CORSO DI LAUREA IN
INGEGNERIA INFORMATICA

SOCIAL NETWORKS ANALYSIS
A.A. 2021/22

INTRODUCTIONS TO NETWORKS

NETWORKS ARE UBIQUITOUS

- Modern society is highly connected in several different ways
  - Global comunication systems
  - Internet
  - Social networks
  - Financial and techonological systems
  - News and media

- All these elements share the idea of network
  - A collection of objects and relations between objects

Why Are We Interested to Networks?

- Networks are a general framework that can be applied to every system consisting of different components linked in any ways
- It can be applied to a large number of very different domains
  - Mathematics, Computer Science, Physiscs, Chemistry, Biology, Engineering, Economy, Sociology, Social Sciences, etc.
- Our technological and economic systems are based on very complex networks
  - Susceptible of disruptions that spread through the network structure and can transform localized breakdowns in cascading failures or global crises
  - It's of huge importance to learn about network effects to deal with such events

What Are We interested in?

We are interested in understanding

- Nature and behaviour of each single component of the system
  - Es. How a computer works? how an individual or a society behaves in a market? Which is the behaviour of a cell?
- Nature of the connections/interactions among different components
  - Es. Communication protocols in Internet, dynamics in social relations, cellular/molecular communication mechanisms
- Communication patterns among system components
  - Can be represented as a network
  - Can give information on the system behaviour and on how components' behaviours are related

Structure of a Network

- The communication pattern captures the structure of a network
- The network structure influences how the system works
  - Selection of the routing paths in Internet and latency of the transmission
  - How people learn, construct their opinions, gather information through their social relations
  - How diseases or trends spread
  - How we choose our friends
  - How organisms evolve and how they interact with their ecosystem

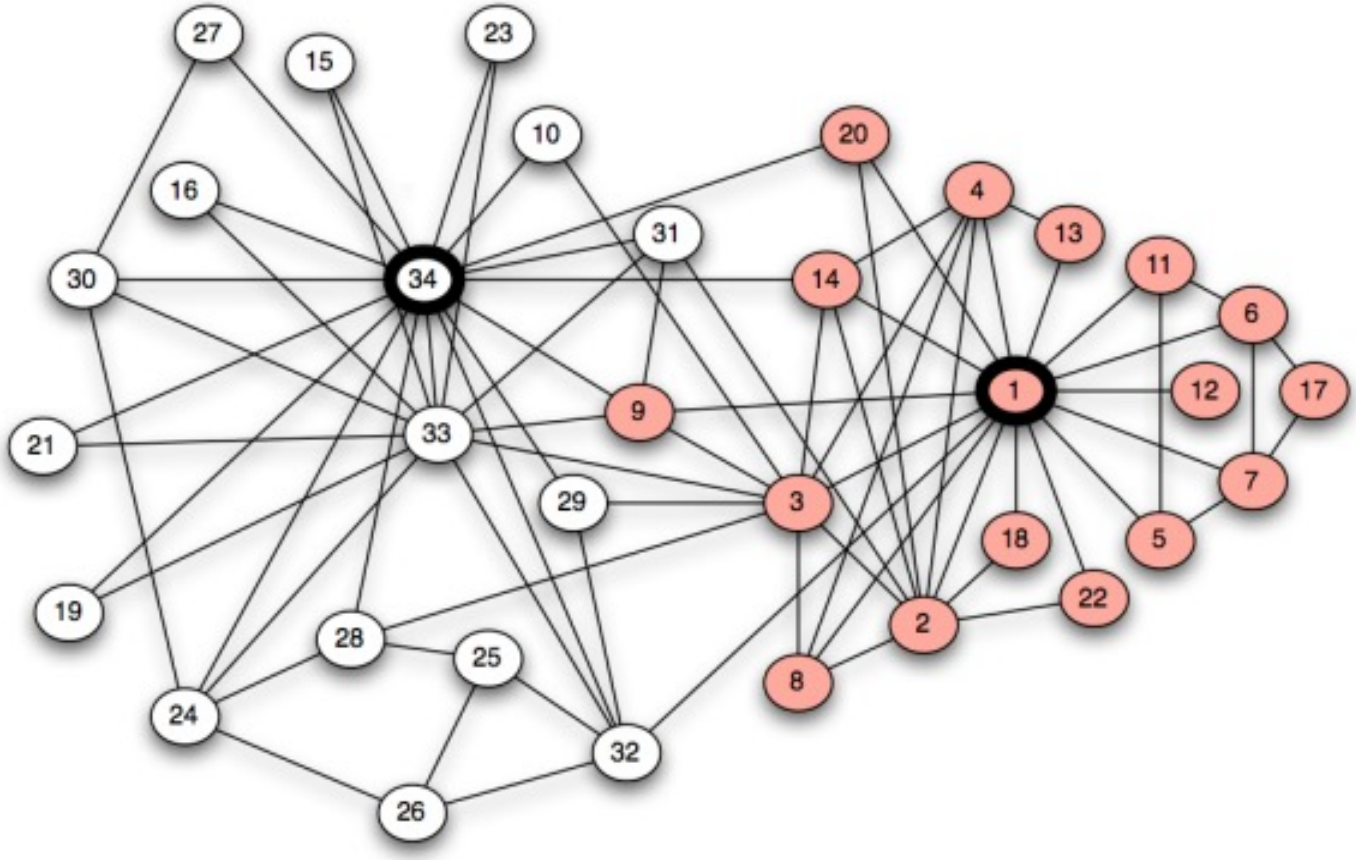
A Multidisciplinar Approach to Network Science

- Understanding highly connected systems requires a set of ideas for reasoning about network structure, strategic behavior, and the feedback effects they produce across large populations
  - Ideas traditionally dispersed across many different disciplines
- Nowadays there exists a large set of mathematical, computational and statistical tools to analyze, model and study networks
  - Developed in different areas to interpret different phenomena …
  - But usable also in areas totally different from their original application field because all based on the abstract framework of network
- **Network Science is a new discipline that studies**
  - *"network representations of physical, biological and social phenomena leading to predictive models of these phenomena "*
- A discipline intrinsecally multidisciplinar
  - Put together ideas and techniques developed by sociologists, economists, mathematicians, physicians, biologists, geneticists, computer scientists

Network Science's Objectives

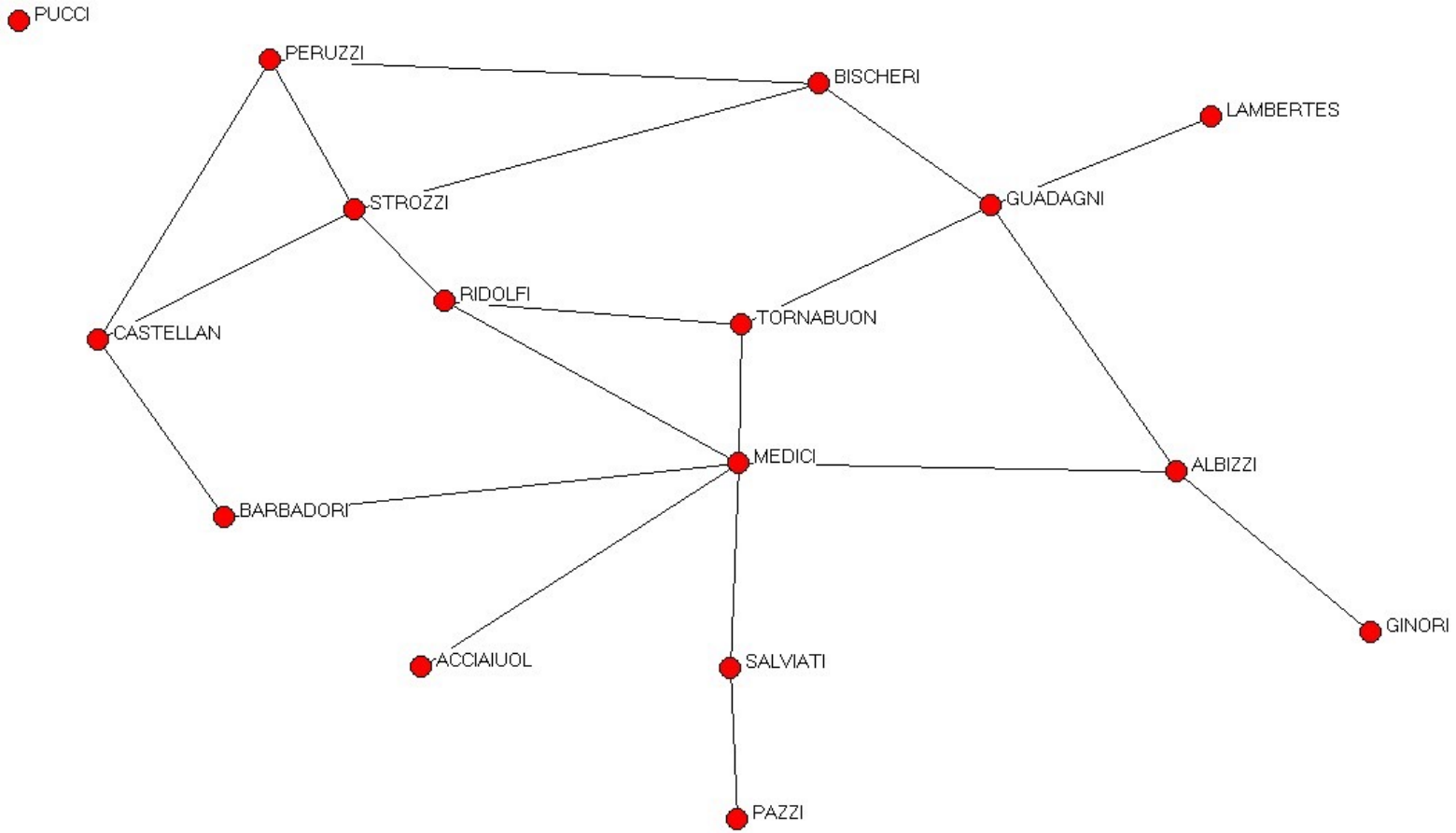
- Networks pervade all our modern society and they have a fundamental role in our lives
  - Our economic and technological systems are based on highly complex networks
- It's crucial to understand phenomena occurring in a network in order to control and guide them
  - Eg. Global markets crisis, spread of computer virus, spread of information, identification of communities
  - What kind of a network emerges from local interactions among simple agents and how this influences agents' behaviour?
- Network Science looks for
  - indices to model structural characteristics of the network
  - efficient algorithms to compute such indices
  - mathematical models to predict the evolution of important processes occurring in a network

Some Network Examples



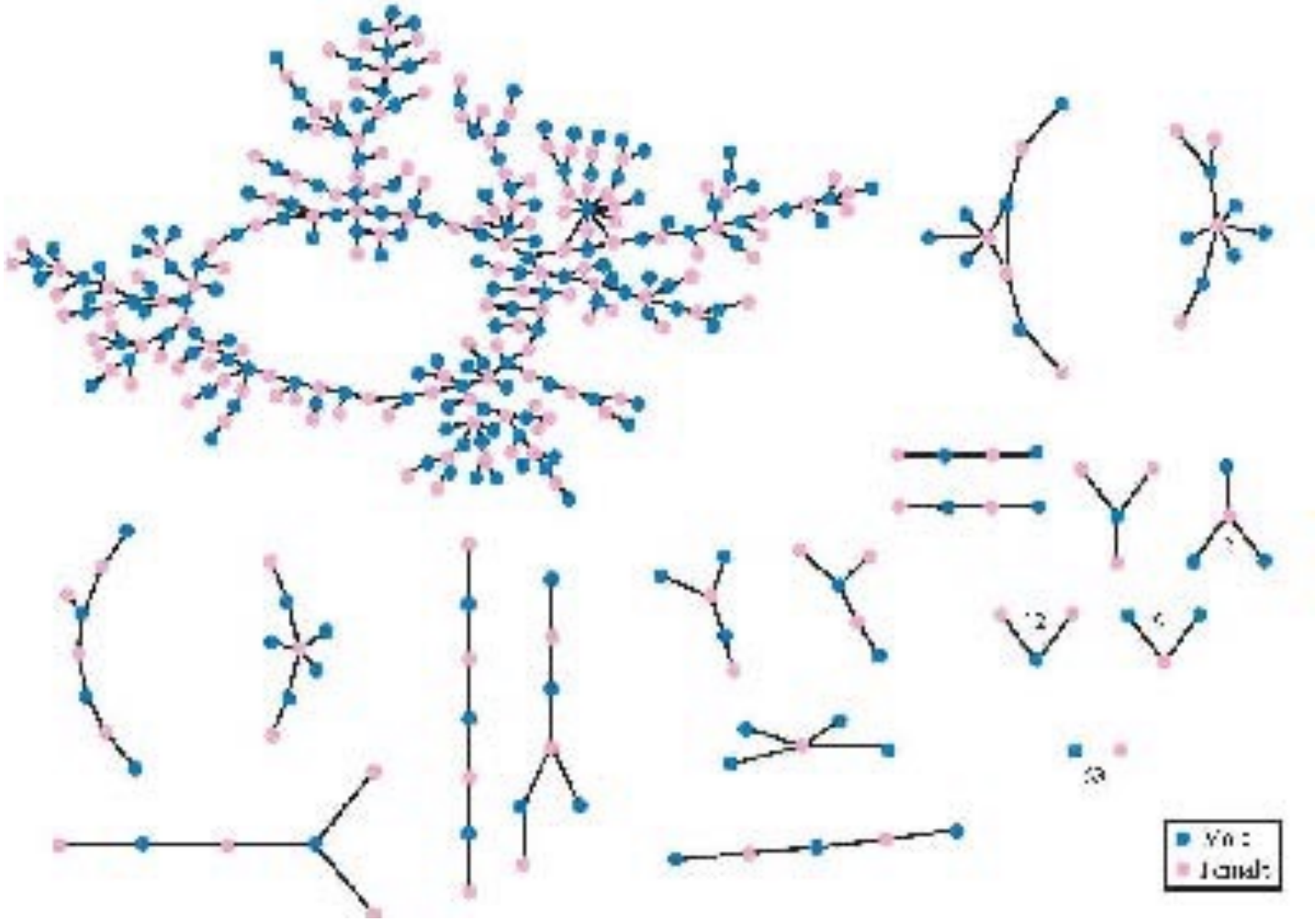
- Network of the friendships among components of a Karate Club
  - Who are the most influential persons in this social network?
  - Is this a conflicting configuration?

Some Network Examples



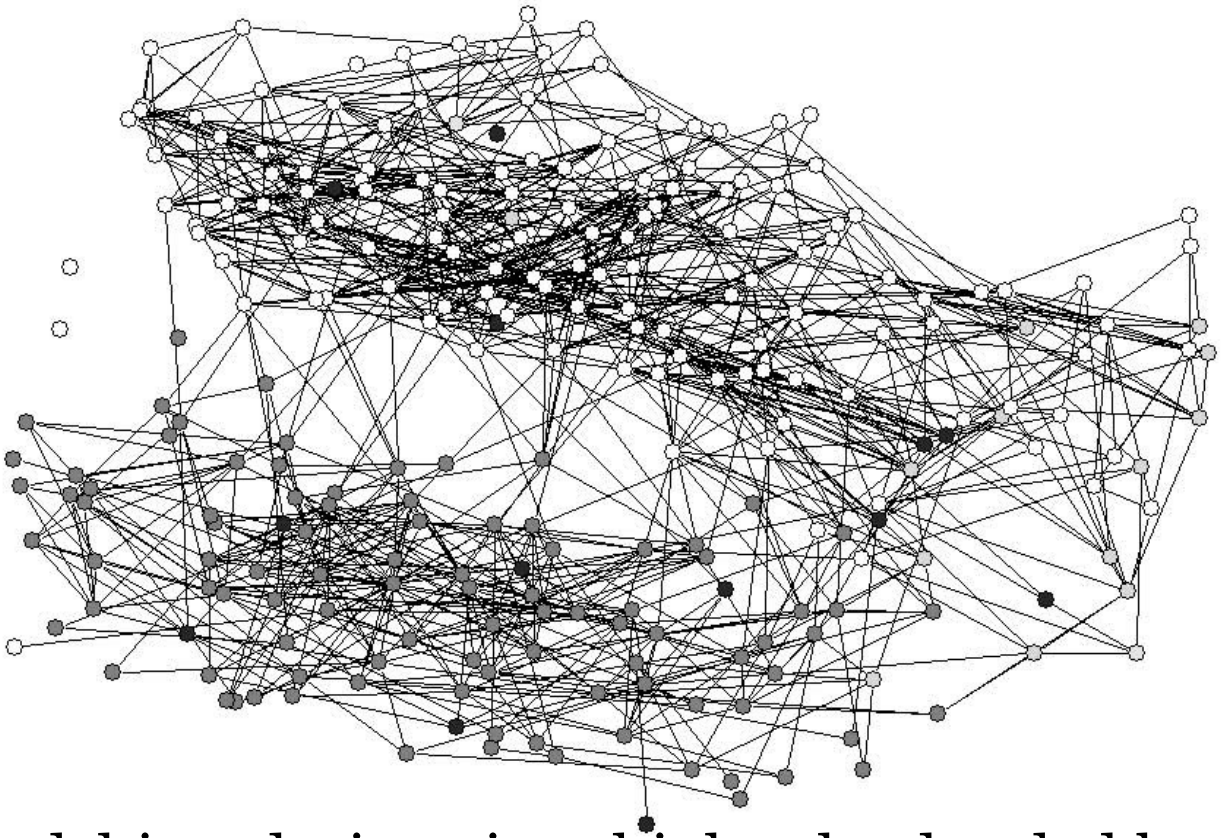
- Network of marriage relations between some powerful florentine families in the rinascimental period
  - Why Medici were so powerful even if they were not the richest family?
  - Was the Medici's rise only due to Fate?

Some Network Examples



- Network romantic relations among students in a high school
  - Why sexual diseases spread so fast?
  - What happens to information?

Some Network Examples

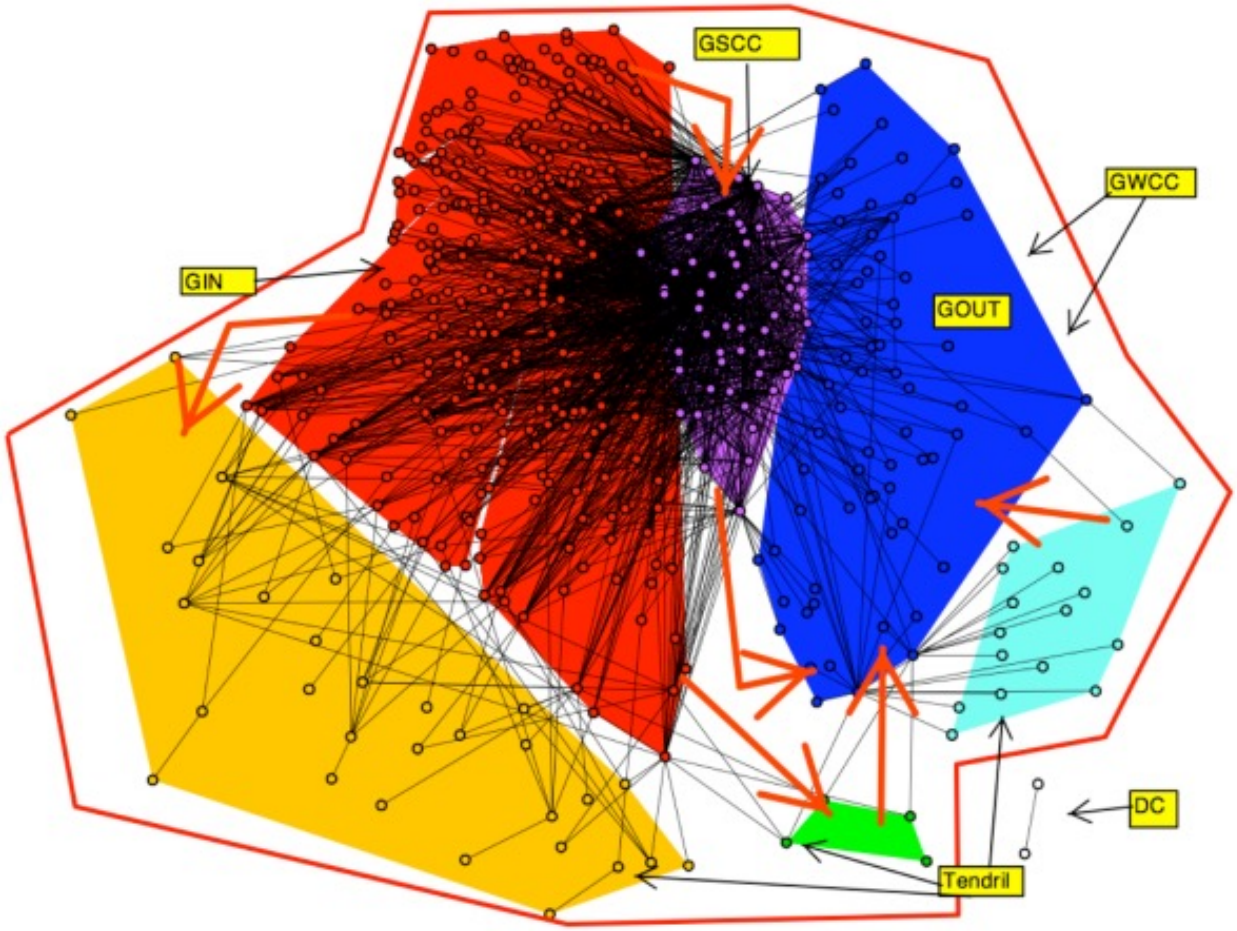


- Friendship relations in a high school coded by race
  - White = White
  - Black = Hispanic
  - Grey = Afro-American
  - Light Grey = Asian and others

Some Network Examples

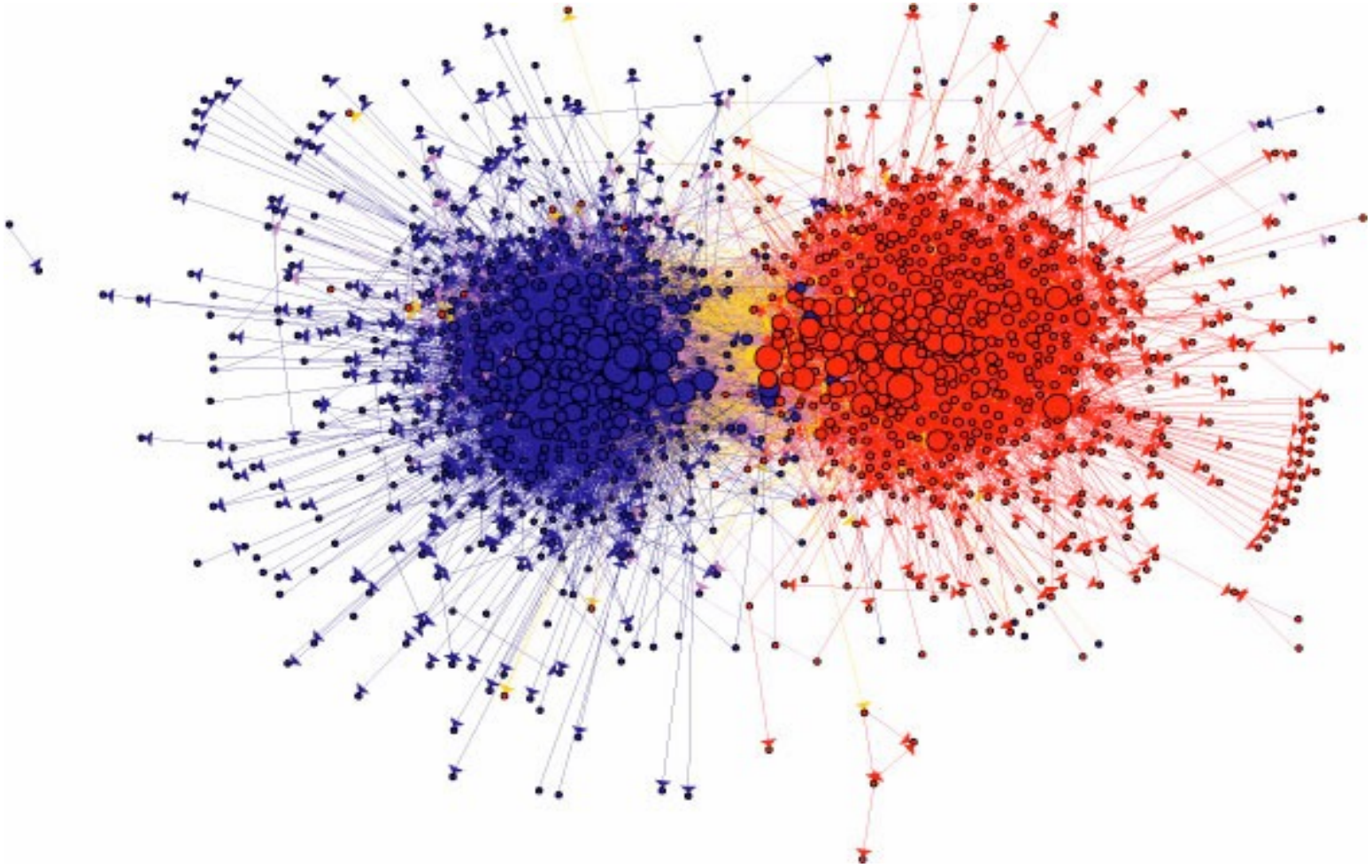
- E-mail communications among HP employers
  - Superimposed to the organizational hiearachy of the society
- A social network is tipically composed of several communities weakly linked

Some Network Examples



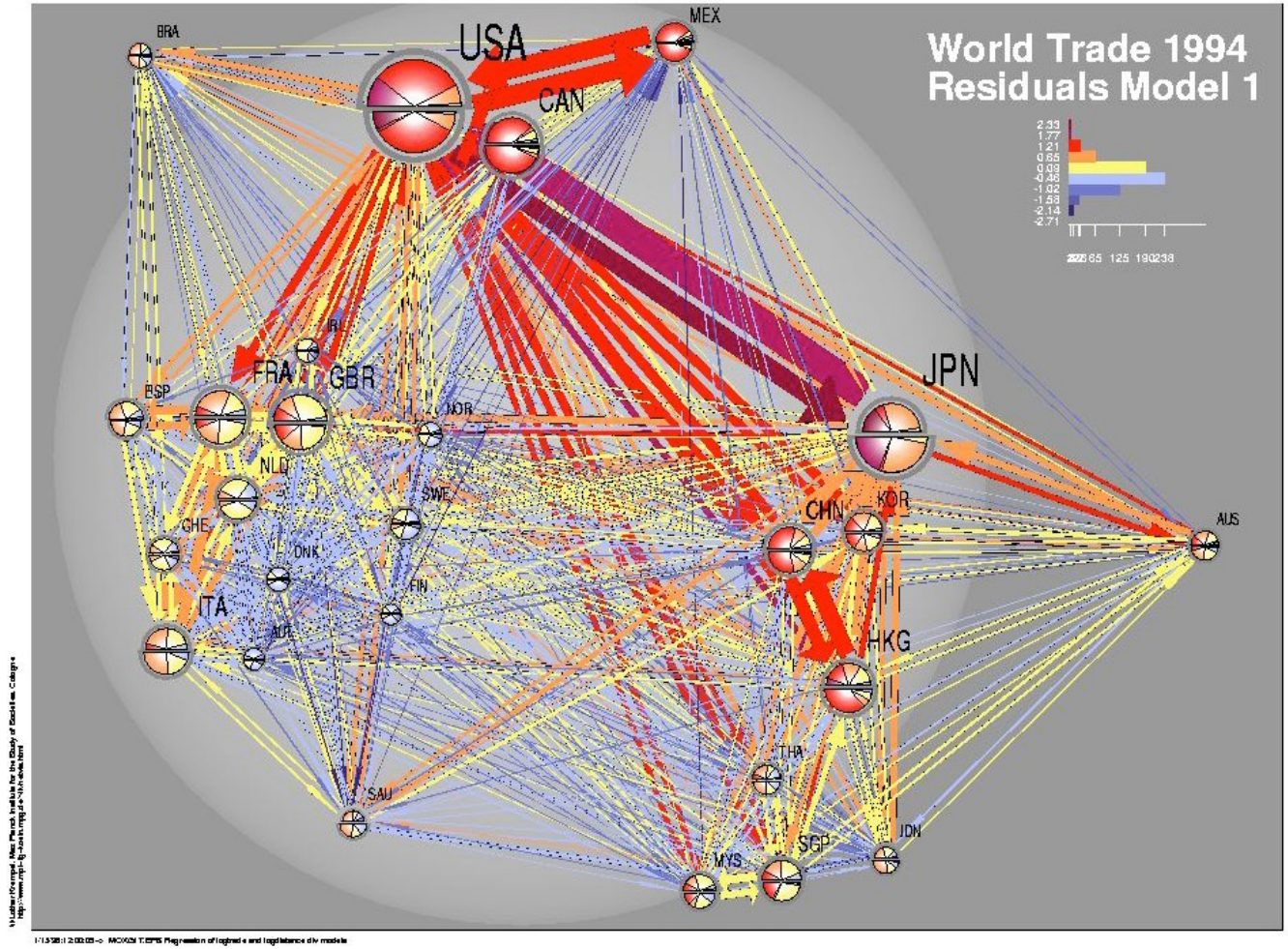
- Network of loans among financial institutions
  - Who are the most influential institutions?
  - What is the role of each institution in the global financial system?
  - Is the system healthy?

Some Network Examples



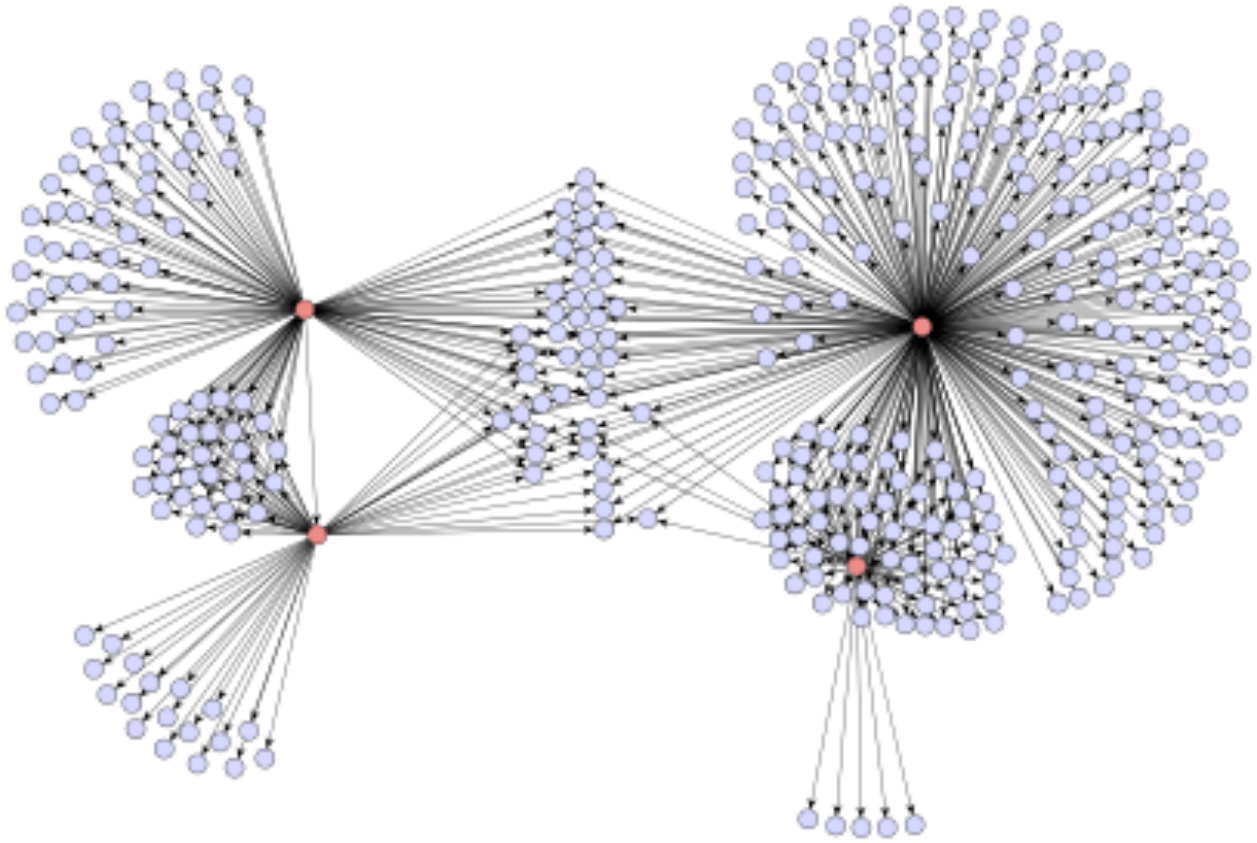
- Links among political blogs during the american presidential campaign in 2004
  - Who are the most influential bloggers?
  - Is it possible to identify clusters?

Some Network Examples



- Network of the international trades among countries
  - Arcs are weighted by the monetary value of the trades
  - Which are the countries that obtain the largest benefits from their position?

Some Network Examples



- Email recommendations for a Japanese book
  - Individuals are influenced in their choices by their neighbours
  - Adoption of a new behaviour can spread through the network structure
  - Who are the best initial nodes for an advertising campaign?

Some Interesting Questions about Networks

- **Which are the structural characteristics of a network?**
  - Difficult to find by hand on very large datasets
  - We need some indices and algorithms for extracting useful information from massive datasets
- **How network components behave?**
  - Strategic behaviour
  - Incentives to induce a given behaviour
  - Complex relations of cause-effect
- **Which dynamics describe the aggregate behaviour in a population?**
  - Why youTube and Facebook are so popular?
  - How our behaviours and our opinions are influenced by our social relations?
  - How information/innovation/influence spread in a network?

Multidisciplinarity of the Network Science

- To study phenomena in a networked-system we need different know-hows
  - Graph Theory (Mathematics) gives tools to describe and analyze network structure
  - Game Theory (Economics) gives tools to describe and analize strategic behaviours of the network components and to define incentives to induce specific behaviours
  - Sociology gives tools to analyze the structural properties of social networks and describe dynamics within a community
  - Computer Science gives tools and techniques to mine in massive datasets

Mining Data

- Data mining is a very generic term that is used in different contexts with different meanings
- In statistics it means to find a probability distribution from which observed data could be extracted
- In machine learning means to train learning algorithms to recognize some features and answer to complex queries
  - Eg. Predict how successful a product will be
- In computer science means to find hidden features in a massive dataset
  - Representing data succinctly and approximately
    - Eg. Pagerank represents the relevance/popularity of a web page
  - Extracting the most prominent features from the observed data
    - Eg. Similar items are used for suggestions in Amazon

Limits of Data Mining

- When you look for some rare events in a data set you will find it even if data are completely random
  - These occurrences are *bogus* and their number of occurrences grows with the size of dataset
  - Eg. Total Information Awareness was a project promoted by american government to mine a huge amount of data (credit card receipts, hotel records, travel data, etc.) for tracking terroristic activities
    - The project gave rise to a large number of false positives and it has been (maybe) killed but the Congress

- Bonferroni's principle helps in recognizing significative features
  - A feature is significative only if the number of its occurrences is sufficiently larger than the expected number of occurrences when data are completely at random

An Opportunity for Cross-Fertilization

- Network science is not only a sum of different disciplines but an extraordinary opportunity for cross-fertilization
  - Each discipline can be unriched by ideas coming from other disciplines
- Economists have theories to model strategic interactions among either a small community of economic subjects or within the whole population
  - Most of the interesting economic interactions develop in an intermediate situation where macroscopic effects come from an intricate pattern of local interactions
- Sociologists study the structure of social networks
  - Their methodologies do not scale to networks of large size
- Computer Science deals with the design of computation systems
  - In the era of Web and social computation we must deal with agents' selfishness
  - They use the system and interact with the other components not as defined by a protocol but according to their interests

A (tentative) schedule of the Course -- 1

- Network analysis
  - Representing network data
  - Mining network data to extract useful information about interesting phenomena
- Markets and strategic interactions
  - Modeling economic interactions that occur in a network
  - Auctions
  - Understanding how inteactions are produced and which are the most influential nodes
- Searching and mining informational and social networks
  - Link analysis
  - Web advertising
  - How search engine work and how they influence agents' behaviour
  - How to orchestrate a marketing campaign on the web or on a social network

A (tentative) schedule of the Course -- 2

- Network Dynamics: population effects
  - Understanding how ideas, trend, information, popularity spread through a network
    - An agent is influenced by her neighbours and she tend to conform
    - Rich-get-Richer phenomenon
- Network Dynamics: structural effects
  - Cascading behaviours
  - Epidemics
  - Social influence and influence manipulation
- Social Learning
  - Bandit algorihms