Network Science

Prof Alberto Paccanaro
Dr Suzana Santos (monitor)

About myself

- Background

  MSc Computer Science (Un. Milan) 

  PhD in Machine Learning (Un. Toronto) 

  Computational Biology/Bioinformatics (Yale Un.) 

- Research

  www.paccanarolab.org

- Courses at FGV

  Técnicas e Algoritmos em Ciência de Dados (UG)

  Fundamentos de Ciência de Dados (MSc)

  Ciência de Redes (PhD)

About Suzana

- BSc Computer Science, USP (2012)
- MSc Computer Science, USP (2015)
  Development of statistical tests and a software to compare gene coexpression networks
- PhD Computer Science, USP (2020)
  Consistent estimation of parameters for random graph models and application in neuroscience
- Postdoc Computational Biology (2020 – present)
  Drug repositioning for viral diseases

Email: suzana.santos@fgv.br

Office hours: Tues (11:00-12:00) and Thurs (14:00-15:00)

About the course

- **Classes:**
  - Tuesdays 14:20-16:00
  - Tuesdays 14:20-16:00

- **Office hours:**
  - Tuesdays 13:00-14:00
  - Thursdays 12:00-13:00

  Please, contact me by email 24 hours in advance.

- **Course website**: on eClass

  Here you will find the slides of the classes, exercises, solutions and any other information about the course

Objectives of the course

Provide a general overview of the main **principles** and **ideas** in Network Science

Provide a good understanding of the main **algorithms**

Provide working knowlegde of the existing **tools**, and the ability to use them on **real-world problems**.

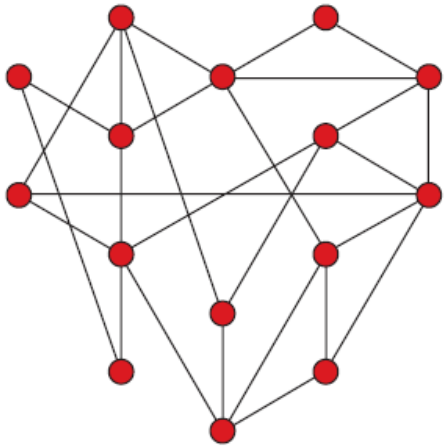
Mathematical foundations

Algos: why they work and how

Applications

OK, but what is the course about ???

The *network revolution*



Start with N nodes and connect each pair of nodes with probability *p.*
(Erdös-Rényi model, 1959)

But some facts did not add up…

Milgram's experiment

(from Wikipedia)

Then we got some data… ☺

1. *Actor*
2. *US power grid*
3. *World Wide Web*

Science, Vol. 286, 1999

Emergence of Scaling in Random Networks

Albert-László Barabási* and Réka Albert

Nature, Vol. 393, 440, 1998

Collective dynamics of 'small-world' networks

Duncan J. Watts* & Steven H. Strogatz

SIGCOMM 1999

On Power-Law Relationships of the Internet Topology

Michalis Faloutsos
U.C. Riverside
Dept. of Comp. Science
michalis@cs.ucr.edu

Petros Faloutsos
U. of Toronto
Dept. of Comp. Science
pfal@cs.toronto.edu

Christos Faloutsos *
Carnegie Mellon Univ.
Dept. of Comp. Science
christos@cs.cmu.edu

Topics

- Introduction: social networks, communication networks, the web, the internet, biological networks
- Basic concepts from Graph Theory;

- Models: the Erdos-Renyi model; small world networks; the scale free property; the Barabasi- Albert model;
- Measures: centrality measures; degree correlation; network robustness;
- Community detection;
- Dynamics over networks: information, influence, epidemic.

- Graph relational learning, node embeddings
- Graph Neural Networks
- Deep Generative Models for networks

Plan of activities

Every week:

- **“theory” class**: algorithms + theoretical and mathematical fundamentals.
- **“lab” class**: labs to implement these algorithms and apply them to real world data (Python)

One project, to be done in pairs, carried out throughout the course.

The project

Select a network of interest, map it out and analyse it.

- Carried out in pairs (professionally heterogenous, if possible).
- *You choose a dataset:* you MUST collect your own data - simply downloading a dataset already prepared for network analysis is not acceptable.

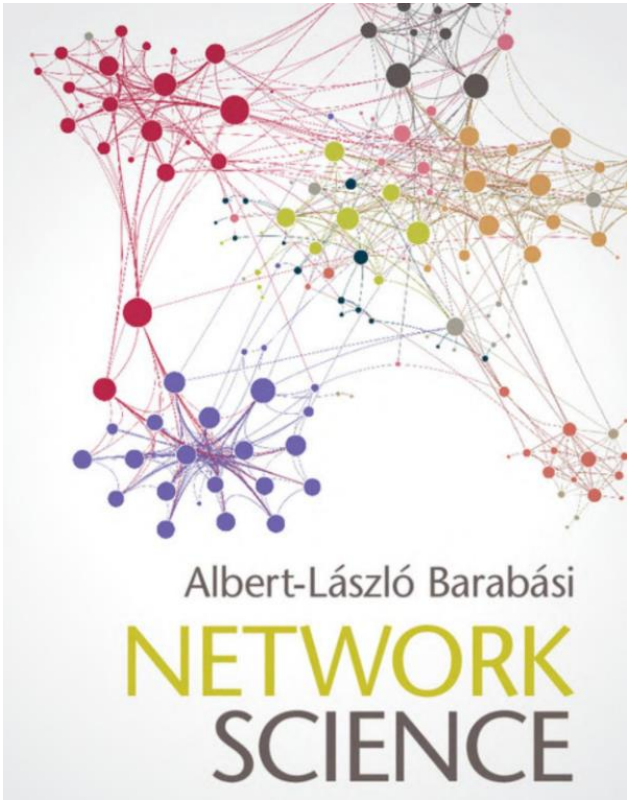
  (*You must experience the choices and compromises one must make in network mapping*).
- Manual mapping is allowed (ingredients of recipes in a cookbook; interaction of characters in a novel). Digital mapping is encouraged (scrapping data from a website or a database and reinterpret and clean the data to make it amenable for network analysis).
- *Preliminary project presentation*: offering a preview of the dataset selected – short presentation, length to be decided (probably, about 10 minutes)
- *Final project presentations:* length to be decided

all the details on pages 5-6 of AL. Barabasi book

Evaluation

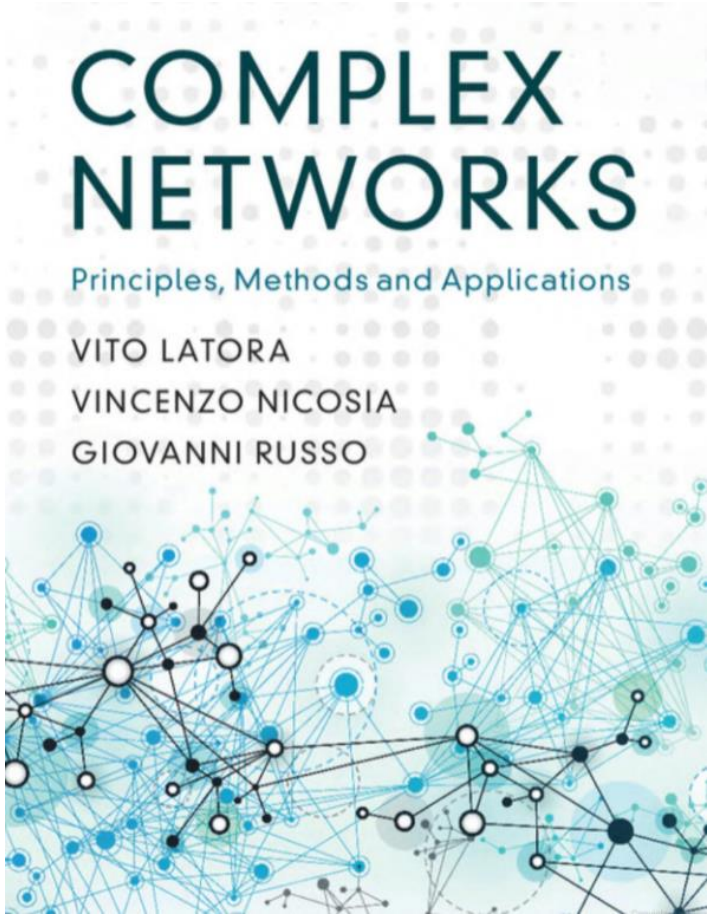
	Date (approx.)	Value	COMMENT
PROJECT preliminary presentation	19/7	**20%**	• 8% of the marks is reserved to the **quality of the presentation**
EXAM	30/8	**30%**	• short (1.5 hours) to check the understanding of the theory
PROJECT final presentation	6/9	**50%**	• 15% of the marks is reserved to the **quality of the presentation;** • 5% *extra* bonus points for the best presentation (regardless of the content) • 2% *extra* bonus points for the second best presentation (regardless of the content)

Books

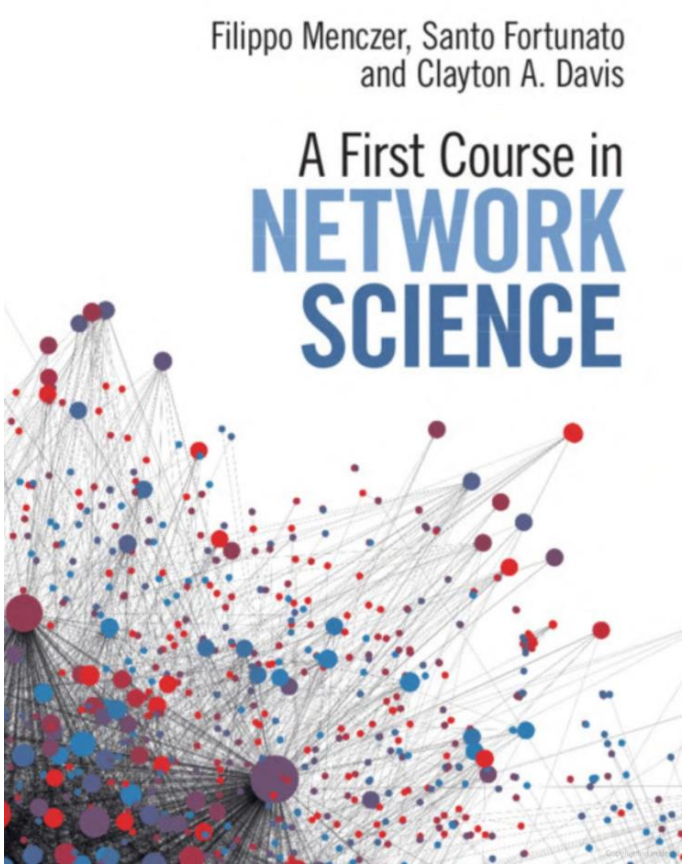


A. Barabási, and M. Pósfai.
Network science
Cambridge University Press, 2016

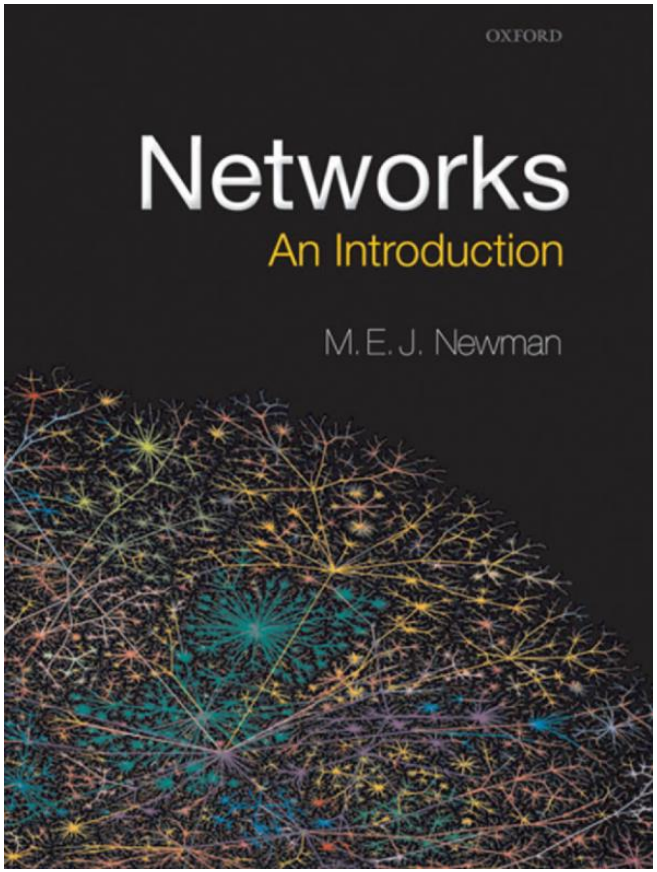
Available online at: http://networksciencebook.com/



Vito Latora, Vincenzo Nicosia, Giovanni Russo
Cambridge University, 2017



Filippo Menczer, Santo Fortunato, Clayton A. Davis
A first course in network Science
Cambridge University Press, 2020



Mark E.J Newman
Networks, an introduction
Oxford University press, 2010

A “summarized” version of this book:
The Structure and Function of Complex Networks – SIAM Review, 2003
will be made available on Eclass

William L. Hamilton

Graph Representation Learning

McGill University, 2020

https://www.cs.mcgill.ca/~wlh/grl_book/files/GRL_Book.pdf

Advices for the course

Don’t catch yourself behind

PROGRAM !!! PROGRAM !!!

PROGRAM !!!

- **How I teach**
  - I follow the book
  - I make slides and I present them during the class

- **Class rules**
  - 5 minutes interval after 40 minutes
  - Ask questions!
  - Don’t be late for class

IMPORTANT
programming test today

during the lab, 16:20-18:00

To evaluate your knowledge of Python/numpy and your programming skills