

Analysis of Czech actors' cooperation network using Gephi



Josef Daňa, Jaroslav Ráček

Department of Computer Systems and Communications,
Faculty of Informatics,
Masaryk University Brno

Presentation outline

- Introduction – network analysis
- Complex networks
- Gephi tool
- Czech actors' cooperation network
- Discussion

Network analysis



Disease spreading



Innovation & Inspiration

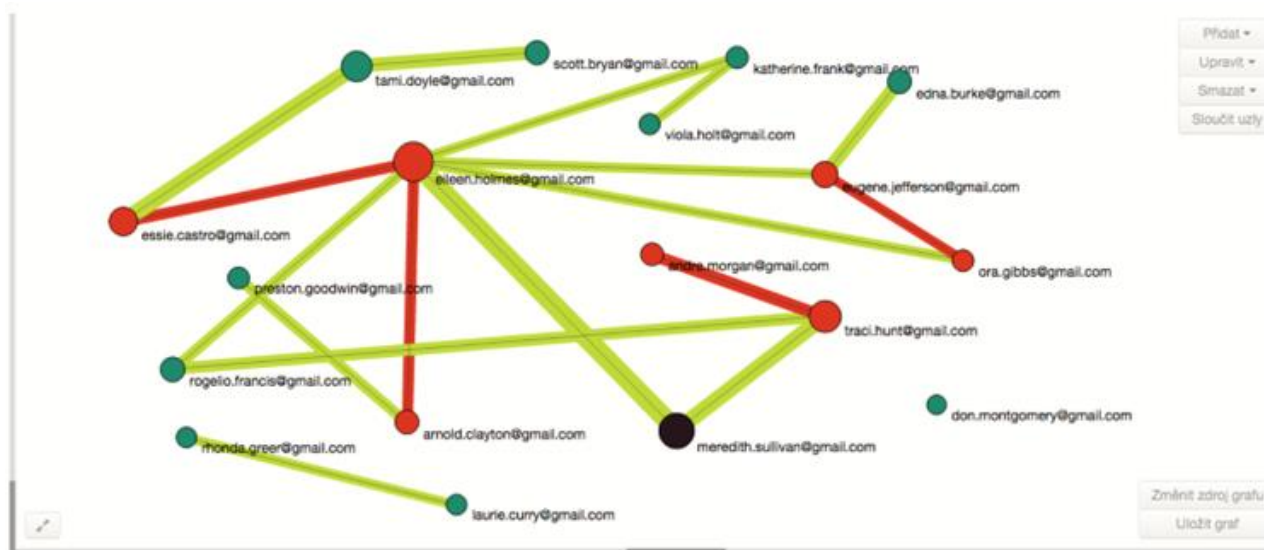


Business cooperation

Network analysis

- Any interaction between non-trivial amount of agents can be perceived as network
- Availability of good data?
 - Structured or semistructured data
 - Co-authorship of papers
 - Cooperation in movies
 - Sport matches records
 - Brain MRI
 - Trade networks
 - ...

Mining unstructured data



RÁČEK, Jaroslav, DAŇA, Josef a FRIŠTIK, Matúš. (2015). Personal Relationships Identification in Unstructured data.

Network analysis

- Revealing the network itself is useful for determining the interaction
- However, it does not tell us much about
 - Importance of nodes
 - Overall structure of network
 - Formation of clusters
 - Processes on the network

→ **Complex networks**

Complex networks

- Network science – interdisciplinary scientific field originating in Chaos and Complexity theory
- Multilayer network of dynamic interaction between many agents
 - Self-organization, emerging phenomena, small-world, self-similarity (fractal structure), etc.

Complex networks



Ant colonies

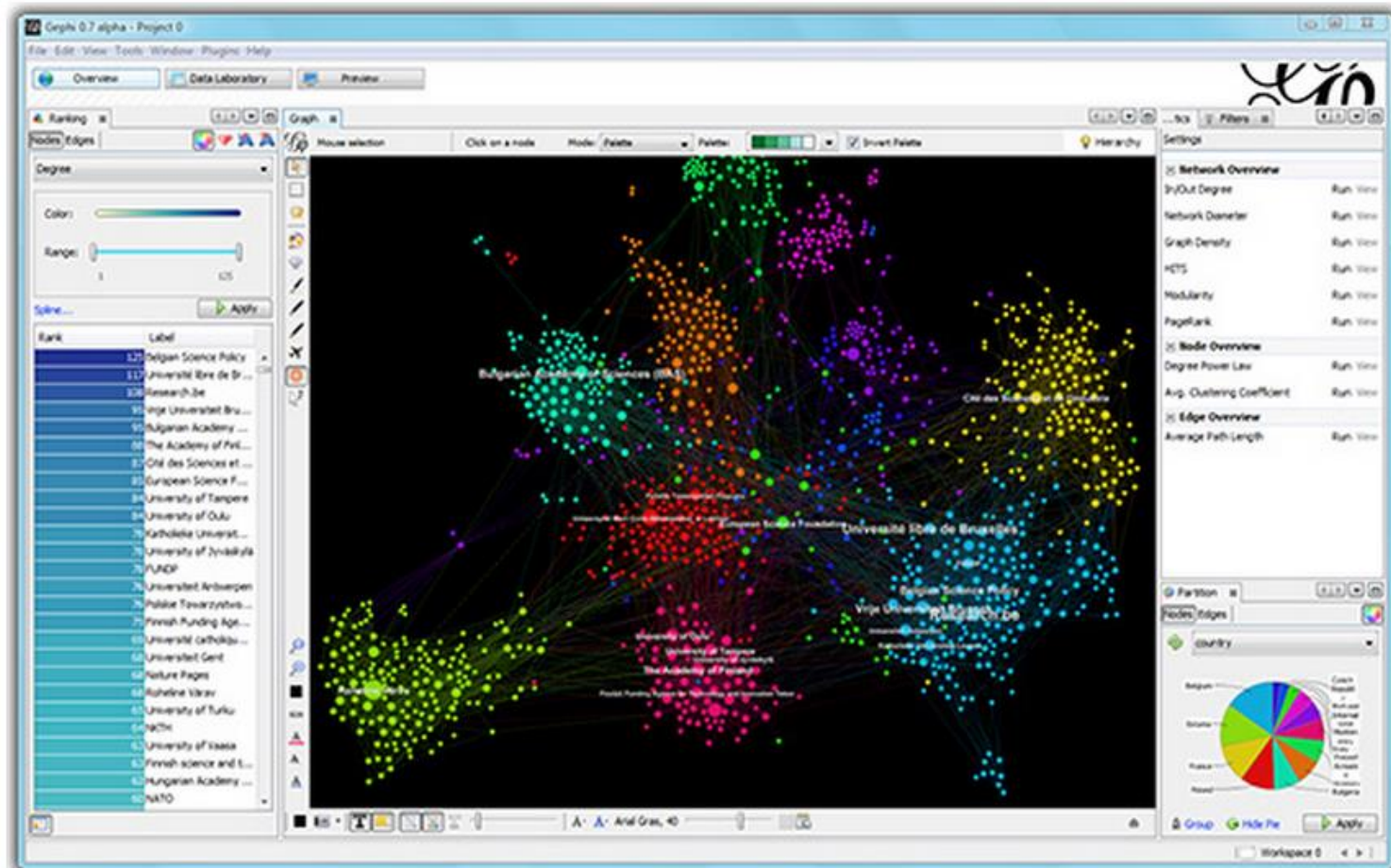


Neural networks



Transportation networks

Gephi tool

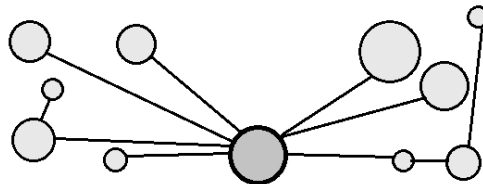


Gephi tool

- Open-source tool for analysing complex networks
 - Capable of visualising temporal (evolving) networks
 - Many community developed algorithms
 - Measuring centralities
 - Network layout
 - Detecting clusters
 - ...
- example: Czech actors' cooperation network

Network centralities

- Degree
 - Number of connections to other nodes
 - Visualization: bigger node = higher degree centrality
- Betweenness
 - Measure of shortest path from one part of network to another
 - Visualization: darker hue = higher betweenness centrality



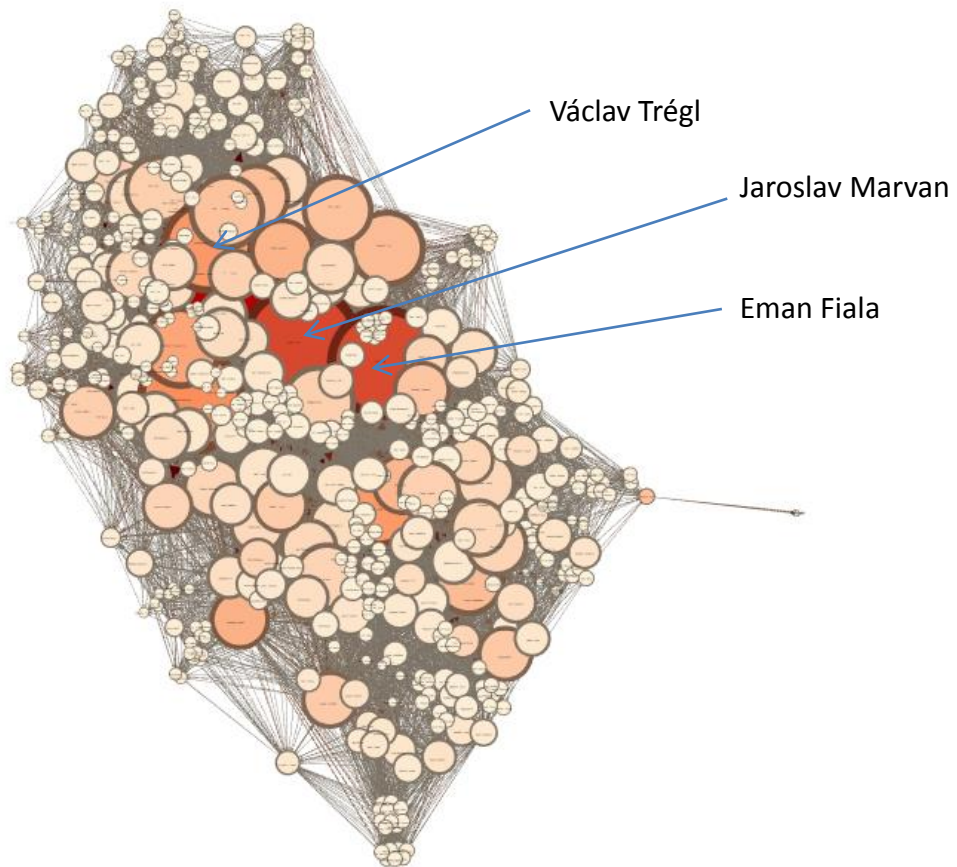
Czech actors' cooperation network

- Gathering data

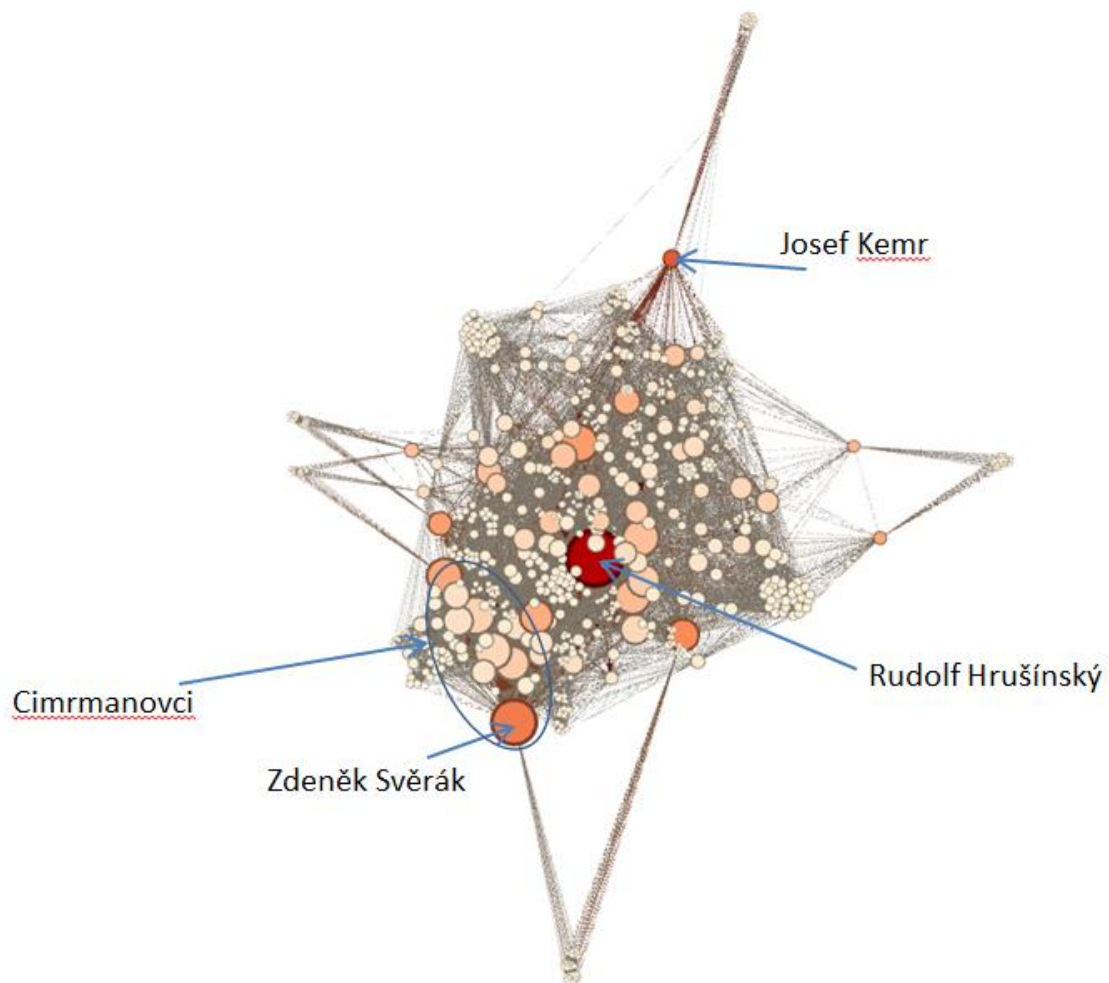


- 50 best rated Czech (Czechoslovak) movies from 3 periods:
 - 1930-1959
 - 1960-1989
 - 1990-2015
- Overall: $N = 1785$ nodes, $E = 43453$ links

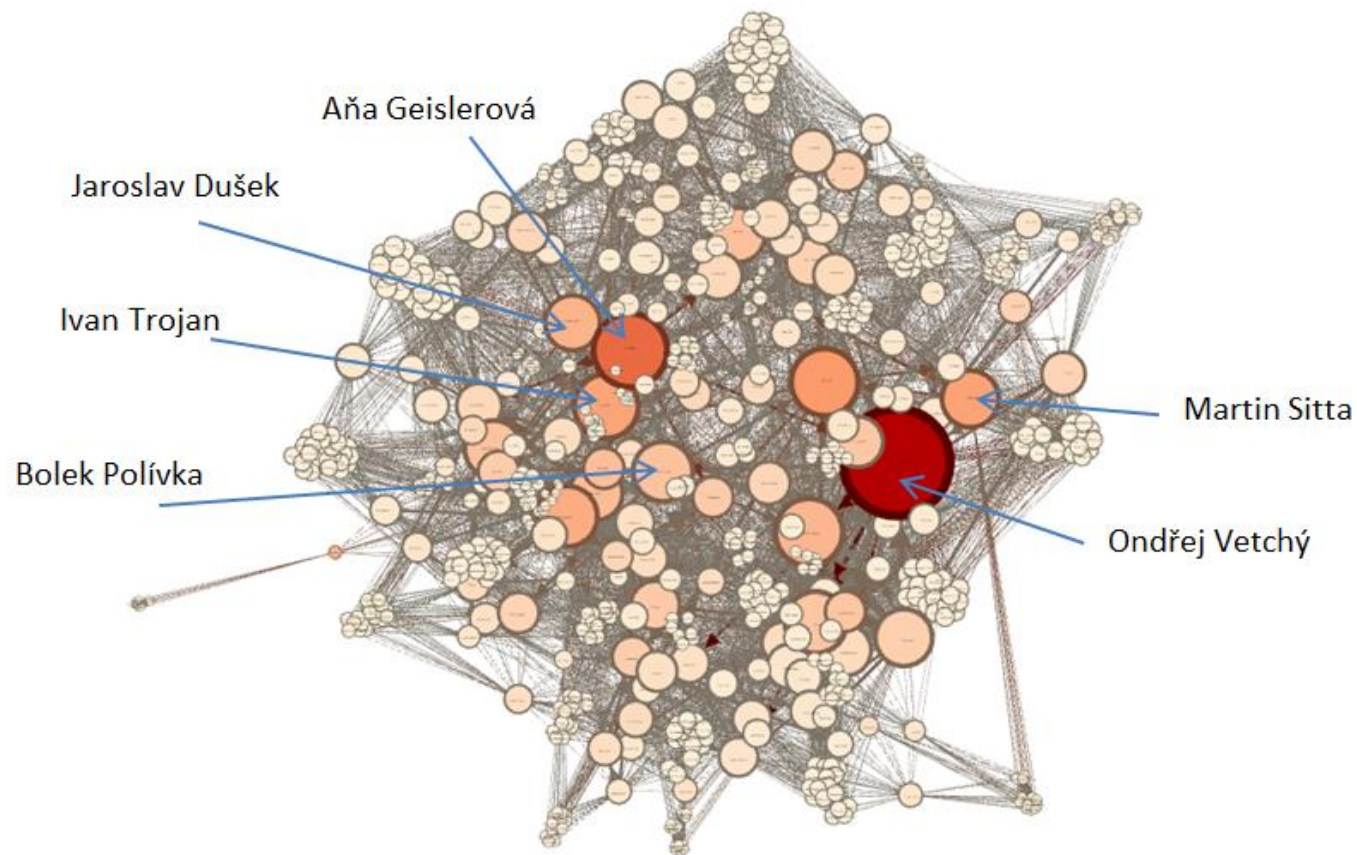
1930-1959



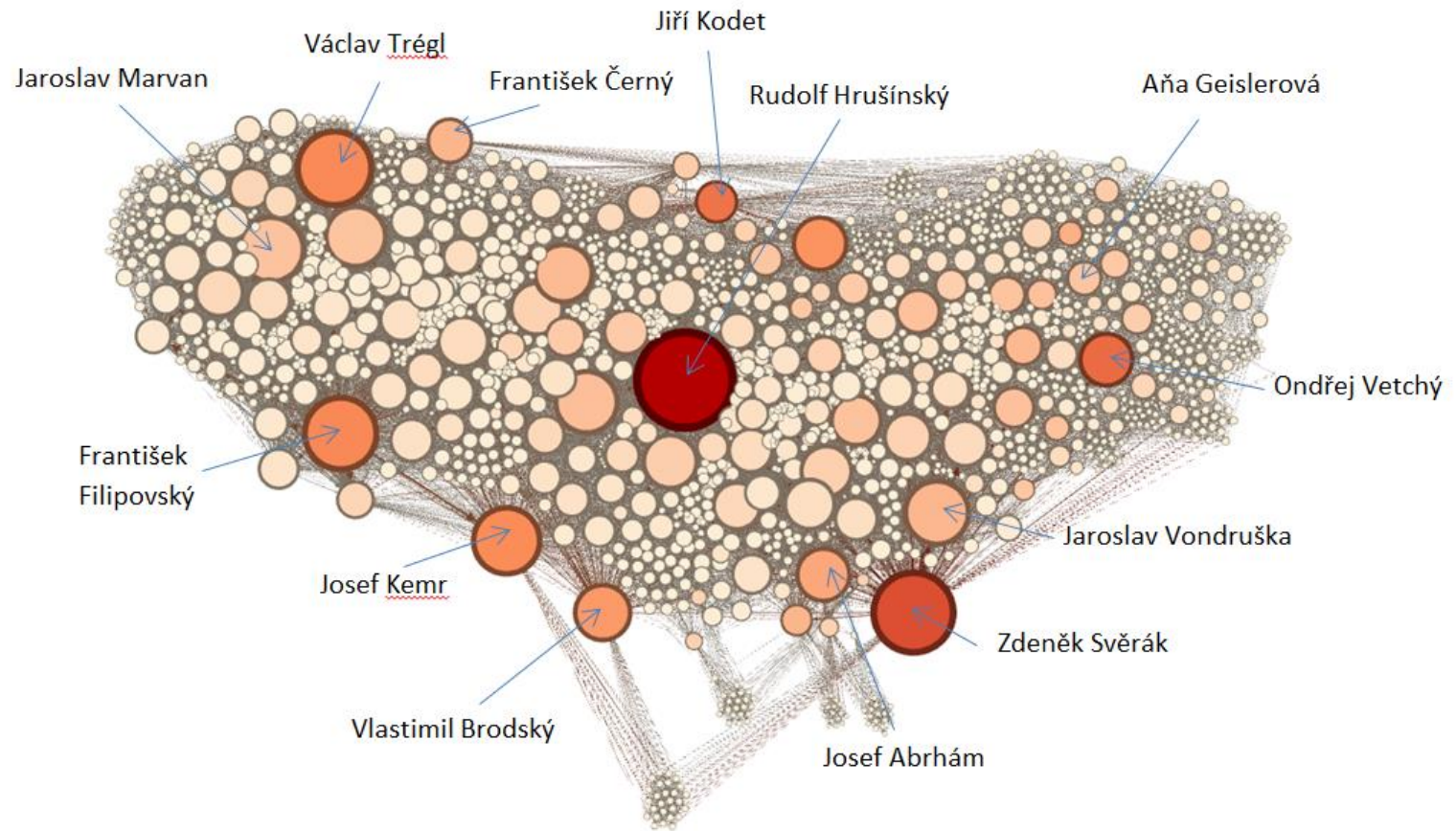
1959-1989



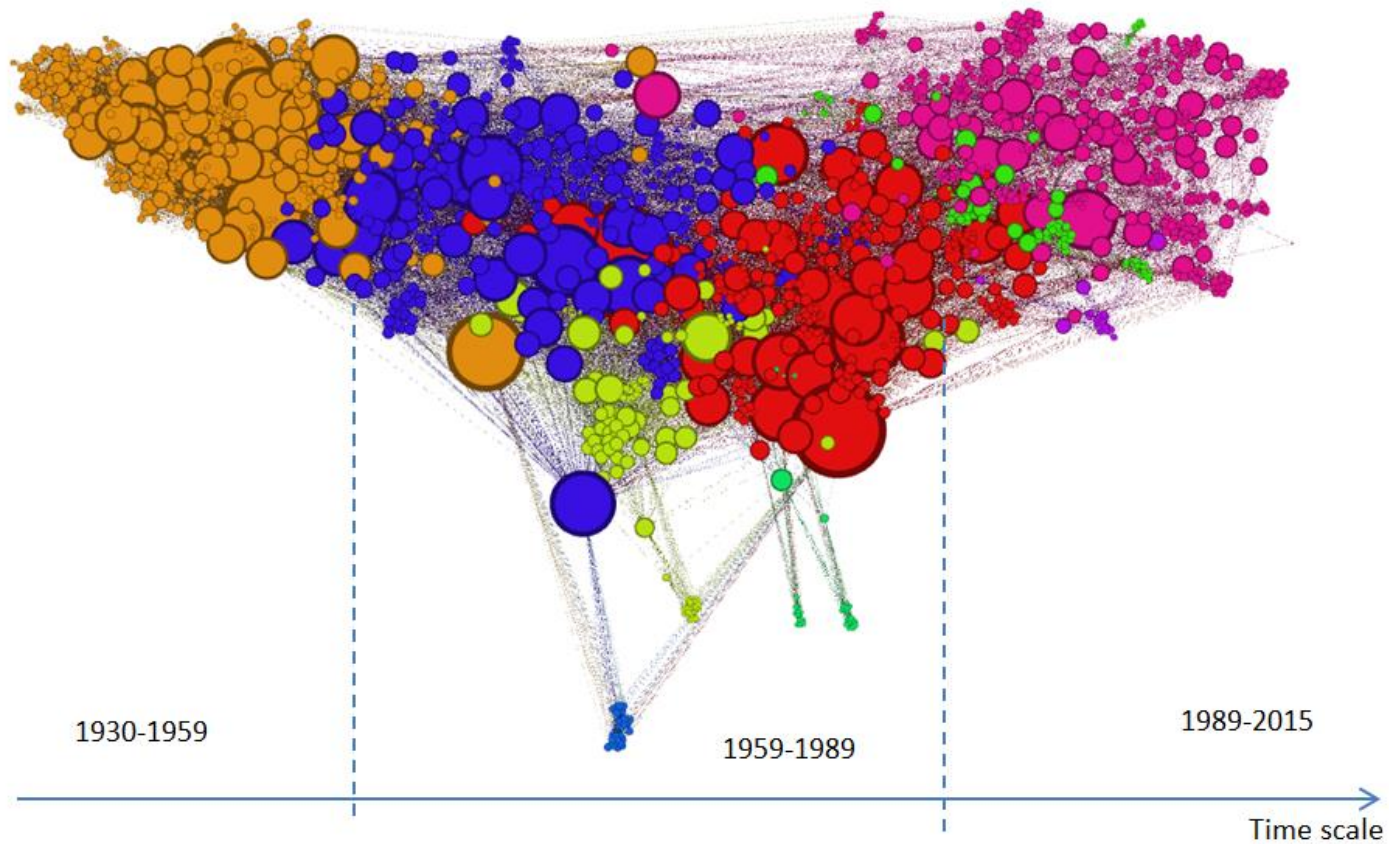
1989-2015



1930-2015



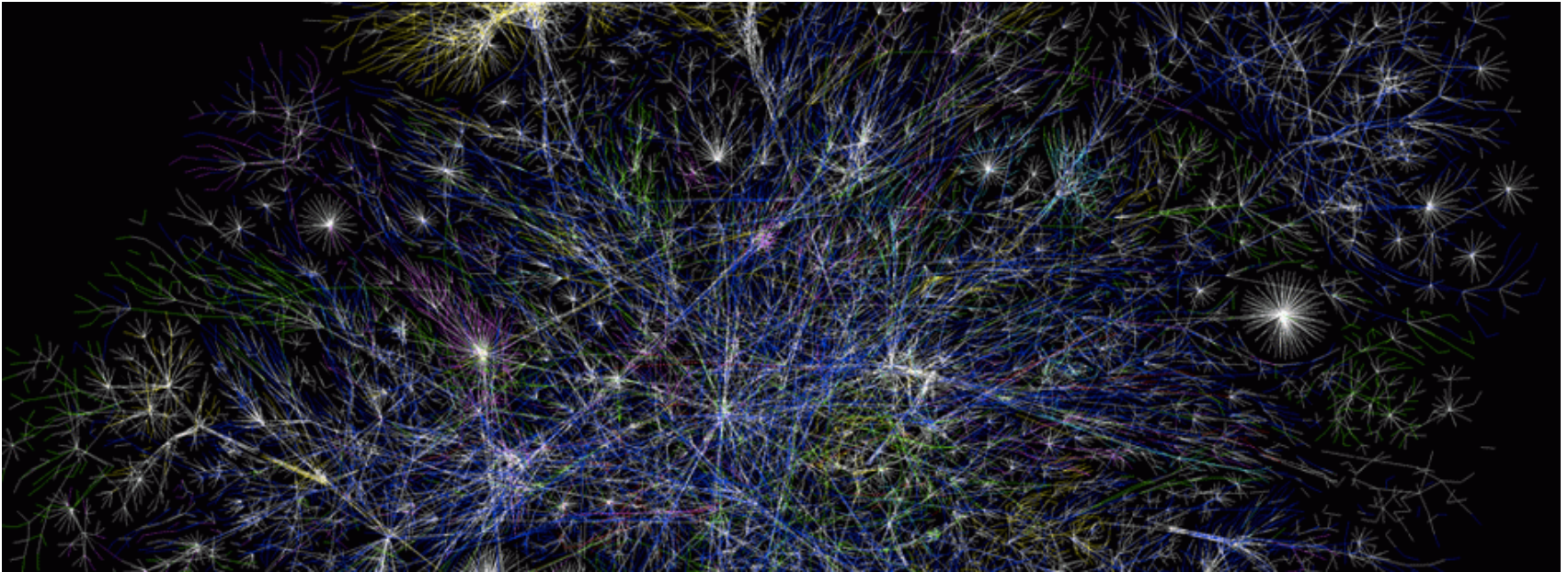
1930-2015 - Modules



Discussion

- Complex networks analysis may provide additional insight into structure and functioning of dynamic interacting systems
- Domain non-specific (biology, medicine, economy, IT, psychology...)
- Biggest issue is the availability of good data

Thank you



jdana@mail.muni.cz

racek@mail.muni.cz