Lecture 0: Welcome to the class!

Who should take it? How to take it?

COMS 4995-1: Introduction to Social Networks
Tuesday September 4th
Lecture zero

- Why social networks matter?
- What is this class about?
- Who should take it?
- What do you need to succeed?
- Getting to know your instructors
- Roadmap
Can you recognize this picture?
All previous bubble episode were retrospectively very important to deploy game changing techs.

Late 80s ... cheap microprocessors, no applications
  - But had brought millions of pcs to business/home

Late 90s ... end of the dot-com boom
  - But the Internet infrastructure was built for future

Mid 2010s peak of the social boom
Today

What are we building for the next generation?

“The best mind of my generation are thinking about how to make people click ads.” J. Hammerbacher

“This Tech Bubble Is Different.”
A. Vance, Businessweek, 04/17/2011
The next generation could be the one with access to an unprecedented amount of behavioral data.

This can solve real problems:
- not just finding a movie or a restaurant
- ensuring energy efficiency
- monitoring our environment
- reduce inequality
- informing social decision
Only convinced by numbers?

+40%  * How much data production grows / year
  o Enough to double every 24 months
    (72h of videos upload on YouTube in 1 min).

€260b * How much data can save on health care
  o In Europe [McKinsey] (U.S. save $300b)

+300-1000%  * How much lifts improve when ads are using behavioral targeting
What are Social Networks?

* Large set of *personal information* about users
  - History of Browsing, Purchasing, Rating
  - Sociological profile (age, gender, location, income)
  - Community of interests

* Large set of *relational information* about users
  - Connections (friendship, collaboration, schoolmate)
  - Contacts (email IM phone calls etc., meeting)
What *primarily* matters is your social environment!
- For Business: how to best advertise a product?
- For Media: how to find most relevant information?
- For Engineers-CS: how to best design an application?
- For Science and Society at large: how to understand human behavior? Take advantage of it?

... 4 (classical) questions, being reinvented *today*
But here’s the thing

* By re-inventing, I mean that the problem becomes way more complex
Why social networks matter?

What is this class about?
Who should take it?
What do you need to succeed?

Getting to know your instructors
Roadmap
Objectives of this class

* Introduce *concepts* used in social networks
  o Connected to important scientific questions
  o and real systems, practical problems
* Manipulate these concepts
  1. Make them familiar
     Proof in class, Problem set to practice/experience
  2. Make them available for your critical eye
     Interpretation case-studies
Why social networks matter?

What is this class about?

Who should take it?

What do you need to succeed?

Getting to know your instructors

Roadmap
Why take this class: Reason #1

* These concepts matter to the companies that you want a job from (including your startup!)
  o Social information is becoming the web’s hottest commodity (Google, Facebook, IBM, Telcos, Media)
  o Users’s data are company’s key differentiating factor

* You (not me) are the social media generation!
  o The game is just starting; it gets harder
  o CS deals with “complexity” deeply and elegantly
  o learning foundational concepts adds to your assets
This topic is fertile for research, here and at large
  o Big data is everywhere, especially in public-funding
  o Many of these data are networks connecting people

Many opportunities within reach here at Columbia
  o Institute for Data science and Engineering
  o Brown Institute for Media Innovation
  o Lots of other departments
  o Overall, good topic for an academic job, eventually!
Computation/communication resources gradually become a commodity

- But collecting/processing YOUR Data is not innocuous
  It may not pollute the air, but it may hurt your freedom, or your fair access to service/knowledge.

Two issues need to be addressed:

1. The privacy tussle: How can a user keep track of what is known about her? And control it?
2. Is data a real currency? One you can own and use?
3. How to exploit them with better algorithms?
The topic offers a great diversity of experience
- You can hack, make proof, interview people, make money, (perhaps) save lives
- If you have broad interests (or wonder what you would like), this could be a way to figure out

Hence, diversity in this class is highly encouraged!
- Special background, atypical majors/minors, are especially enriching to our discussions
- Previously: CS, EE, Mech, Civil, Med. school
Before starting the trip
Before starting the trip

* There is no textbook!
* Or maybe there are too much
Why NOT to take this class?

* Bad choice for maximum GPA / minimum workload
  o Combines programming, maths, interpretation
  o New, not a single textbook, may at time be unfamiliar
Take this class if you’ll enjoy it, it’s not a requirement!

* Bad choice if you don’t tolerate contradictions
  o How come two models predict different results?
    What is THE model for network structure?
At least, be reassured that the instructor sympathizes
How is your 1st social network class?
Social processes are (1) on graph, (2) messy
- Getting comfortable with definition is not sufficient
- You need to use them to prove results

Prerequisites:
- Graph theory: vertices, edges, paths, shortest paths,
- Probability: expectation, ind., condit., Markov Chains
- Calculus (bounds on infinite sums, etc.)
- Linear Algebra: eigenvalues, spectral theory
- Programming: one program. language

Why NOT to take this class?
More clarifications

* This class is not “playing on FB and get credit for it”
  o although many of your friends may think it is

* This class is not “how to build large web systems.”
  o although some of what we see matter for it
  o Take Prof. Geambasu’s class or classes on cloud

* This class does not cover game theory / deep ML
  o Very important topics, covered elsewhere,
  o We will only discuss these occasionally (no prerequisite)
The topic is **broad**: “CS-theory, Networking, Sociology, Physics”

- This is why the course focuses on algorithmic prop.

The topic seems (at times) **immature**:

- “What is THE model? How to tell the cause?”

- Algorithmic research problems have an impact

Involves **substantial** maths & programming
How to succeed in this class?

* Attend the lecture (remotely, for CVN students)!
* Come to our office hours: easy, close to the course!
  Tuesday 8:15-9am (A. Chaintreau)
  Possible to attend through Skype, Hang-out, if requested before
* Your grade: (no extra-credit)
  o ~3 homework assignments
  o Activities: data contests, blogs, etc.
  o 1 (short) midterm, 1 final
    Generally speaking: hard questions, generous grades
More on the course

* Our wiki: TBD
  Slides + Assignments + Readings + Additional References
  To ask and answers all questions to the class and instructors
  Unless otherwise specified we promise you a 24h maximum delay

* Our blog:  http://social-network-and-computing.tumblr.com
  Where you can read and write post, mostly about news

* Integrity Rules:
* The “Apple” Policy:
A bit about myself

- Started research in 1999 in Bay area
  ... I could not even order a beer then!
- Attended ACM SIGCOMM since 2000
- Studied at ENS-INRIA in Paris (Ph.D in 2006)
  - Interns at Sprint, Alcatel, IBM, Intel
  - Worked 5 years for Technicolor (formerly Thomson)
- Works on Mobile, Social Networks, Privacy
  - Previously, models of TCP, Peer-to-peer, Human mobility
  - Emphasis on performance of networked algorithm
Meet the TAs!

* Xiao
  * From Shanghai
  * MS CS Mach. Learn.
  * Work on
    o Social Media
    o Ads & Privacy
    o Mobile Computing
  * Worked at twitter this summer
  * English, Mandarin
  * Office Hour TBD
Contents:

- Structure (September)
  * small world, weak tie, homophily, balance, ...
- Dynamics (October)
  * epidemics, influence, wisdom of the crowds ...
- Discover (November)
  * crawling, source detection, reconciliation
- The 10 papers that will make you a social expert
- Some case studies, discussion within/outside lectures