

# Wissenschaftliche Zusammenfassung meiner Dissertation

Philipp Fischbeck

**Titel:** On Realistic Random Network Models for Algorithm Evaluation

Networks are widely used across many fields to represent interactions, such as the web, social interactions, infrastructure, or biological connections. Algorithms and processes on these networks have many uses, for example to find the shortest path in a road network, or to understand the spreading of information through a social network. In order to better answer such questions, we should understand how algorithms and processes are affected by features of the networks. One useful tool for this can be found in random network models, serving as proxies for real-world networks. Their features can be controlled via model parameters. While these models are usually easier to understand and analyze than real-world networks, their usefulness depends on how realistic they are.

Thus, we look at well-known network models and evaluate whether they are useful for predicting algorithm behavior on real-world networks – the external validity. We focus on many widely-used algorithms and processes, and establish that the models are good proxies for real-world networks. We observe that heterogeneity and locality are important features for explaining and predicting algorithm behavior.

In order to use random network models as proxies for real-world networks or to investigate model fit, one should utilize the model configuration capabilities optimally. As the models can be configured via parameters that affect the network features, this entails choosing the best parameters for some target network features. However, the exact relation between the configuration and the resulting network features is not clear. We present a method for estimating the best-fitting parameter configuration for commonly used network models, given some target features of the resulting networks. Our iterative method based on stochastic approximation needs only few samples and works reliably across a range of parameter configurations.

Furthermore, we take a closer look at the features and substructures of networks. In particular, based on methods from clustering and outlier detection, we consider the idea that the edge set actually comes from two edge sets with differing features. With this idea, we investigate the well-known bootstrap percolation process on this combination of networks, which had previously only been considered on single network models. We theoretically and empirically consider the process on two combined network models, with differing percolation thresholds for the two edge sets, and see an interesting emerging behavior of a slow and a rapid percolation phase. In addition, we consider ways to separate edge sets based on locality features. Assuming the edges are a combination of two network models, we show how the edges can be separated by a simple metric, even with imbalanced set sizes.

Overall, we evaluate common random network models, improve methods to utilize them, and contribute towards the design and understanding of new, combined network models.