Analysis of Large Networks
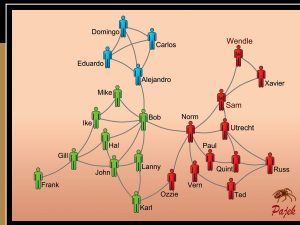
Syllabus

- Network
- Types of networks
- Size and density
- Walk, chain, path, closed walk, cycle, closed chain, loop
- Path length, the shortest path, diameter, average path length
- k-neighbour
- Program Pajek

- Definition of Cluster and Partition
- Cut-outs
- Components
  - weakly connected components
  - strongly connected components
  - biconnected components
- Cores
- Triads and cliques
- Global views
  - reduction
  - hierarchy
- Local views
  - cut-out
  - context

- Centrality and centralization
  - degree
  - closeness
  - betweenness
  - proximity prestige
  - hubs and authorities
- Brokerage roles
  - coordinator
  - itinerant broker
  - representative
  - gatekeeper
  - liaison

Test

Additional types of networks:

- Multiple relations networks
- Temporal networks
- Two-mode networks, multilevel networks
- Signed networks

For eachy type:

- Examples
- Presentation in Pajek
- Typical analyses and visualisations

- Generating networks
  - Networks obtained from WWW (WoS, Web Crawlers)
  - Networks obtained from texts
    - semantic networks
    - associations
    - dictionary networks
    - words transformation networks
  - Networks obtained from movies (Lindenstrasse)
- Excel2Pajek, text2Pajek

- Acyclic networks
- Properties of acyclic networks
  - source and sink vertices
  - topological sort
  - visualization in layers
- Computing traversal weights in citations networks
  - SPC - Search Path Count
  - SPLC - Search Path Link Count
  - SPNP - Search Path Node Pair
- Searching for main paths
  - local main paths
  - global main paths
  - key-routes

- Collecting genalogical data
- GEDCOM standard for storing genalogical data
- Representing genealogies as networks
  - Ore-graph
  - p-graph
  - bipartite p-graph
- Relinking marriages
  - relinking index
  - blood marriages
  - relinking marriages
- Searching for relinking marriages
- Special genealogies
  - Students and their PhD thesis advisors
  - Noble families
  - American presidents

Additional measures and approaches for analysing large networks:

- Clustering coefficient

- Short cycles

- Islands

- Community detection

- E-I Index

- Comparing partitions
  - Cramer's V
  - Rajski coefficients
  - Adjusted Rand Index

- Standard network models
    - Bernoulli-Poisson and Erdös–Rényi model
    - Degree conditional Bernoulli-Poisson model
    - Small Word model
    - Preferential attachment - Scale free model
- Differences among models
    - giant component
    - average degree
    - diameter
    - average distance
    - clustering / transitivity
    - centralization
- Monte Carlo simulations and fitting to real networks

- Collaboration networks - Erdős number
- Animations - PajekToSVGAnim
- Communication networks
- Food webs
- Protein-protein interaction networks