UMEÅ UNIVERSITET

Through the Coding-Lens
Community Detection and Beyond

Christopher Blöcker

Akademisk avhandling

som med vederbörligt tillstånd av Rektor vid Umeå universitet för avläggande av filosofie doktorsexamen framläggs till offentligt försvar i sal NAT.D.470, byggnad Naturvetarhuset, fredag den 21 oktober, kl. 09:00.
Avhandlingen kommer att försvaras på engelska.

Fakultetsopponent: Dr. Jari Saramäki, Full Professor, Department of Computer Science, Aalto University, Aalto, Finland.

Integrated Science Lab, Department of Physics

Organization
Umeå University
Department of Physics

Document type
Doctoral thesis

Date of publication
30 September 2022

Author
Christopher Blöcker

Title
Through the Coding-Lens – Community Detection and Beyond

Abstract
We live in a highly-connected world and find networks wherever we look: social networks, public transport networks, telecommunication networks, financial networks, and more. These networks can be immensely complex, comprising potentially millions or even billions of inter-connected objects. Answering questions such as how to control disease spreading in contact networks, how to optimise public transport networks, or how to diversify investment portfolios requires understanding each network's function and working principles.

Network scientists analyse the structure of networks in search of communities: groups of objects that form clusters and are more connected to each other than the rest. Communities form the building blocks of networks, corresponding to their sub-systems, and allow us to represent networks with coarse-grained models. Analysing communities and their interactions helps us unravel how networks function.

In this thesis, we use the so-called map equation framework, an information-theoretic community-detection approach. The map equation follows the minimum description length principle and assumes complete data in networks with one node type. We challenge these assumptions and adapt the map equation for community detection in networks with two node types and incomplete networks where some data is missing. We move beyond detecting communities and derive approaches for how, based on communities, we can identify influential objects in networks, and predict links that do not (yet) exist.

Keywords
community detection, map equation, Huffman coding, network centrality, link prediction

Language
English

ISBN
978-91-7855-827-8 (print)
978-91-7855-828-5 (pdf)

Number of pages
103 + 4 papers