbors in the network, the adoption of a new product or innovation can cascade through the network structure. Here, e-mail recommendations for a Japanese graphic novel spread in a kind of informational or social contagion. (Image from Leskovec et al. [271].)

Social contagion

a closed community can stop diffusion

epidemic

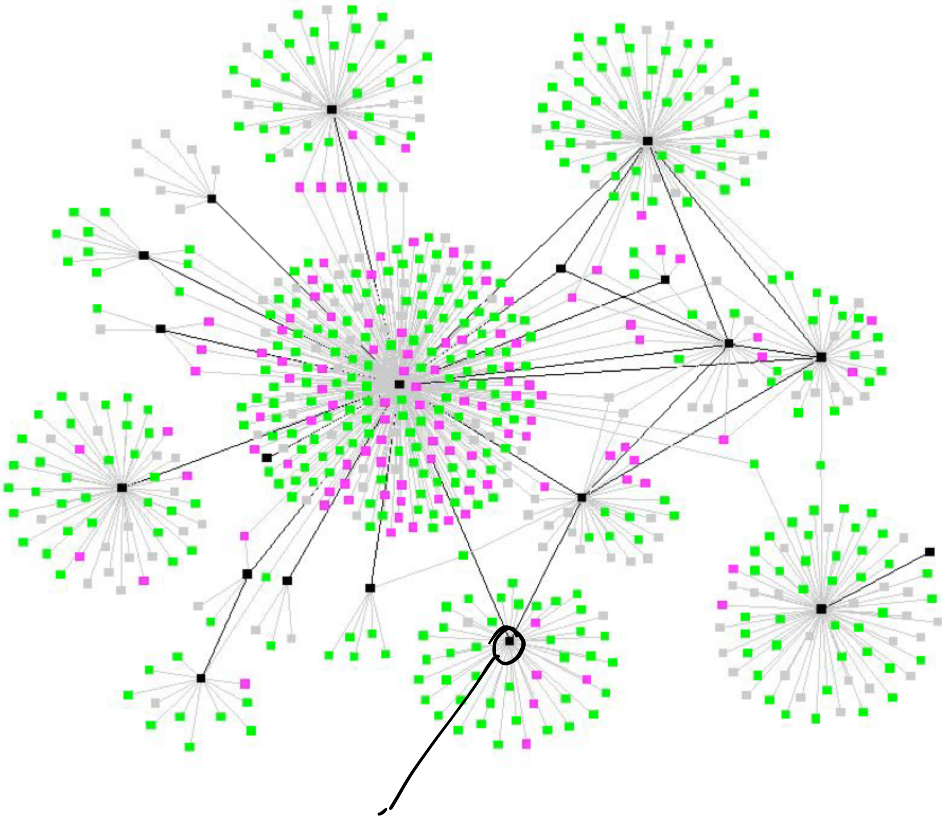


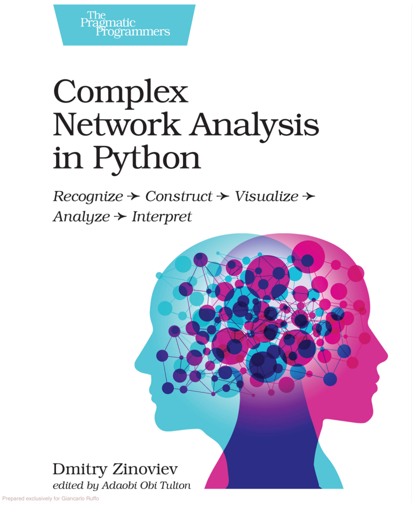
Figure 1.12: The spread of an epidemic disease (such as the tuberculosis outbreak shown here) is another form of cascading behavior in a network. The similarities and contrasts between biological and social contagion lead to interesting research questions. (Image from Andre et al. [16].)

Cascading behavior here too!

Recap (2)

Network-level dynamics are similar and insights from the study of biological epidemics are also useful in thinking about the process by which things spread in networks

Tools



Ch. 1 & 2

Complex Network

Analysis

CNA

Jupyter
(notebook)

ANACONDA Distribution

Package	Used version	Package	Used version
python	3.6.5	networkx	2.1
matplotlib	2.0.2	community	0.10
nltk	3.2.5	numpy	1.13.3
pandas	0.22.0	pygraphviz	1.3
wikipedia	1.4.0	scipy	1.0.1
toposort	1.5		

GEPHI

a network visualization
Tool developed in Java

Examples and real datasets

Technological networks	Communication systems; transportation; the Internet; electric grid; water mains
Biological/ecological networks	Food webs; gene/protein interactions; neural system; disease epidemics
Economic networks	Financial transactions; corporate partnerships; international trade; market basket analysis
Social networks	Families and friends; email/SMS exchanges; professional groups
Cultural networks	Language families; semantic networks; literature, art, history, religion networks (emerging fields)

Comparison

	graph-tool	iGraph	NetworkX	NetworKit
Implementation language	C/C++	C/C++	Python	C/C++
Language bindings	Python	C, Python, R	Python	C++, Python
Installation effort	Hard	Medium	Easy	Medium
OpenMP support	Yes	No	No	Yes
Relative slowdown[5]	1	1–4	40–135	N/A
Built-in community detection	Yes	Yes	No	Yes
Built-in advanced layouts	Yes	Yes	No	Yes