Social and Technological Networks

Rik Sarkar

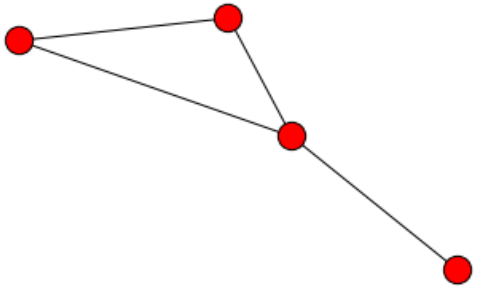
University of Edinburgh, 2017.

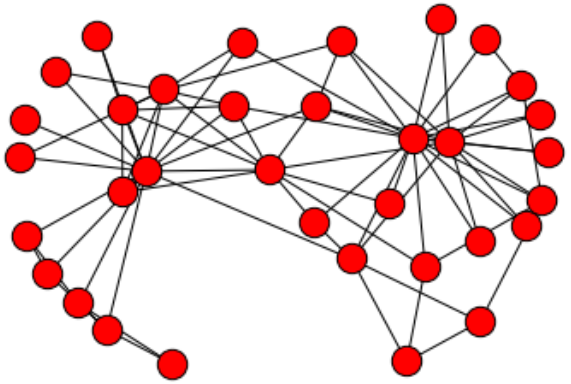
Course specifics

- Lectures
  - Tuesdays 12:10 – 13:00
    - 7 Bristo Square, Lecture Theatre 2
  - Fridays 12:10 – 13:00
    - 1 George Square, G.8 Gaddum LT
- Web page
  - http://www.inf.ed.ac.uk/teaching/courses/stn/
- Lookout for announcements on the web page

Network

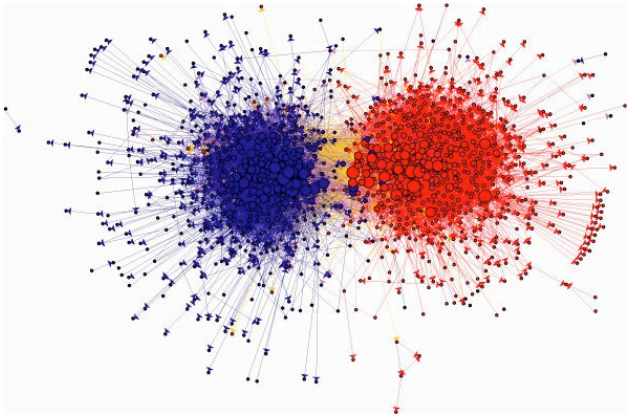
- A set of entities or nodes: V
- A set of egdes: E
  - Each edge e = (a, b) for nodes a, b in V
  - An edge (a,b) represents existence of a relation or a *link* between a and b





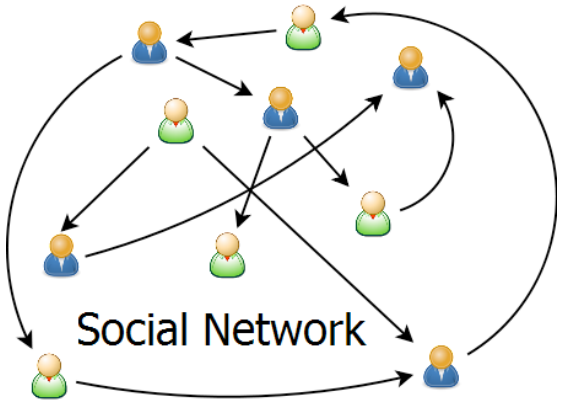
Networks are everywhere

- There exist different relations between components components in a system
  - There is a network
- Properties of the network determine properties of the system
- Analysis of data from the system must take the network into consideration.



Example: Social networks

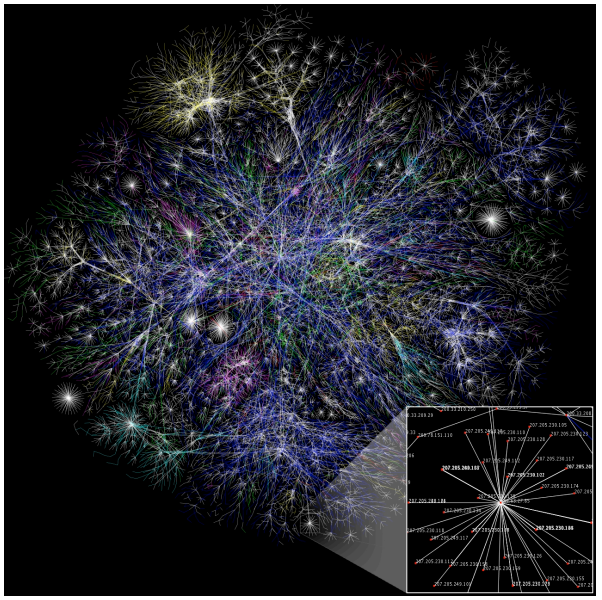
- Facebook, Linkedin, twitter..
- Nodes are people
- Edges are friendships



- The network determines society, communities, etc..
- How information flows in the society
- How innovation/influence spreads
- Who are the influential people
- Predict behaviour

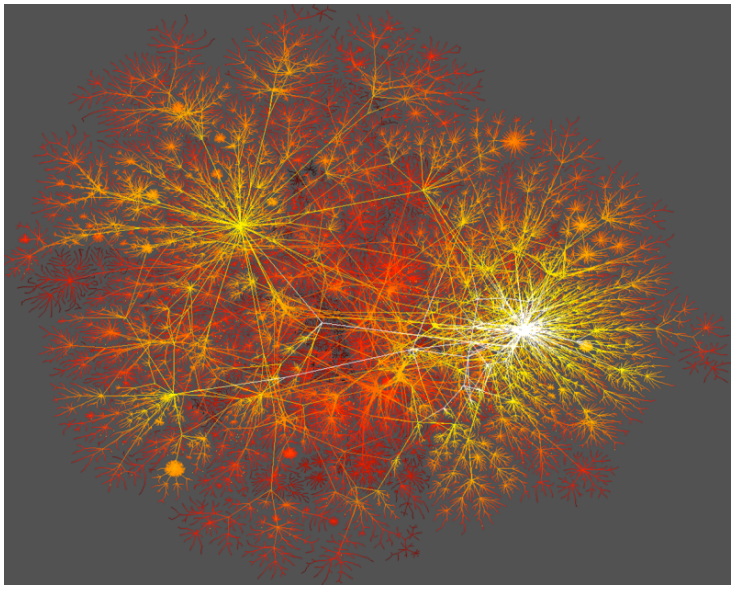
World wide web

- Links/edges between web pages
- Determines availability of information
- Important pages have more links pointing to them
- Network analysis is the basis of search engines



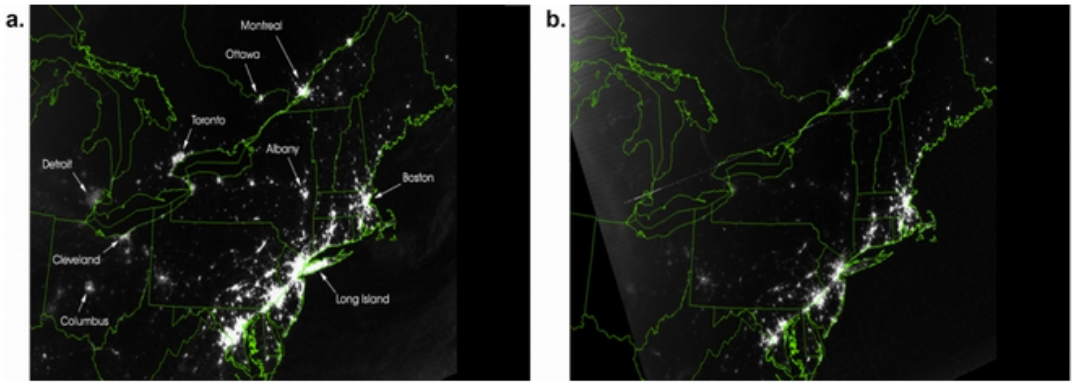
Computer networks

- What can we say about the internet?
- How reliable are computer networks?



Electricity grid

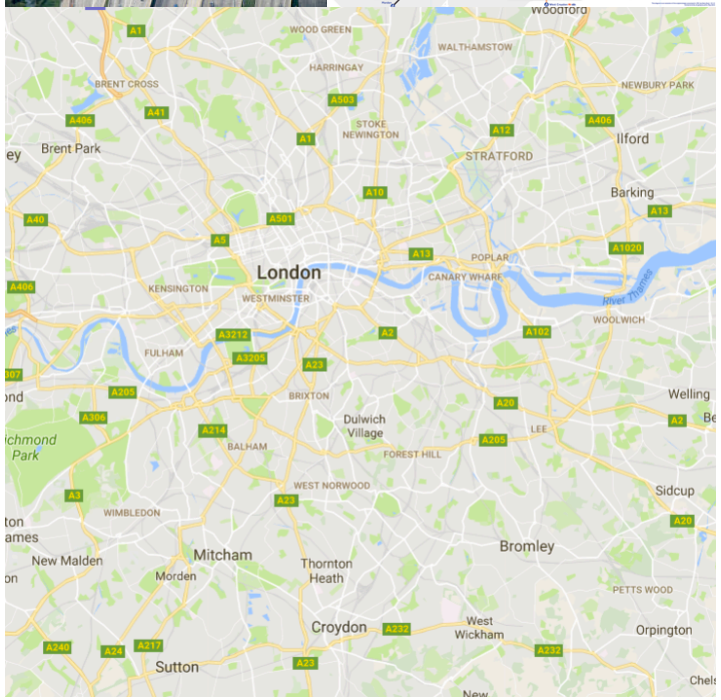
- Network of many nodes, redistributing power
- Critical infrastructure
- Failure can disrupt … everything
- Small local failures can spread
  - Load redistributes
  - Trigger a casdade of failures
- Network strcuture is critical



From Barabasi: Network Science

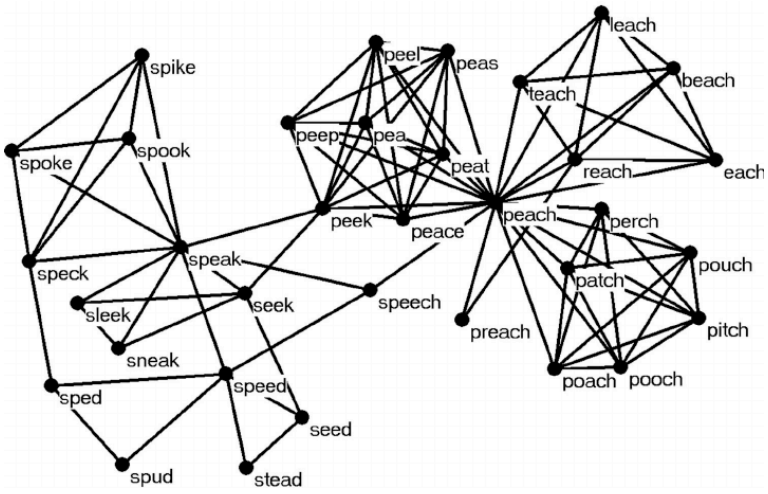
Road network and transportation

- Mobility patterns of people
  - Location data
- Failure cascades
- Traffic needs
- Suggest bus routes
- Suggest travel plans
- Traffic engineering
- Increasing importance
  - More vehicles
  - Self driving cars



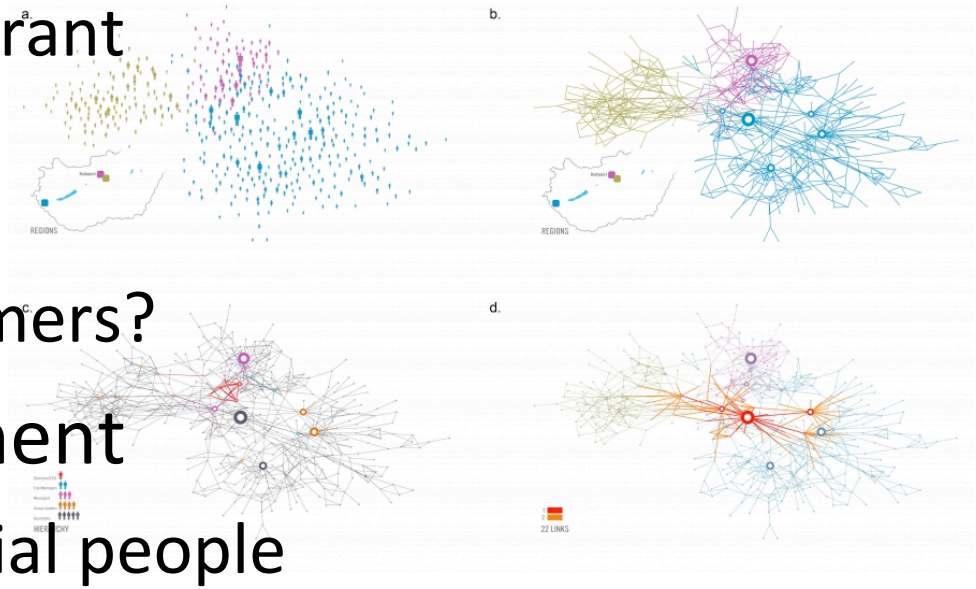
Linguistic networks

- Networks of words
- Show similarities between languages
- Show differences between languages
- Document analysis



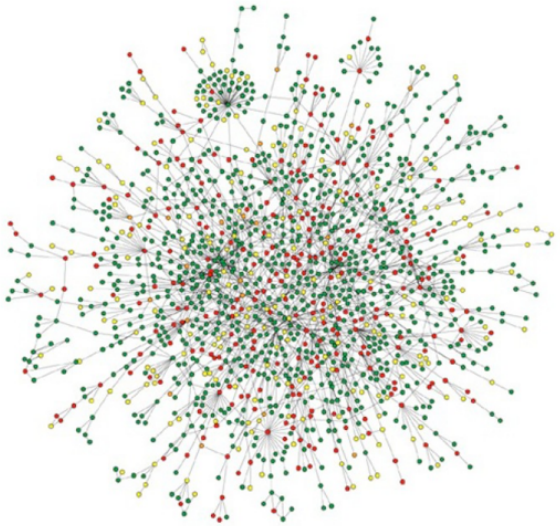
Business and management and marketing

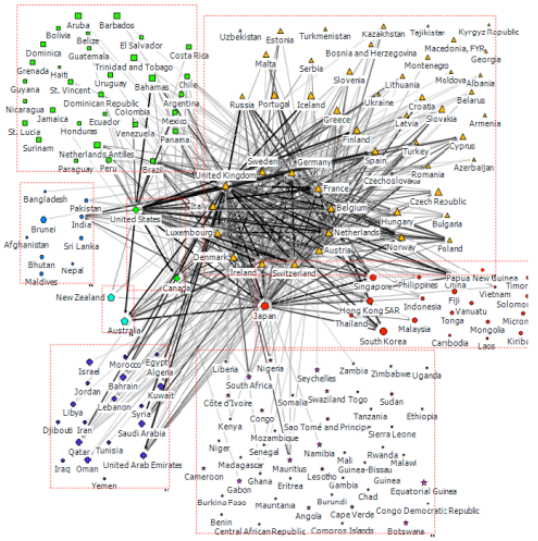
- Business
  - What makes a restaurant successful?
  - Nearby restaurants? Community of customers?
- Marketing/management
  - Who are the influential people in spread of ideas/products?



Other networks

- Chemistry/biology
  - Interactions between chemicals
  - Interactions between species
  - Ecological networks
- Finance/economies
  - Dependencies between institutions
  - Resilience and fragility
- Neural (Brain) networks





Why Network science? Why Now?

- Many of these systems have similar underlying characteristics
- Network science studies these general properties

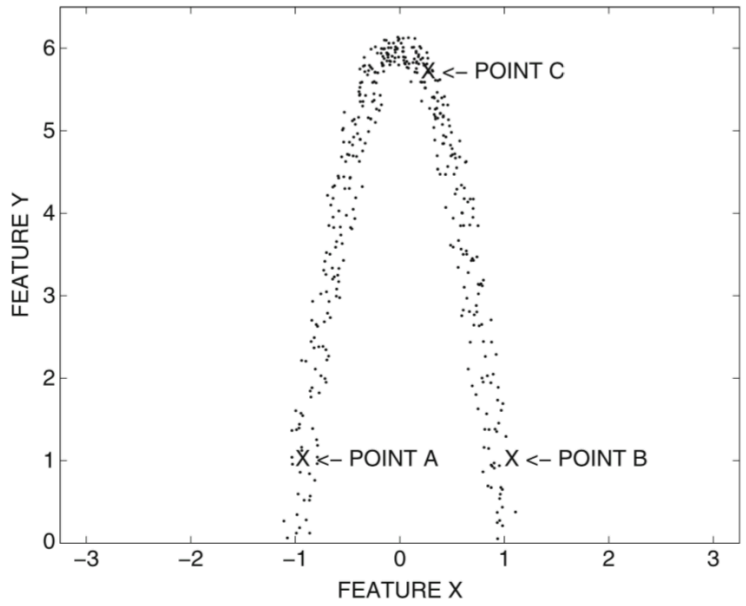
- We now have many tools: algorithms, graph theory, optimization…
- Last decade or so a lot of network-type data has become available
  - www – search engines etc
  - Location data: traffic and road data
- We can now look at this data and search for theories

Network analysis in data science

- Data getting more complex
- Many types of data are not points in R^d space
  - Data carry relations – networks
  - Simple classification inadequate
  - Network knowledge can make ML more accurate, efficient
  - E.g. data from social network or social media, www, IoT and sensor networks

Network analysis in data science

- Networks reflect the *shape* of data
- Connect nearby points with edges
- Analyse resultant network



The breadth of network science

- Tied to real systems
  - Anything in network science has impact on multiple real things
- Data driven
  - Need good data-handling techniques, optimizations, approximations
  - Get to learn data driven thinking
  - Study of algorithms, data mining
- Mathematical and rigorous
  - Emphasis on precise understanding, provable properties. Clear thinking.
  - Exactly what is true and what is not, what works and what doesn't, in exactly which circumstances

Topics of study

- Random graphs: the most basic, unstructured simple networks
  - What are their properties? What can we expect?
  - Erdos renyi graphs
  - Construction of random graphs
- Power law and scale free networks
  - Distribution of degrees of nodes
  - Power law occurs in many places: www, social nets etc..
  - What is the process that generates this? How do we know that it is the right process?
- Metrics and distance measures in networks
  - Basis of classification, clustering, route planning etc

Topics of study

- Small world networks
  - Milgram's experiment
  - What is the deal with six degrees of separation
  - How are people so well connected?
- Web graphs and ranking of web pages
  - Google's origins and pagerank
  - How do you identify important web pages?
  - Analysis of the algorithm: do they converge? Can they give a clear answer?
- Spectral methods

Topics of study

- Strong and weak ties in social networks, social capital
  - How does information spread in a social network?
  - How do you make use of your position in a network?
  - Which contacts are useful in finding jobs? Why?
- What are the communities (close knit groups)?
  - How do communities affect social processes?
  - Clustering/unsupervised learning

Topics of study

- Cascades – things that spread
  - Node failures
  - Epidemics, diseases
  - Innovation – products, ideas, technologies
- How can we maximize a spread?
  - Who are the most influential nodes?
  - How can we identify them?
  - Submodular optimization

Topics of study

- Shape of networks
  - What is the shape of internet?
  - What are bow tie and tree-like networks?
  - What does it mean to say a network is tree-like?

The course

- Is not about:
  - Facebook, Whatsapp, Linkedin, Twitter…
  - Making apps

The course

- Is about:
  - Understanding mathematical measures that define properties of networks
  - Mathematics and algorithms to compute and analyze these properties
  - How machine learning applies to networks, and vice versa

Our approach

- Clearly define different aspects of networks
  - What is a random graph?
  - What exactly is a small world?
  - How do you define 'community' or clustering in networks?
  - How do you define influential nodes?
- Design algorithms to analyze networks
  - Find communities, find influential nodes
  - Understand the properties of these algorithms
  - When do they work, when do they not work
    - Why?

Our approach

- Test ideas on real and artificial networks
  - Data driven understanding
  - Do real networks have the properties predicted by theory?
  - Do the algorithms work as well as expected?

Project

- 1 project. 40% of marks
- Given: Around Oct 5 to 10.
- Due: Around Nov 15.
- Choose from one of several projects
- **Objective: Try something new in network science.**
- Given problem statement, try your own ideas on how to solve it
  - No unique solution.
- We will give you a topic. You have to
  - Formulate it as a precise network problem
  - Find a way to solve it
  - You are allowed to try different problems and approaches
  - Or define your own topic
- Submit code and ≈3 page report
- Marked on originality, rigor of work (proper analysis/experiments), clarity of presentation

Possible types of projects

- Given a dataset from a particular social/ technological area, find a way to solve a particular problem
  - Devise a prediction method
  - Find interesting properties of specific networks
  - Design of efficient algorithms to compute network properties
- Programming is useful for evaluation/ experiments
  - We will use python in class (recommended)
  - You can use other languages (python, java, c, c++)
- Theoretical work is also great. But must have analytical approach such as proofs

Theory Exam

- Standard exam, 60% of marks
- Explain phenomena, devise mechanisms, prove properties…
- Last year's paper online..

Lectures

- Slides will be uploaded after each class
- Lecture notes will be given covering some material left over
- Exercise problems will be given covering important material
- Ipython (jupyter) notebooks will be uploaded
- Do the exercise problems to make sure
  - You understand things
  - You can solve analytic problems
- Solutions will be given later for important problems
  - Check that your solution is right
  - Check that your writing is sufficiently precise

Pre-requisites

- Probability, distributions, set theory
- Basic graph theory and algorithms
  - Graphs, trees, DFS, BFS, minimum spanning trees, sorting
- Asymptotic notations: Big O.
- Linear algebra
    - Matrix operations
    - (preferably) Eigen vectors and eigen values
- Sample problems online

- Take notes in class. Not everything is on slides!

Course learning expectations

- Formulate problems
- Plan and execute original projects
- Use programming to analyze network data
- Use theoretical analysis (maths) to understand ideas/models
- Present analysis and ideas
  - Precisely
  - Unambiguously
  - Clearly

- Have fun playing with ideas!