

Projects on (Social) Network Analysis

Frank Takes

Leiden University, The Netherlands

Bachelorklas — 11 november 2019

Frank Takes — Projects — Bachelorklas — 11 november 2019

Bachelor and master projects



- Topics: social network analysis and network science (context: graph search, graph mining, graph theory, data mining, data science, algorithms, datastructures)
- Quantitative studies with concrete algorithms, (large) real-world datasets and experiments
- Project work in C++ and/or Python
- Thesis in LATEX
- Relevant BSc courses: algorithms, data structures and data mining
- Interested? Contact me at f.w.takes@liacs.leidenuniv.nl or walk by Snellius room 157b







- Many packages for (social) network analysis exist
- Mostly in C++ and Python
- Examples: NetworKit, NetworkX, Graph-tool, teexGraph
- Used, among other things, for computing what the central players in social networks are
- Question: what is the best package in terms of performance?
- Type of work: programming, running experiments as part of a comparative analysis, working with large datasets

Project 2: Motif counting visualization



- Motifs: higher order patterns in networks / handful of nodes and edges in a particular configuration
- Building blocks of social networks
- Problem: hard to interpret motif analysis results for domain scientists
- Question: how can we build a graphical interface to interpret motif detection results?
- Type of work: implementing a cross-platform GUI that calls C++ code to count the frequency of each motif, and display this interactively

Motifs



 $M_{1,3,1}^{2}$ $M_{1,3,2}^{2}$ $M_{1,3,3}^{2}$ $M_{1,3,3}^{2}$ $M_{1,3,13}$ $M_{1,3,14}$ $M_{1,3,14}$ $M_{1,3,15}$ M1,3,16 M1,3,17 M1,3,18 € 2 - 2 👥 2 🛒 M1.3.19 M1.3.20 M1.3.21 € - 2 ----M1,3,22 M1,3,23 M1,3,24 ₹ 2 ··· M_{1,3,25} M_{1,3,26} M_{1,3,27}



- Compare a number of Breadth First Search (BFS) algorithms for traversing real-world graphs
- Starting point: several existing C++ algorithms and obtained results
- Use large-scale social network data
- Question: which algorithm has the best performance on various types of network datasets?
- Type of work: Running C++ code on multi-CPU machines

Breadth First Search





Frank Takes — Projects — Bachelorklas — 11 november 2019