Modeling, analyzing, and learning from mechanisms of community building in networks

Over the past decades, we have witnessed a dramatic change in where information is provided and how information is shared. Driven by social media platforms that rely on sophisticated filtering algorithms to select who sees what, have resulted in information bubbles and communities sharing a bubble (Apprich et al., 2018; Pariser, 2011). Social relationships, both offline and digital, are commonly modeled as networks. The research conducted in this PhD project aims at developing a transdisciplinary understanding of community formation in networks, which are composed of individuals (agents), sharing rather similar judgements (rules) and opinions. This will be achieved by analyzing existing networks, modeling the dynamics of how communities emerge, to better understand why they are formed, and to critically reflect on why and how network modeling and algorithms used, e.g., by social media platforms have contributed and are contributing themselves to their formation. The ultimate goal is to understand how the modification of individual rules could lead to modified network structures. Applying this knowledge with a scientifically ethical approach, guided by media historical considerations, could provide guidelines for a “transformative change”.

The proposed PhD topic has been jointly developed by an interfaculty JKU workgroup from SOWI (Prof. Sophie N. Parragh) and TNF (Prof. Kurt Hingerl), and UAK researcher Prof. Clemens Apprich. Combining the expertise of these three researchers allows to advise a transdisciplinary PhD topic in network science. This joint expertise allows to describe the topology and model (qualitatively as well as analytically) stationary networks, to investigate global (i.e., emergent) and dynamic changes due to a variation of local rules respectively network nodes, to analyze and critically reflect on the origins of algorithms contributing towards community formation, and to potentially identify paths for future action.

Networks consist of vertices and connections usually called edges (possibly directed) between these vertices. Contributions to network science arise from graph theory (discrete mathematics), statistical mechanics (physics), sociometry, data mining and information visualization (IT), among others. Their application is manyfold: from social science and political economy models to media design, from scheduling and path planning algorithms, from understanding "frustrated spin networks" to brain models, from telephone and power grids to internet social networks; to name just a few.

In social network analysis (Wassermann and Faust, 1994), an important task concerns the detection of communities in the network. Finding (the minimum number of) cliques in a graph relates to well-known graph-theoretical problems, such as vertex coloring, which are known to be difficult from a computational perspective. In order to apply optimization technology to the identification of communities, a measure needs to be defined, which can be used in an objective function. A commonly deployed measure is modularity (Newman and Girvan, 2004). It is defined as the share of edges among all edges in the network that connect vertices within groups minus the expected fraction, if edges were randomly distributed. A high modularity value indicates that the identified community structure is strong. Fortunato (2010) gives an overview of concepts and algorithms applied to community detection. Recent methodological advances are also rooted in the operations research field (Gschwind et al., 2015, 2021).

In physics, the first models for emergent behavior and network formation evolved from chaos theory and “Synergetics” (Hermann Haken), based on the nonequilibrium thermodynamics Fokker-Planck partial differential equations for transitions in the multidimensional probability space. As a major feature, it turned out that nonlinear feedback can yield emergent behavior, limit cycles and (physical) “slaving” of the short-living systems through the long-living systems, and thereby patterns. This approach, developed in nonequilibrium statistical physics, has been soon after applied to sociological and economical questions, e.g., to compute opinion forming in networks (Weidlich and Haag, 1983). In parallel, the economist and political scientist (and Nobel prize winner) Thomas Schelling set up an agent based discrete model, which shows the dynamics of (racial) segregation, just due to local (and even weak) in-group preferences (Schelling, 1971). It turned out that the segregation of social groups resembles the alignment of ferromagnetic spins, which is often described by the Ising model. If all (or the majority of, or a cluster of) spins align parallel, a phase transition occurs. In 2021 Giorgio Parisi received the Nobel prize *“for the discovery of the interplay of disorder and fluctuations from atomic to planetary scales”.* Instead of citing his seminal theoretical contributions to physics (e.g., spin glasses), we just mention his popular science book *“The Flight of the Starlings*” (Penguin-2021), where he draws analogies between physics, animals, humans, society, etc.

The idea to analyze social or technical systems as networks goes back to the work of the Austrian psychologist Jacob Levy Moreno (1951). His sociometry describes societies on the basis of their formal and informal structures, which consist of a network of social relationship (Yuk and Halpin, 2013). Soon, just about anything was imagined as a network, especially since the network in its most abstract form can be applied to almost everything. This raises the question of new modes of subjectification, which have emerged on the basis of the current network dispositif. Since digital information and communication technologies increasingly co-determine the formation of a networked subjectivity, a better understanding of the socio-technical framework is needed, in which this subjectivity is embedded. According to Tiziana Terranova (2022), we are witnessing a fundamental shift in how the social is constructed through networks, necessitating a critical engagement with the history and structure of networks.

Taking a transdisciplinary approach, within in this thesis, the following research questions (RQ) shall be addressed:

- **RQ 1:** What can be learned from social network analysis (focusing on community detection) and what are its limitations in terms of modeling and analysis?
- **RQ 2:** What can be learned from agent-based approaches for modeling and analyzing the emergence of communities in social networks?
- **RQ 3:** How could a modification of individual rules change the network structure, to open up bubbles and increase inter-community connections?

RQ 1: To answer RQ 1, quantitative approaches towards network analysis will be used to detect communities (see Fortunato (2010) for an overview). The most promising approaches will be determined and the obtained results will be analyzed from a multidisciplinary perspective. Also, potential limitations concerning the methods themselves as well as their interpretation will be investigated.

RQ 2: To answer RQ 2, methods rooted in physics will be used to model social networks. Physical systems as spin glasses, which are frozen or (quenched) into disordered states are out of equilibrium and trapped in a few non-ergodic local minima, surrounded by large (energy) barriers (the time average of a single system is not equivalent to the instantaneous average of many - *n* - equivalent systems), *and this resembles, at least heuristically,* the echo chambers and bubbles in social networks and media. After understanding in more mathematical detail the concepts of the replica method (Parisi, 2006), its application to social networks will be investigated and the current mechanisms contributing to community building will be analyzed.

RQ 3: In order to answer RQ 3, based on the insights gained and the answers to RQs 1 and 2, modifications of the networks and their energy landscapes will be tested, to identify potential mechanisms which would - in theory - allow agents to jump out of these bubbles and traps, at least in simulations. However, it should be mentioned that such knowledge contains high risks.

It can be used to guide a transformative change, where individuals then behave differently to before, e.g., in terms of CO_2 production. But the same concepts and techniques could potentially also be misused to manipulate individuals. Therefore, ethical as well as media historical considerations will be used to guide the presentation of such research results.

Potential data sets/code

https://snap.stanford.edu/data/

References

Apprich, C., Cramer, F., Hui Kyong Chun, W., & Steyerl, H. (2018). Pattern discrimination. meson press/University of Minnesota Press, Minneapolis/London.

Fortunato, S. (2010). Community detection in graphs. *Physics reports*, *486*(3-5), 75-174.

Gschwind, T., Irnich, S., Furini, F., & Calvo, R. W. (2021). A branch-and-price framework for decomposing graphs into relaxed cliques. *INFORMS Journal on Computing*, *33*(3), 1070-1090.

Gschwind, T., Irnich, S., Furini, F., & Calvo, R. W. (2015). *Social network analysis and community detection by decomposing a graph into relaxed cliques* (Technical Report, Gutenberg School of Management and Economics, No. 1520).

Hui, Yuk & Halpin, Harry (2013). ‘Collective Individuation: The Future of the Social Web’, in Unlike Us Reader. Social Media Monopolies and Their Alternatives, ed. Geert Lovink et al. (Amsterdam: INC), 103–116.

Moreno, Jacob L. (1951). Sociometry, Experimental Method and the Science of Society: An Approach to a New Political Orientation (Boston: Beacon House).

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, *69*(2), 026113.

Pariser, E. (2011). The filter bubble: How the new personalized web is changing what we read and how we think. Penguin.

Parisi, Giorgio (1995). “On the replica approach to spin-glass theory” Philosophical Magazine **B71**, 471-478 doi: 10.1080/01418639508238538

Schelling, Thomas C. (1971). "Dynamic models of segregation". The Journal of Mathematical Sociology. **1**, 143–186; doi:10.1080/0022250x.1971.9989794 (1971)

Terranova, Tiziana (2022) After the Internet. Digital Networks between Capital and the Common (Cambridge, MA: MIT Press).

Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications. Cambridge University Press.

Weidlich, W. & Haag, G.: *Concepts and Models of a Quantitative Sociology. The Dynamics of Interacting Populations.* Springer Series in Synergetics, Vol. 14, Berlin, (1983)