Socioeconomic Networks

IV124

Josef Spurný & Eva Výtvarová

Faculty of Informatics, Masaryk University

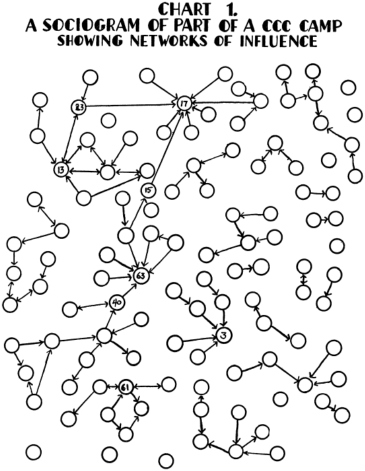
May 12, 2023

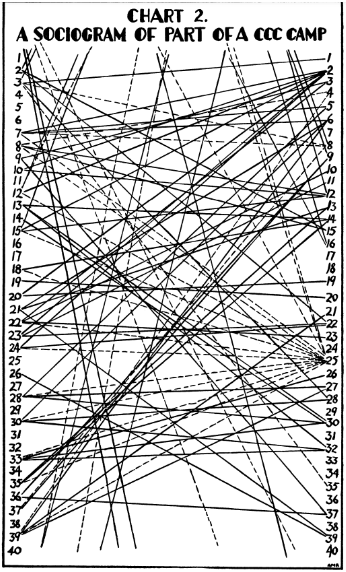
Social Networks

Numerous variations

- granularity of nodes: individuals, groups
- semantics of ties: friendship, kinship, formal, political, sexual, ...

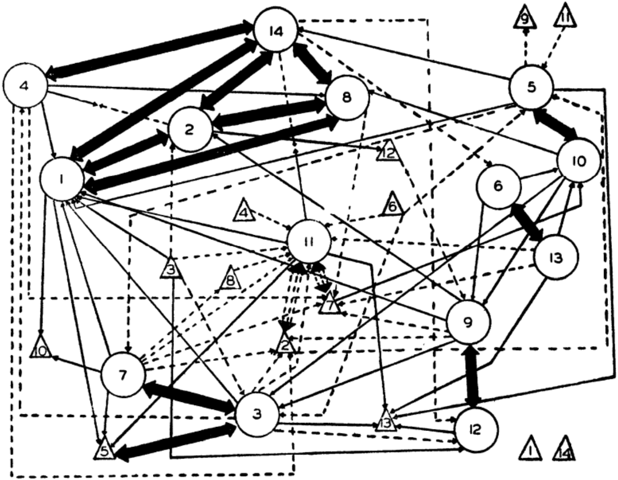
- rich history (since the 1930s): sociogram
- methodological resource for applications in other disciplines





[1]

[1] Murphy, A. J. (1941). A Study of the Leadership Process



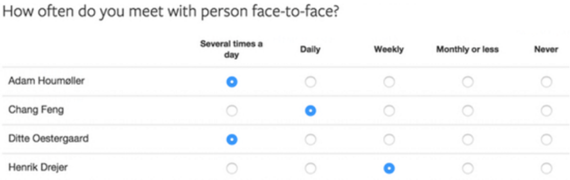
SOCIOGRAM I.—Positive choices (unbroken lines) and rejections (broken lines) made by pupils when asked which four persons they would like to have sit near them in school. Circles represent girls; triangles, boys.

2

[2] Dahkle O. & Monahan T. (1949). Problems in the Application of Sociometry to Schools

Collecting Network Data – Pre-ICT

- observation, declared groups
- but mostly questionnaires


How often do you meet with person face-to-face?

	Several times a day	Daily	Weekly	Monthly or less	Never
Adam Houmøller	●	○	○	○	○
Chang Feng	○	●	○	○	○
Ditte Oestergaard	●	○	○	○	○
Henrik Drejer	○	○	●	○	○

[3]

Types of questions asked:

- who do you contact for help?
- which colleague you meet outside work?
- who disturbs you from work?

[3] Socilyzer.com

Collecting Network Data – Data Science Approach

Analyzing large quantities of interaction data from various resources.

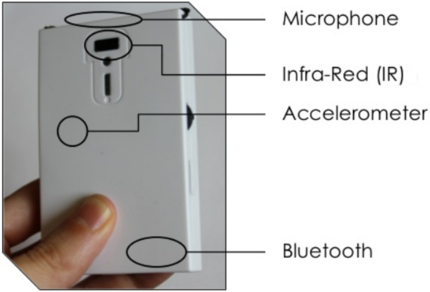
- records of electronic communication (phone, email)
- datamining collaborative systems (e.g. ticketing systems, GitHub, enterprise systems)
- harvesting social networks or social media (FB, Twitter)

Sociometric badges[4]

a wearable electronic device that captures social interaction data to gain insight into group dynamics, leadership roles, etc.

Types of collected data:

- speech patterns
- body movements
- proximity to others
- temporal aspects of interactions (e.g, Eda often responds to Eva, Viktor often interrupts Josef...).



[4] https://hd.media.mit.edu/badges/

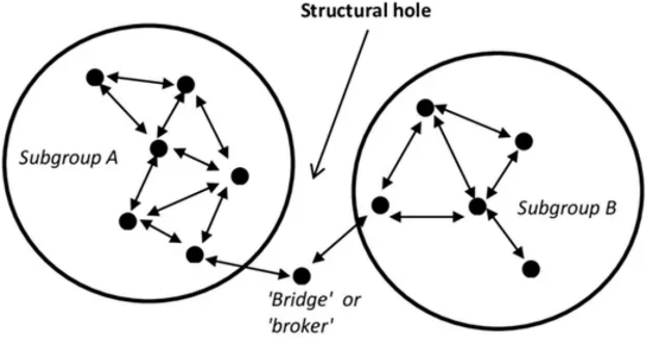
Metrics I.

actors – ties, UCINET

- **homophily (assortativity)**: ties with actors with similar vs. different attributes
- **multiplexity**: different types of ties
- **reciprocity**
- **network closure**: completeness of neighborhood triads, local transitivity
- **propinquity**: tendency of actors to form more ties with geographically closer actors

Metrics II.

- **bridge, centrality, density, distance, tie strength, clustering coefficient**
- **clique**: every two nodes in the clique have a tie; closed groups
- **structural holes**: absence of ties between two parts of the network, alternative concept of social capital

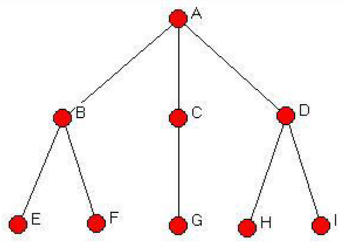


Block Modeling

- revealing **positions** (statuses) and **roles** (relationships between actors) in the network
  - → teacher vs. students
- based on network equivalence
- why do we care about equivalence?
  - → **equivalent nodes often share similarities** (attitudes, behavior...)
  - → **equivalent nodes are substitutable** (one QA specialist may be replaced with another when becoming ill, communication patterns remain the same...)
  - → **simplified representation** of the network: data reduction to structural patterns

Block Modeling – Equivalence

- **structural**: nodes with ties of the same intensity and to (from) the same nodes; interchangeable; equivalent nodes doesn't need to be connected to each other
  - → team leaders – consult with team, report to manager
- **regular**: same ties to other nodes in the same position; multiple ways to divide nodes into equivalent classes



Block Modeling – Similarity

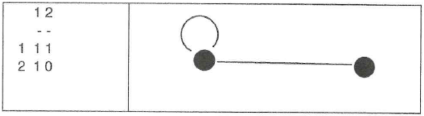
- node equivalence is a local property – we only need to know incoming and outgoing links of selected node
- perfect equivalence does not occur in practice $\rightarrow$ measures of similarity
  - correlation / Euclidean distance
  - network neighborhood, geodesic distances

Block Modeling – Procedure

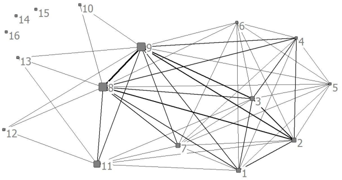
1. definition of equivalence
2. measure of equivalence of nodes and ties
3. **representation of equivalent classes**
4. adequacy of representation

Block Modeling – Representation of Equivalent Classes

- rearrangement of the adjacency matrix according to the node membership in positions
- equivalent nodes are grouped together – groups create **blocks** or all ones or all zeros
- then, the block can be reduced to single value (either one **1-blocks** or zero **0-blocks**) = smaller, so called **image matrix**
- **image graph** = visualization of the image matrix



Block Modeling – Example



Blokmodel procedury CONCOR se čtyřmi pozicemi

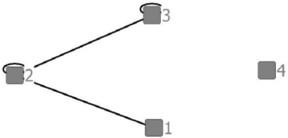
	12	13	9	11	8	7	10	4	5	2	1	3	6	14	15	16
12			1	1	2											
13			1	1	1											
9	1	1		3	10	3	1	3	1	5	3	5	2			
11	1	1	3		3	1				1	1	1				
8	2	1	10	3		3	1	3	2	5	3	4	2			
7			3	1	3			1	1	2	1	2	2			
10			1		1											
4			3		3	1			1	3	2	3	1			
5			1		2	1		1		1	1	1	1			
2			5	1	5	2		3	1		4	6	1			
1			3	1	3	1		2	1	4		4	1			
3			5	1	4	2		3	1	6	4		1			
6			2		2	2		1	1	1	1	1				
14																
15																
16																

Tabulka č. 2: Do bloků uspořádaná matice s vyznačenými pozicemi podle CONCOR

Image matice podle CONCOR

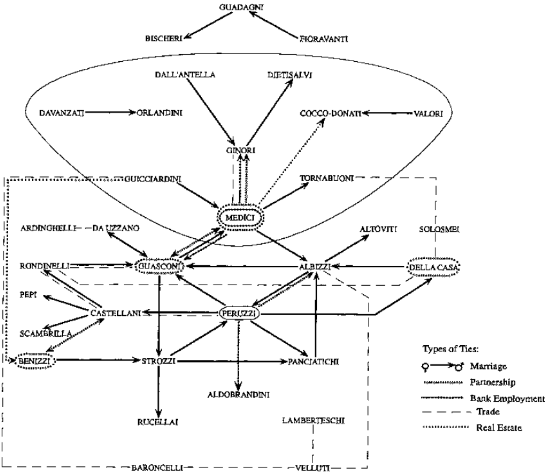
	1	2	3	4
1	0	1	0	0
2	1	1	1	0
3	0	1	1	0
4	0	0	0	0

Tabulka č. 5: Image matice podle CONCOR s kritériem maximální hodnoty (= 1)



Graf č. 3: Image graf čtyř pozic podle CONCOR

Block Modeling – Examples



[5]

[5] Padgett & Ansell. 1993. *Robust Action and the Rise of the Medici, 1400-1434*

Organizational Structure of Companies

Firm/organization as a social network[6]

- shift from hierarchical models to realistic descriptions
- based on real data (e.g., records of communication)
- questions about the influence of actors, network robustness, ...
- network efficiency (e.g., for information diffusion): direct economic impact
- question of social capital
- question of performance (individual/team/overall) with regard to network structure, centralities

[6] Dodds PS. et al. (2003). Information exchange and the robustness of organizational networks.

Online Collaboration Case Study

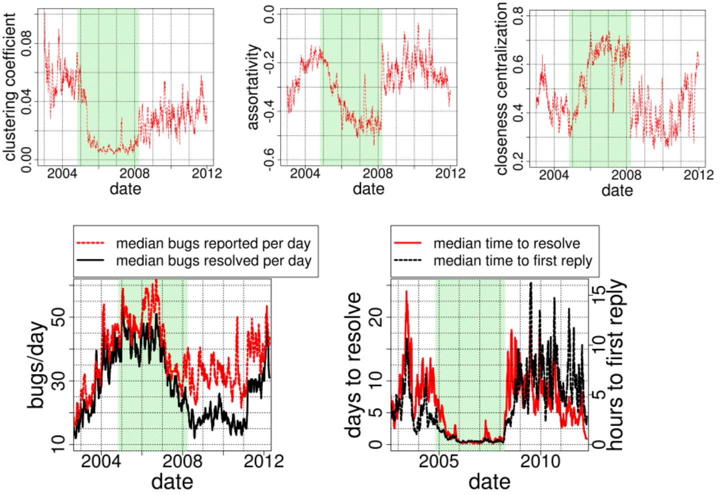
Zanetti et al. (2013). **The Rise and Fall of a Central Contributor**

Open-source community maintaining and updating a Linux distribution

(a) October 2004 (b) July 2006 (c) May 2008

Online Collaboration Case Study

Zanetti et al. (2013). **The Rise and Fall of a Central Contributor**



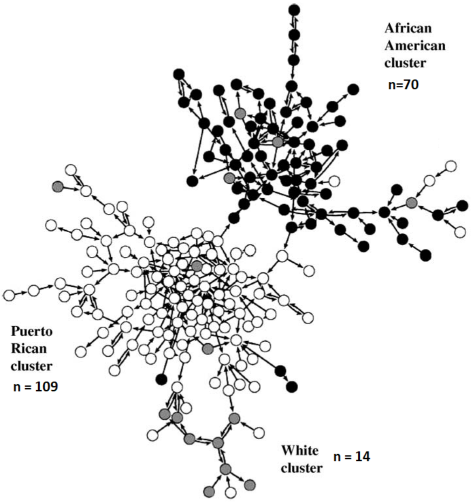
Social Groups: Disease Spread

- Biological process occurring on a social network
- Sexually transmitted diseases (through sexual partners), needle-sharing among drug users, ...

Goal:

- Creating realistic models for disease propagation in the population
- Providing information for effective prevention or stopping the spread[7]

[7] Weeks et al. (2002). Social Networks of Drug Users in High-Risk Sites: Finding the Connections.


African
American
cluster
n=70
Puerto
Rican
cluster
n = 109
White
cluster
n = 14

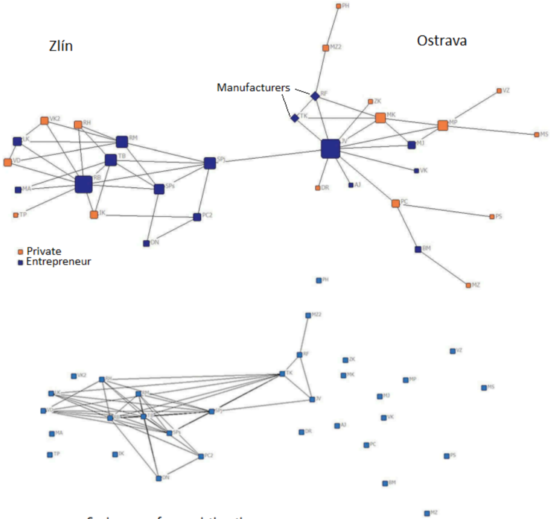
Criminal Networks

Basic dimensions of criminal networks[8]

- Trust – induced by reciprocity, often based on pre-existing ties (kinship)
- Material transfer – drugs, counterfeit goods, illegal profit... (training, planning & logistics...)
- Collaboration & Coordination – be in the same place at the same time

[8] Diviak. (2020). Criminal networks: actors, mechanisms, and structures

Criminal Networks - Methanol Affair[9]



Sociogram of pre-existing ties.

[9] Diviak, Dijkstra, Snijders. (2019). Poisonous connections.

SNA in Sports

Performance of **basketball** teams in NBA[10] depends on key "star" players:

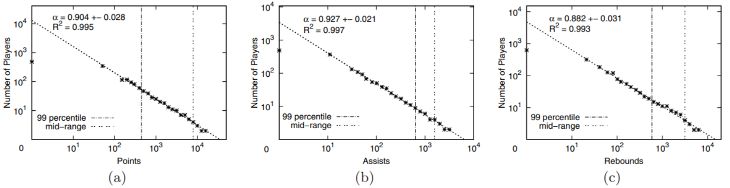


Figure 1: Distribution of points, assists and rebounds of NBA players.

On the contrary, in **football**[11], team performance is linked to intensity of interactions and low network centralization.

[10] Vaz de Melo et al. (2008). Can Complex Network Metrics Predict the Behavior of NBA Teams?

[11] Grund. (2012). Network structure and team performance: The case of English Premier League soccer teams

Economic Networks

Possible applications:

- financial systems
- business and supply chains
- marketing – contextual recommendations
- enterprise management – social networks
- labor market as a social network
- games on the network

Knowledge sharing networks

- Knowledge sharing is a key factor in organizational performance
- Social links are channels transporting information and knowledge

How to increase knowledge flow?[12]

- For knowledge diffusion, target central employees even though they are not central in offical hierarchy
- Deploy method for connecting employees who face similar challenges
- Create formal and informal shortcuts in network to connect cognitively distant communities
- Identify and bridge structural holes

[12] Daňa, Caputo & Ráček. (2020). Complex network analysis for knowledge management and organizational intelligence.

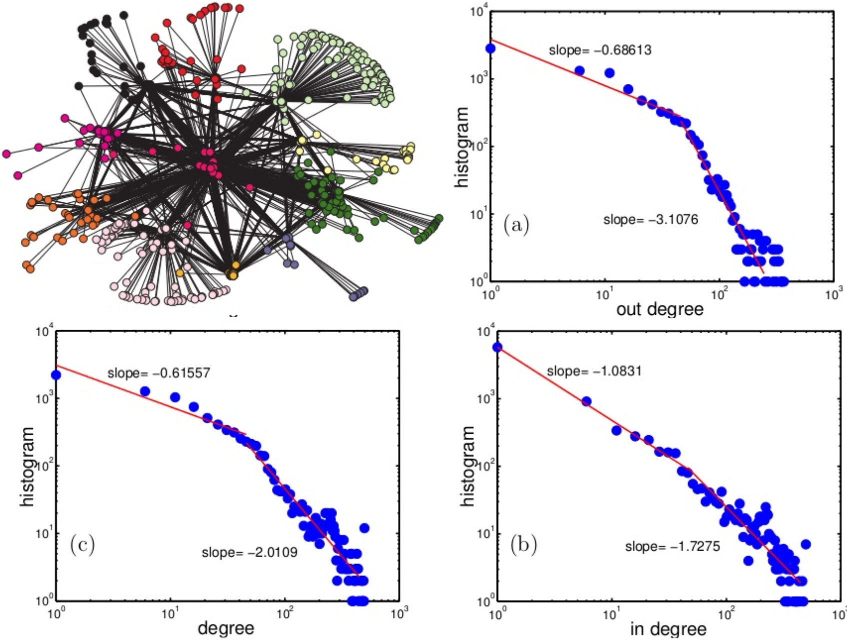
Interbank Loans[13]

- nodes: individual banks
- edges: presence of financial obligation
- edge directionality: from debtors to creditors

Example: Austrian banking sector

- low clustering coefficient (not a small world)
- the amount of obligations follows a power law
- the degree distribution indicates a scale-free network
- hierarchical structure: regional vs. national

[13] Boss, M. et al. (2004). Network topology of the interbank market.


slope= −0.68613
slope= −3.1076
(a)
histogram
out degree
slope= −0.61557
slope= −2.0109
(c)
histogram
degree
slope= −1.0831
slope= −1.7275
(b)
histogram
in degree

Resilience of Banking Networks

Network topology:

- risk sharing vs. risk of cascading failure

Cascade of insolvencies

- modeled as SIS infection
- the infection process takes into account the bank's capital
- main result: for specific nodes, the minimum capital can be determined based on their position in the network topology[14] (currently only context-free risk).

[14] Amini, H. et al. (2012) Resilience to contagion in financial networks.

Recommendation Networks

- nodes – commercial items, edges – co-occurrence
- projection of a bipartite network
- goal: to provide access to a wide (and otherwise confusing) range of products

Properties of the network[15]:

- small-world – related to the transitivity of similarity
- scale-free – indicating preferential attachment
  - if edges are constructed based on descriptors (genres, etc.), the distribution is exponential

[15] Cano. P. et al. (2005): The Topology of Music Recommendation Networks.

Personal Recommendations: Cascades[16]

- **nodes** – users; **edges** – recommendations sent during purchases (discount for referral)
- local cascades are extracted up to a distance of h
- cascades are sorted using approximate isomorphism

Results

- roughly power-law distribution of cascade size
- most of them are very short, unit cascades are predominant

[16] Leskovec J., et al. (2006). Patterns of Influence in a Recommendation Network

Id	Graph	Nodes	Edges	Book R	Book F	DVD R	DVD F	Music R	Music F	Video R	Video F
G_1		2	1	1	86,430	1	36,863	1	11,518	1	1,425
G_2		3	2	2	10,573	4	3,238	2	492	5	33
G_3		3	2	3	5,089	2	5,147	3	389	3	61
G_4		3	2	6	1,593	5	2,419	5	115	22	4
G_5		3	3	4	3,115	3	4,746	4	201	2	63
G_6		4	3	5	2,769	15	505	6	55	20	5
G_7		4	3	8	726	25	416	7	30	27	4
G_8		4	3	10	598	7	909	8	25	0	0
G_9		4	3	12	398	33	312	13	12	0	0
G_{10}		4	3	13	362	22	424	9	18	26	4
G_{11}		4	3	18	156	37	276	53	4	0	0
G_{12}		4	3	29	82	24	418	28	8	0	0
G_{13}		4	3	92	21	12	549	54	4	0	0
G_{14}		4	4	9	625	11	552	31	7	13	8
G_{15}		4	4	22	112	16	495	10	15	0	0
G_{16}		4	4	23	111	20	435	57	3	0	0
G_{17}		4	4	26	85	17	485	83	2	0	0
G_{18}		4	4	30	79	9	706	32	7	29	3
G_{19}		4	4	37	64	38	273	24	9	0	0
G_{20}		4	4	47	51	955	28	0	0	0	0
G_{21}		4	4	90	21	857	31	0	0	0	0
G_{22}		4	4	91	21	1368	20	0	0	0	0

MUNI
FACULTY
OF INFORMATICS