

Mining Big Corporate Network Data

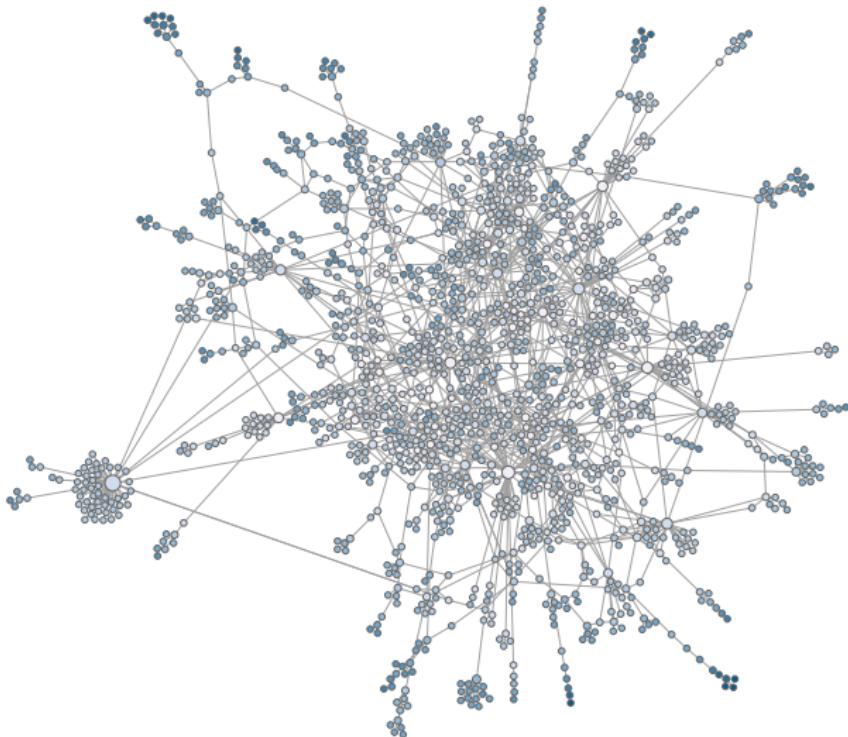
Frank Takes

LIACS — Universiteit Leiden

Lunchlezing
23 maart 2016

Network data

Network data

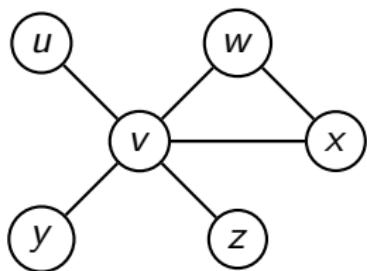


Networks

- **Network** (graph) consisting of:
 - **Nodes** (vertices/points/objects/entities): data objects
 - **Edges** (links/lines/interaction/ties): relationships between objects

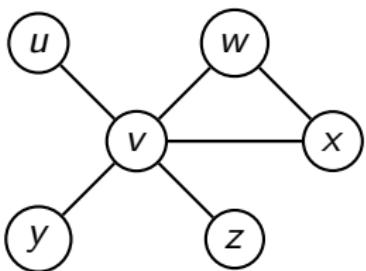
Networks

- **Network** (graph) consisting of:
 - **Nodes** (vertices/points/objects/entities): data objects
 - **Edges** (links/lines/interaction/ties): relationships between objects



Networks

- **Network** (graph) consisting of:
 - **Nodes** (vertices/points/objects/entities): data objects
 - **Edges** (links/lines/interaction/ties): relationships between objects



- Directed vs. undirected
- Weighted vs. unweighted
- Multi-mode vs. one-mode
- Attributes & annotations

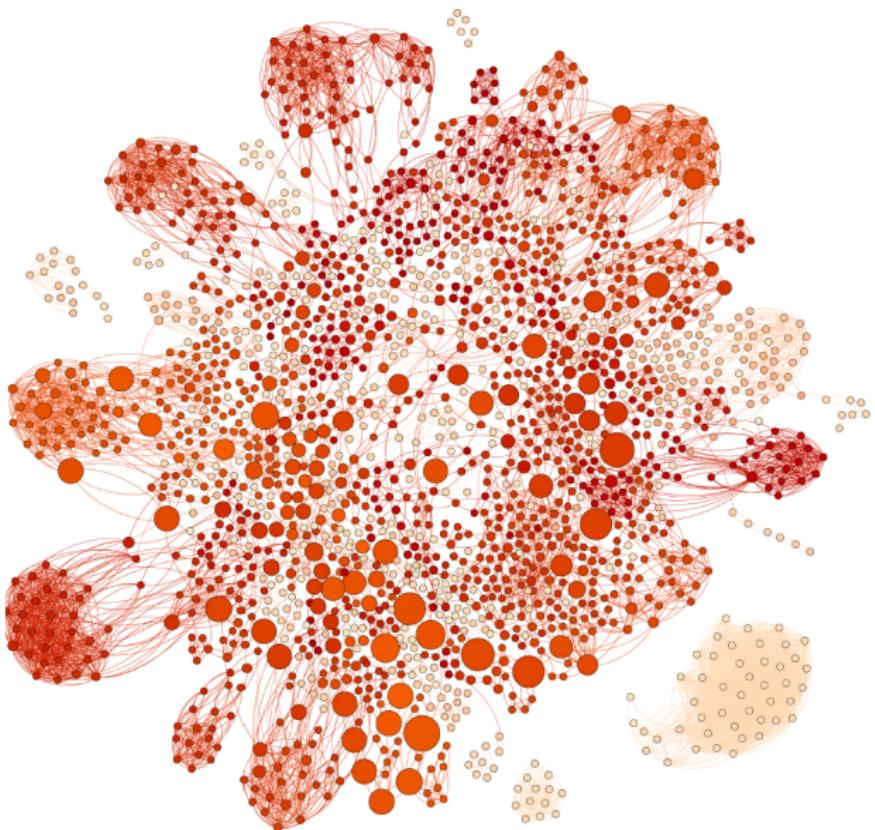


Figure : Sample of online social network with 1,876 nodes and 8,070 edges

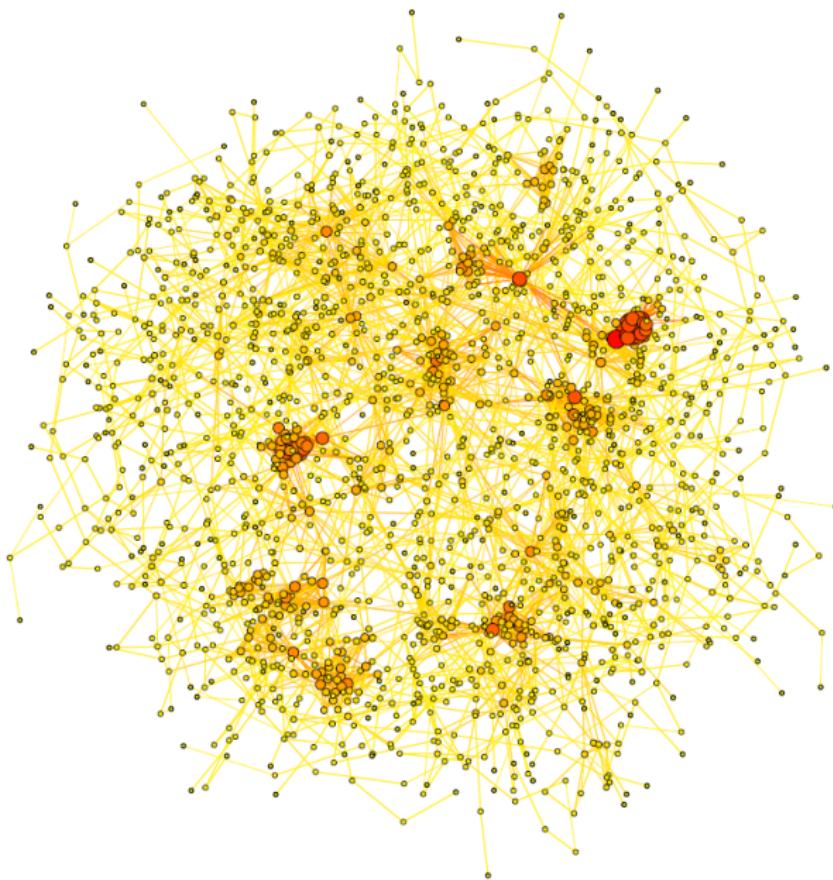


Figure : Sample of online social network with 1,876 nodes and 8,070 edges

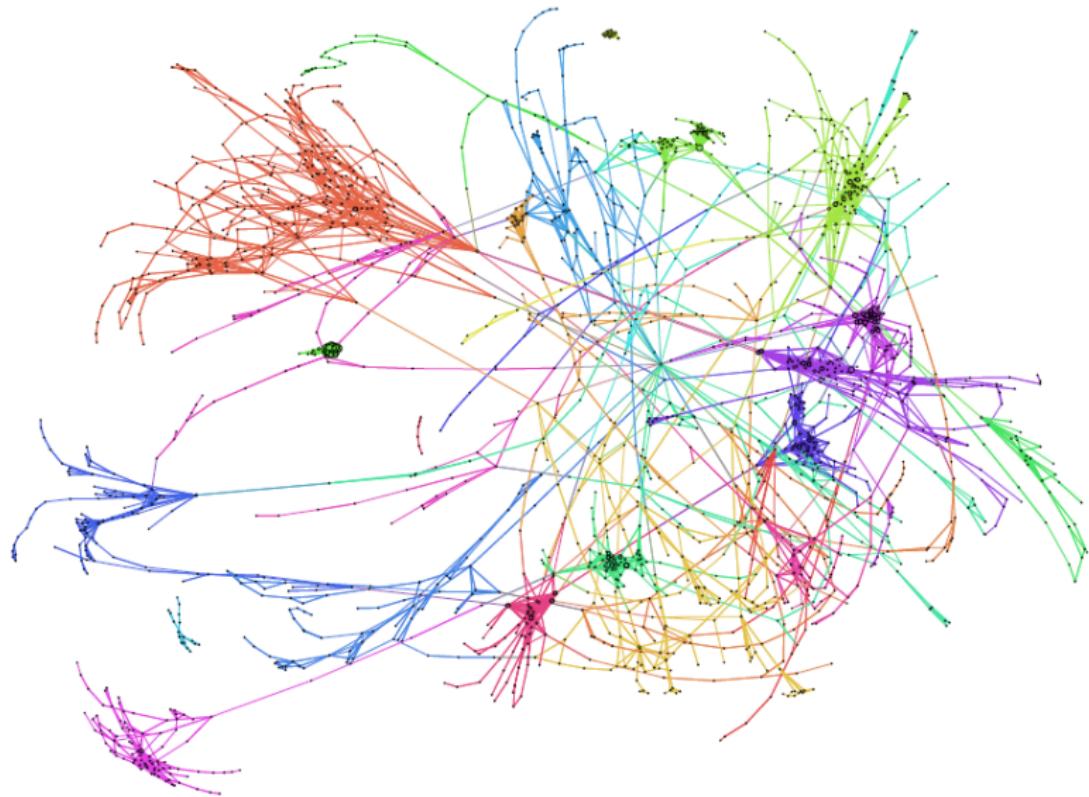


Figure : Sample of an online social network with 1,876 nodes and 8,070 edges

Network science

- Data analysis
- Data mining
- Data science

Network science

- Data analysis
- Data mining
- Data science
- Graph mining
- Network mining
- Social network analysis

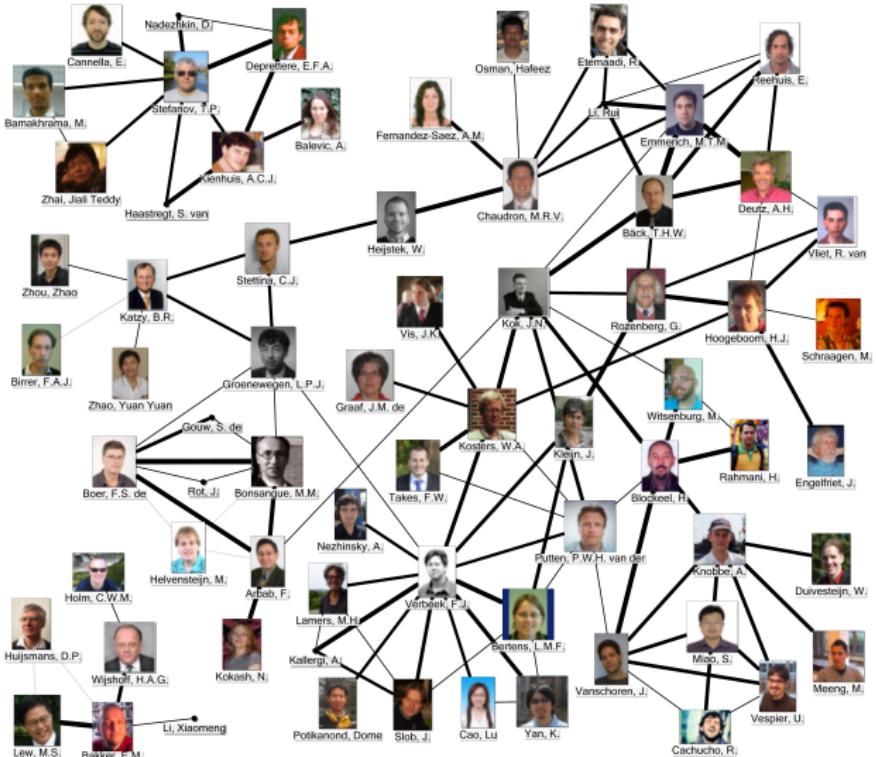
Network science

- Data analysis
- Data mining
- Data science
- Graph mining
- Network mining
- Social network analysis
- **Network science**

Real-world networks

- Online social networks
- Scientific citation network
- Research collaboration networks
- Webgraphs
- Biological networks
- Communication networks
- Trade networks
- Corporate networks
- ...

LIACS staff collaboration network



Corporate networks

- Nodes are firms/corporations/companies
- Edges
 - trading
 - borrowing/lending
 - ownership
 - **board interlocks**

Corporate networks

- Nodes are firms/corporations/companies
- Edges
 - trading
 - borrowing/lending
 - ownership
 - **board interlocks**
- Board interlock network / interlocking directorate network:
 - Node attributes: firm name and country
 - Edge attributes: shared director names and count (weight)

Board interlock network

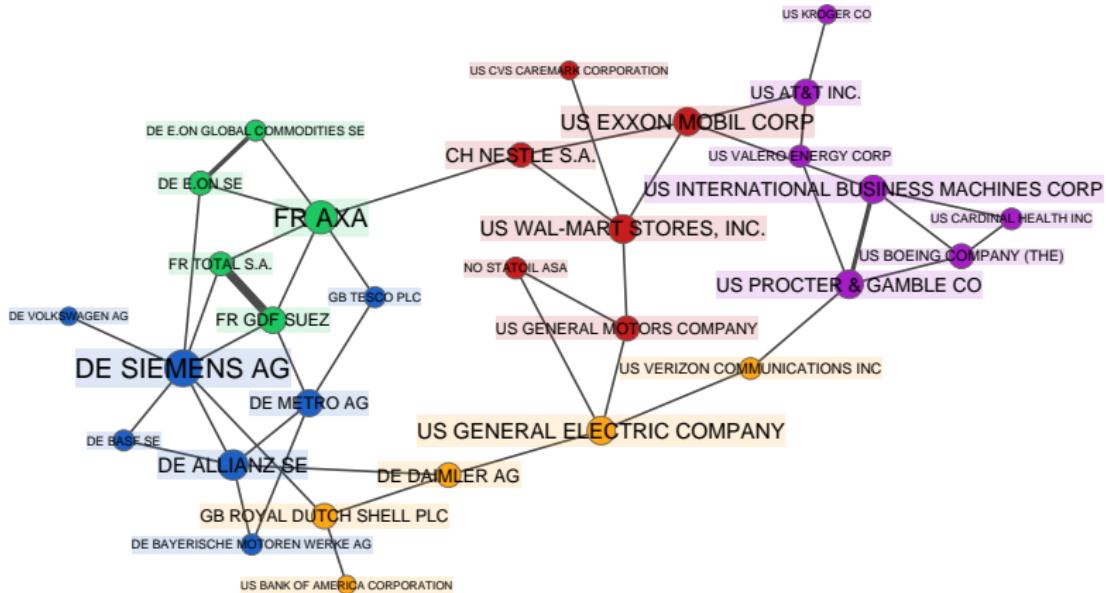


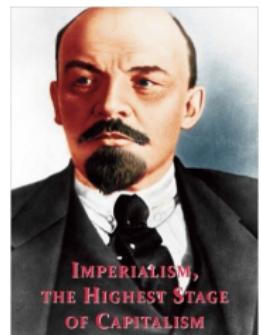
Figure : Board interlock network of 30 firms

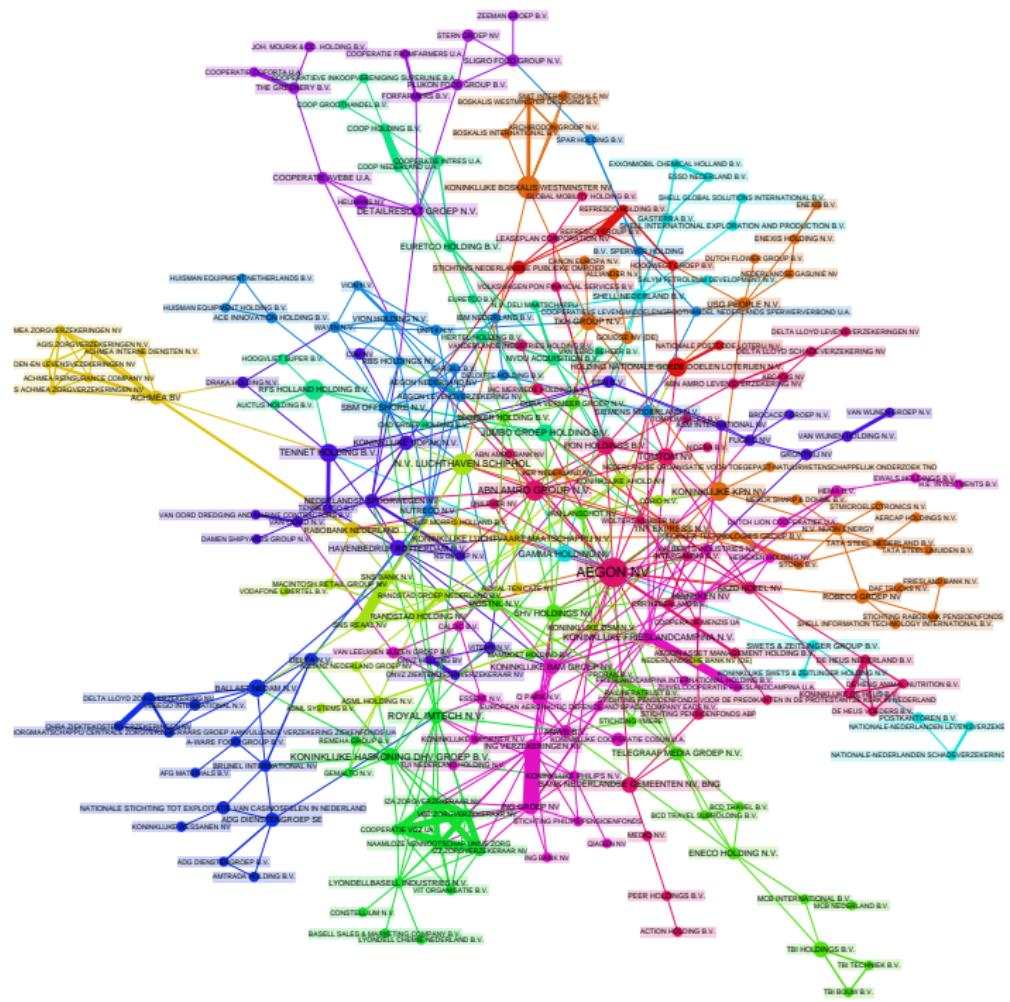
Board interlocks

- **Board interlock:** there is a relationship between firms because they share a board member or director

Board interlocks

- **Board interlock:** there is a relationship between firms because they share a board member or director
- Vladimir I. Lenin, *Imperialism, The Highest Stage of Capitalism*, 1916.
- "... a personal union, so to speak, is established between the banks and the biggest industrial and commercial enterprises, the merging of one with another through the acquisition of shares, through the appointment of bank directors to the Supervisory Boards (or Boards of Directors) of industrial and commercial enterprises, and vice versa."





Board interlocks

■ Causes of interlocks:

- Collusion
- Cooptation and monitoring
- Legitimacy
- Career advancement
- Social cohesion

■ Consequences of interlocks:

- Corporate control
- Economic performance
- Access to resources

Board interlocks

■ Causes of interlocks:

- Collusion
- Cooptation and monitoring
- Legitimacy
- Career advancement
- Social cohesion

■ Consequences of interlocks:

- Corporate control
- Economic performance
- Access to resources



M. Mizruchi, What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates, *Annual review of Sociology* 22: 271–298, 1996.

Board interlock network



Figure : Global corporate network: over 1,500,000 board interlocks

Topological properties

Global network	
Nodes	391,967
Edges	1,711,968
Density	$2.229 \cdot 10^{-5}$
Average degree	8.746
Components	55,616
Giant component	
Nodes	238,859 nodes (60.9%)
Edges	1,533,030 (89.5%)
Density	$5.374 \cdot 10^{-5}$
Average degree	12.83
Average distance	7.775
Radius	18
Diameter	34

Topological properties

Global network	
Nodes	391,967
Edges	1,711,968
Density	$2.229 \cdot 10^{-5}$
Average degree	8.746
Components	55,616
Giant component	
Nodes	238,859 nodes (60.9%)
Edges	1,533,030 (89.5%)
Density	$5.374 \cdot 10^{-5}$
Average degree	12.83
Average distance	7.775
Radius	18
Diameter	34

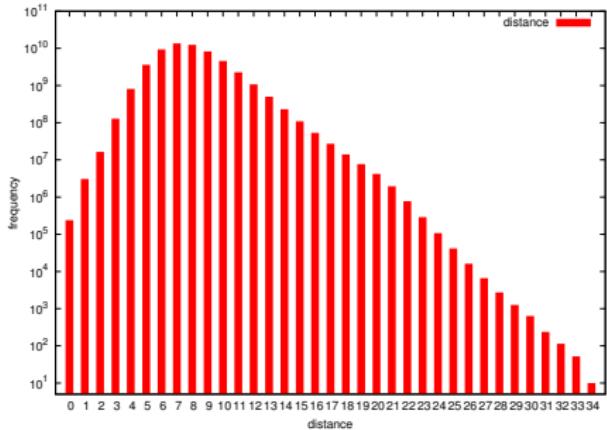


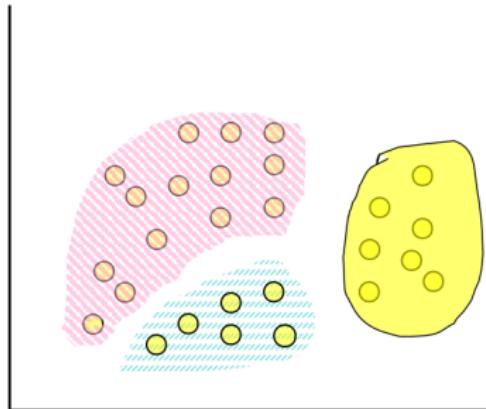
Figure : Distance distribution

Community detection

Based on: E.M. Heemskerk and F.W. Takes, The Corporate Elite Community Structure of Global Capitalism, in *New Political Economy* 21(1): 90–118, 2016. [dx.doi.org/10.1080/13563467.2015.1041483](https://doi.org/10.1080/13563467.2015.1041483)

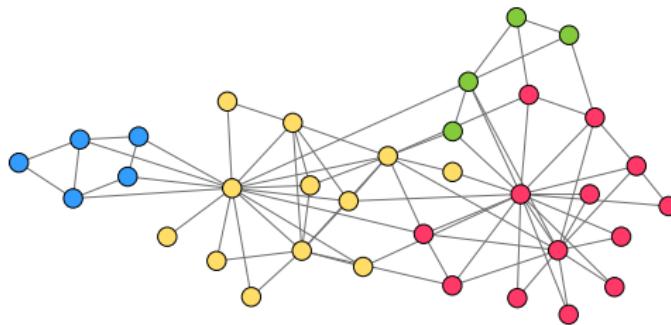
Related: clustering

- Data mining (on “flat” data)
- Unsupervised learning
- **Clustering:** group similar objects based on shared attributes



Community detection

- **Community:** set of nodes connected more strongly with each other than with the rest of the network
- Community detection algorithms:
 - Clique-based methods
 - Hierarchical clustering
 - Divisive algorithms (centrality-based)
 - **Modularity maximization** algorithms



Modularity

- **Modularity:** numerical value Q indicating the quality of a partitioning \mathcal{C} of a network's nodes

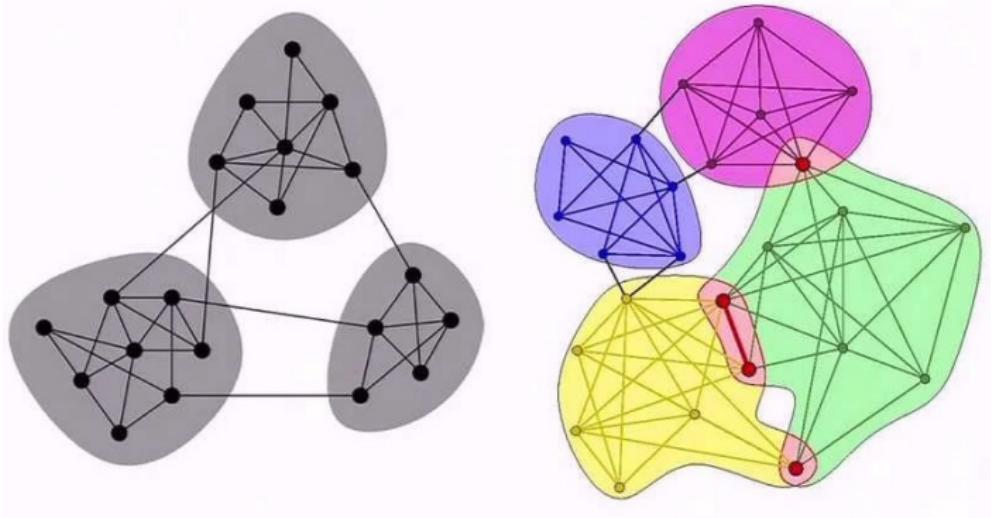
$$Q = \sum_{C \in \mathcal{C}} \left(\frac{|E(C)|}{m} - \left(\frac{\sum_{v \in C} \deg(v)}{2m} \right)^2 \right)$$

($E(C)$ are community C 's internal edges and $\deg(v)$ is v 's degree)

- **Community:** subset of nodes for which the fraction of links inside the community is higher than expected in a random network
- Optional resolution parameter r indicates how “tough” the algorithm should look for communities
- Algorithms maximize the modularity score Q given some r (using hill climbing, heuristics, genetic algorithms, and many more)

V.D. Blondel, J-L. Guillaume, R. Lambiotte and E. Lefebvre, Fast unfolding of communities in large networks, 2008.

Partitions vs. communities



J. Leskovec, Affiliation Network Models for Densely Overlapping Communities, MMDS 2012.

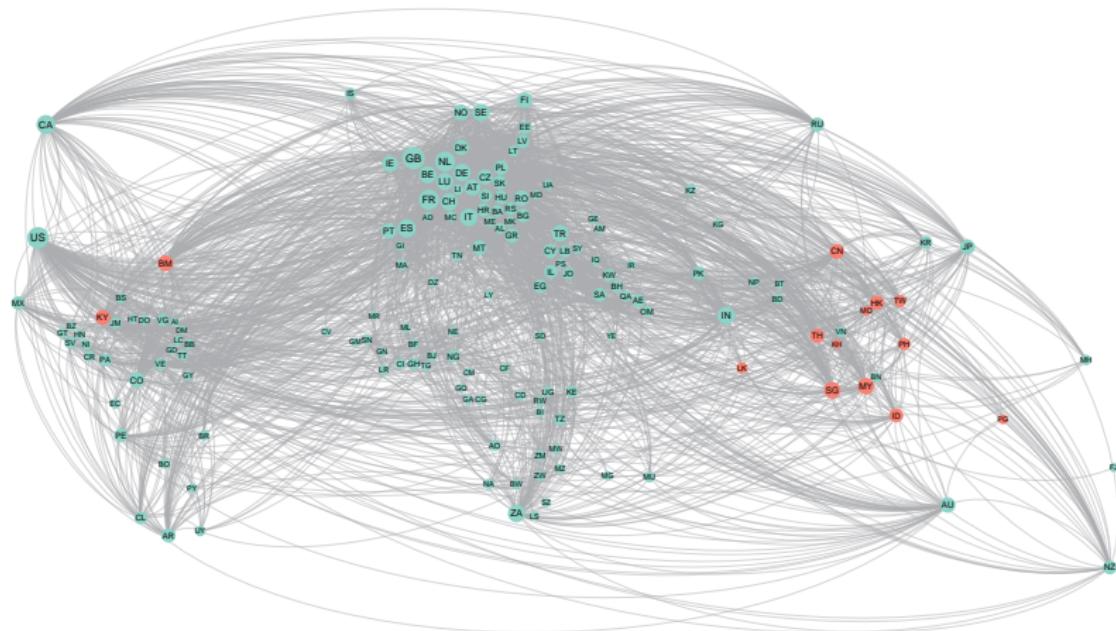
Corporate country network

- Aggregation of corporate network
- Nodes are **countries**
- Edges represent board interlocks between firms in these countries
- Layout on geographic Mercator map
- Apply community detection at different resolutions

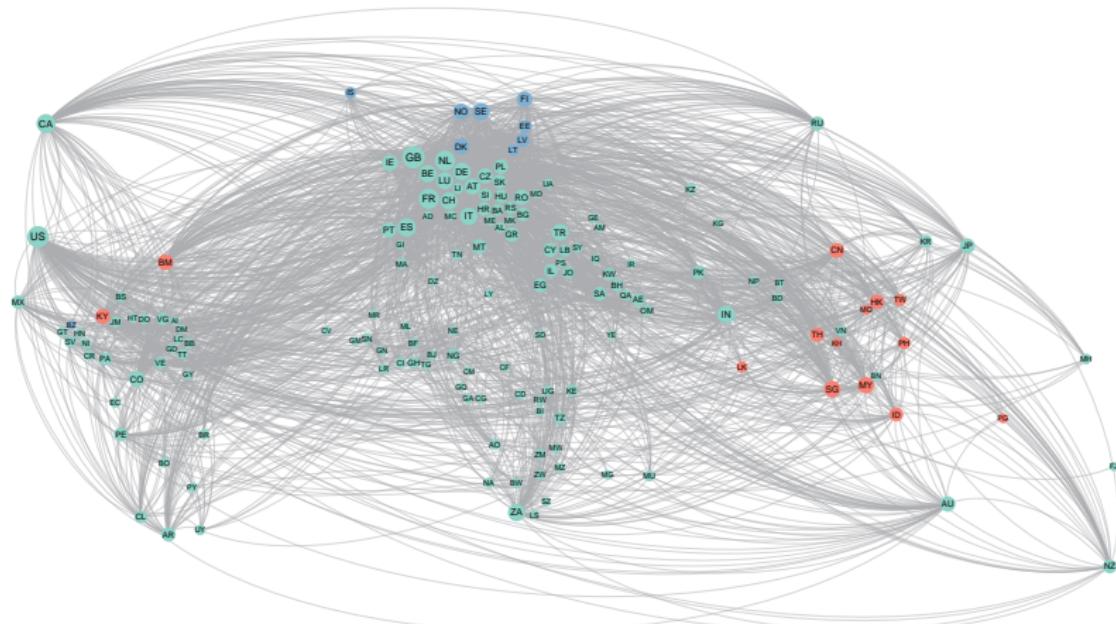
Country network communities



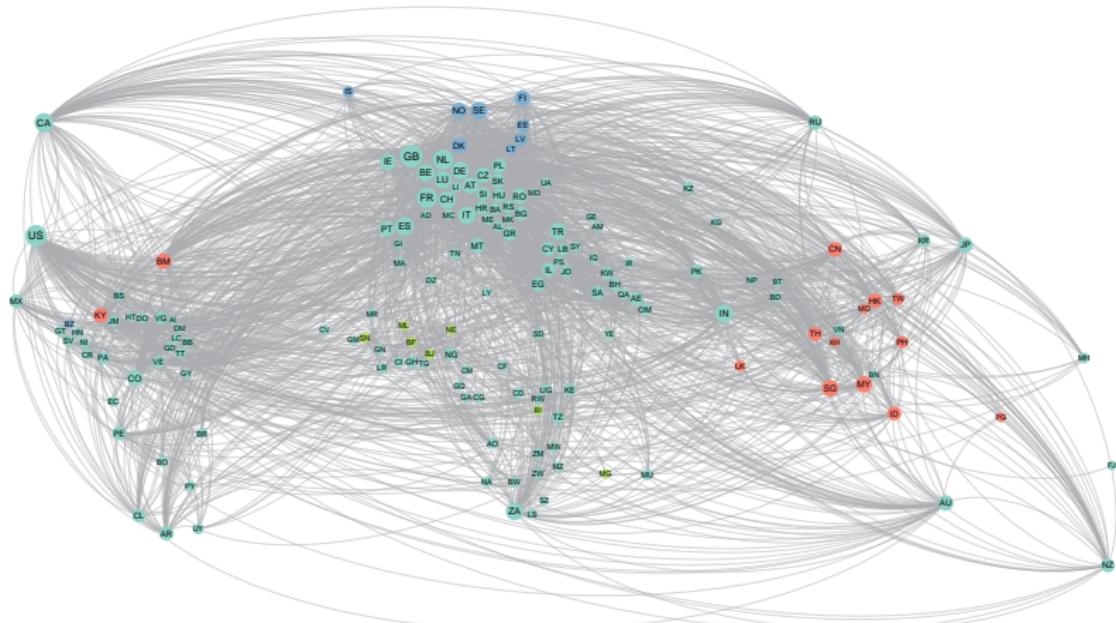
Communities, resolution = 2.0



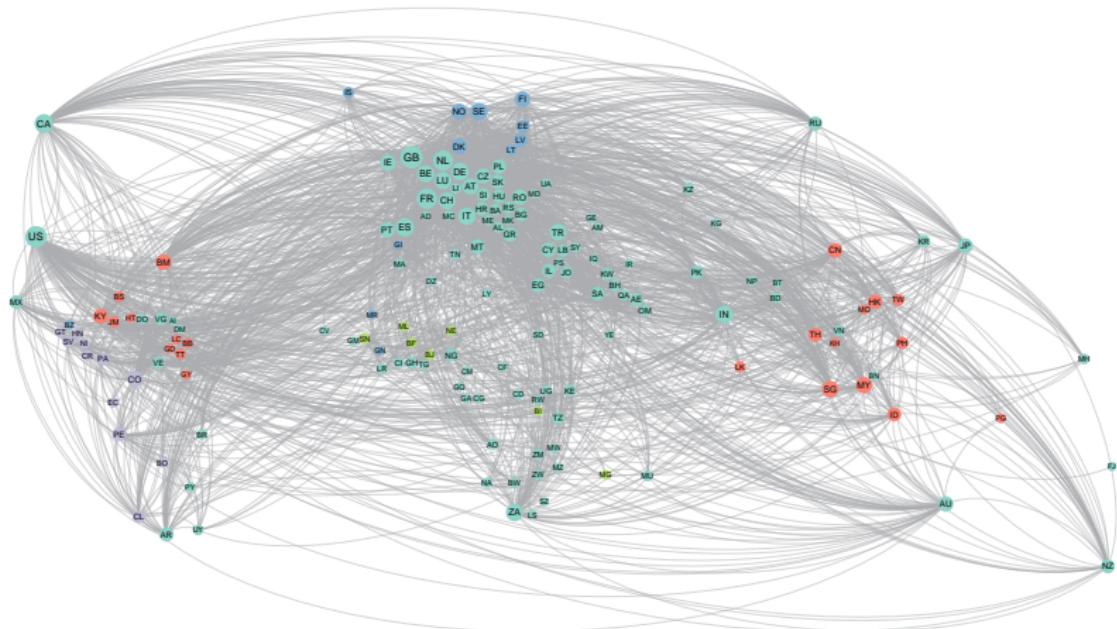
Communities, resolution = 1.75



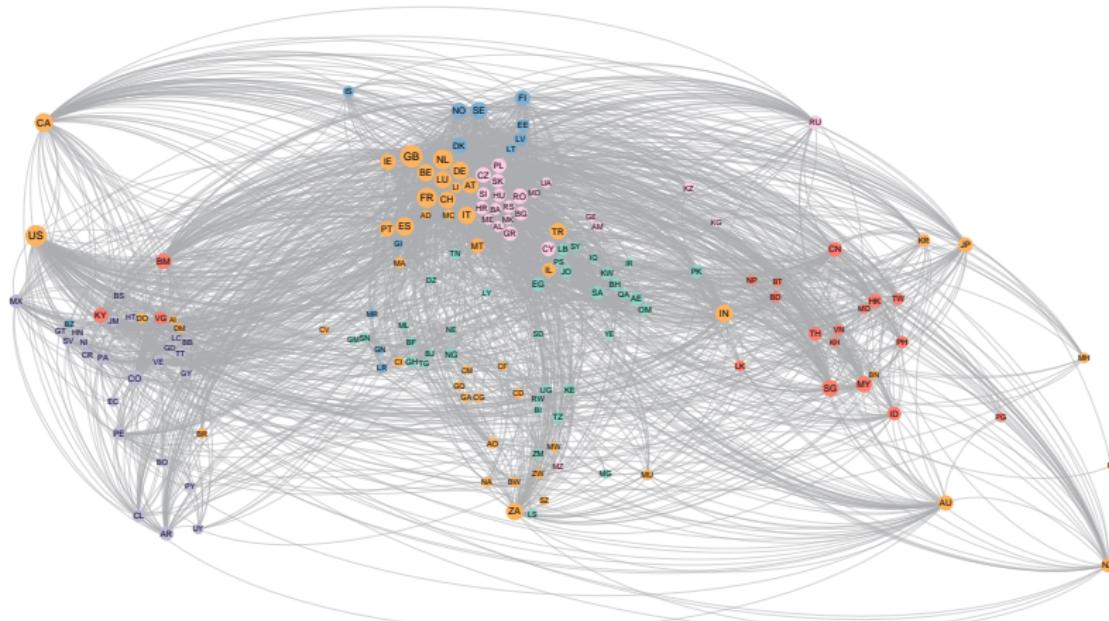
Communities, resolution = 1.55



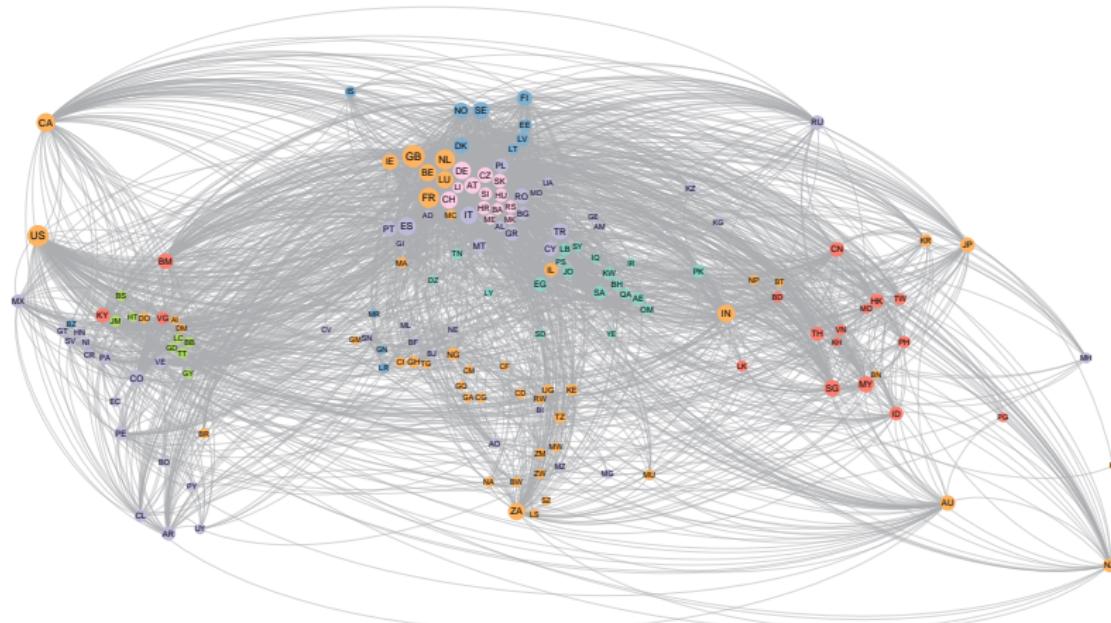
Communities, resolution = 1.5



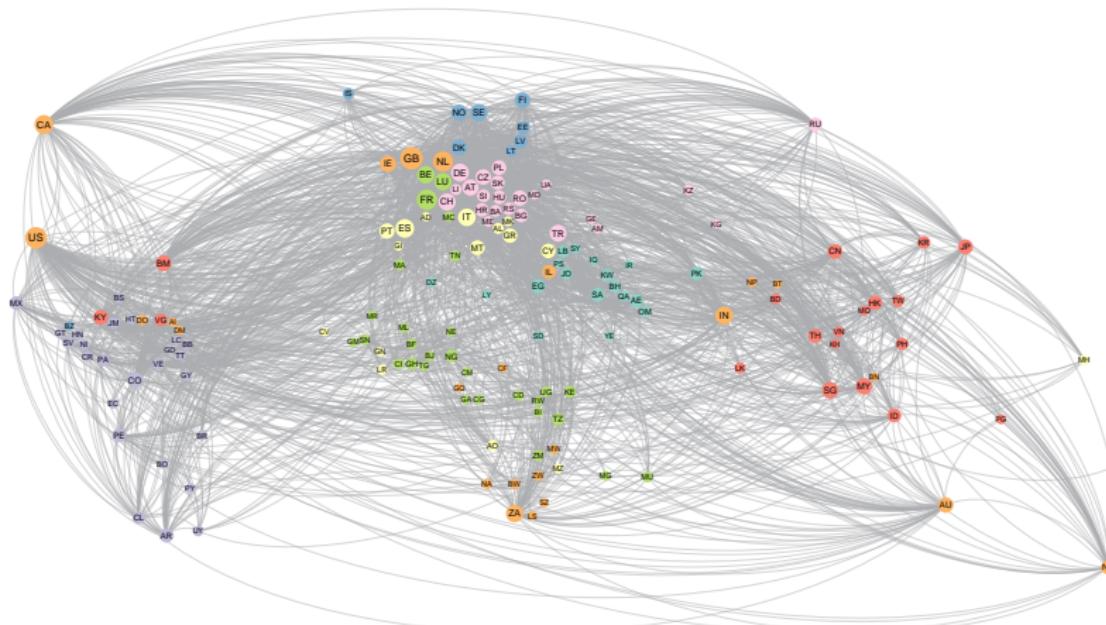
Communities, resolution = 1.0



Communities, resolution = 0.75



Communities, resolution = 0.5



Appearing communities per resolution

- 1 Eastern Asia cluster
- 2 Nordic and Baltic cluster
- 3 Former French colonies
- 4 Latin America
- 5 Eastern/Westers Europe
- 6 LatAm ties with Spain and Portugal
- 7 UK/US ties with NL, financial world

Appearing communities per resolution

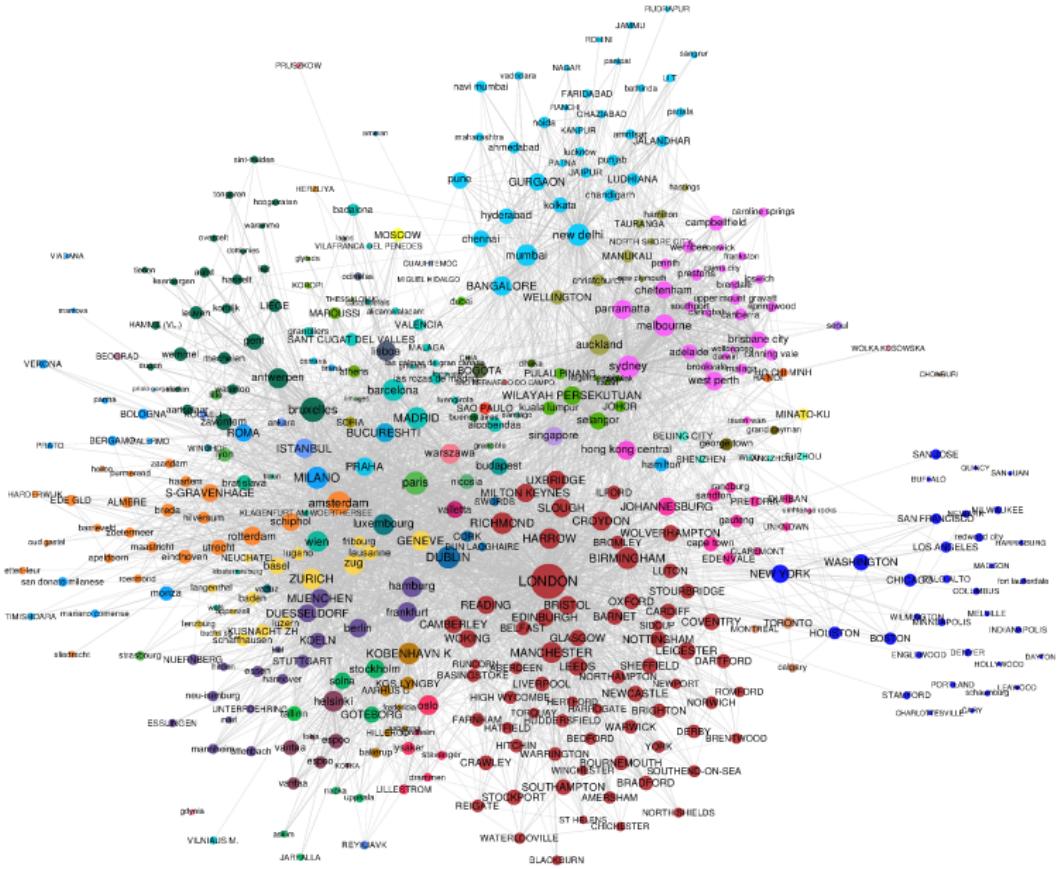
- 1 Eastern Asia cluster
- 2 Nordic and Baltic cluster
- 3 Former French colonies
- 4 Latin America
- 5 Eastern/Westers Europe
- 6 LatAm ties with Spain and Portugal
- 7 UK/US ties with NL, financial world

- Can we go into more detail?

Appearing communities per resolution

- 1 Eastern Asia cluster
- 2 Nordic and Baltic cluster
- 3 Former French colonies
- 4 Latin America
- 5 Eastern/Westers Europe
- 6 LatAm ties with Spain and Portugal
- 7 UK/US ties with NL, financial world

- Can we go into more detail? **City networks** (ongoing research)



City network communities





Conclusion

- **Network science** is a new way of studying data, analyzing systems of interaction rather than sets of objects
- **Corporate networks** provide interesting insight in corporate power and control across the globe
- **Network topology** of the global board interlock network shows that we live in a “small” corporate world
- **Community detection** is able to mine patterns in corporate networks. They can be interpreted based on regional, historical and cultural similarities as well as fiscal and administrative reasons

Thank you!

- Questions?

<http://franktakes.nl>

Thank you!

■ Questions?

<http://franktakes.nl>



Research in collaboration with the University of Amsterdam, see <http://corpnet.uva.nl>