Complex Network Analysis of Fitness Landscapes

Dissertation
zur Erlangung des Grades eines Doktors der
wirtschaftlichen Staatswissenschaften
(Dr. rer. pol.)
des Fachbereichs Rechts- und Wirtschaftswissenschaften
der Johannes Gutenberg-Universität Mainz

vorgelegt von
Sebastian Herrmann
in Mainz

im Jahre 2016

ABSTRACT

The concept of fitness landscapes originated from evolutionary biology and is relevant for numerous disciplines. In metaheuristics for combinatorial optimization, fitness landscapes are frequently used to study the structure of problems. A novel approach is to analyze fitness landscapes by "local optima networks" (LONs). A LON compresses the features of fitness landscapes in a complex network. The nodes are the local optima. The edges model the potential transitions between the local optima basins. Edges are directed and weighted by transition probabilities. Studies towards a deeper insight and exploitation of LONs are rare. The contribution of this thesis is to demonstrate how local optima networks can be used to study the structure and the search difficulty of combinatorial optimization problems for metaheuristics. For our experiments, we mainly used the Kauffman NK model of fitness landscapes. The thesis consists of four papers. In the first and second paper, we show that the PageRank centrality of the global optimum in LONs is a reliable predictor of search difficulty ($R^2 > 90\%$) for local search based metaheuristics. This is possible because PageRank is a variant of Eigenvector centrality. A LON approximates the stochastic process of an algorithm in the fitness landscape. The LON graph's matrix of edge weights is equivalent to the transition matrix of a finite-state Markov chain. The Eigenvector of the transition matrix reflects the stationary distribution of a random walk across the Markov chain. Hence, the scalar value of the global optimum approximates the probability to visit this node during search. In the third paper, we applied the Markov cluster algorithm for community detection to LONs. This reveals a structure of multiple clusters and supplements the big valley hypothesis, which states that good solutions are often contained in a single, giant cluster. The existence of multiple clusters is related to search difficulty: the size of the cluster containing the global optimum is strongly correlated to the success rate of iterated local search. This offers an explanation for failures of this metaheuristic: the perturbation operator is too weak to escape from one cluster to another. Fourth, a method is introduced to represent a landscape by a coarse-grained barrier tree. Barrier trees (disconnectivity graphs) have so far been used to study the barriers which an algorithm needs to pass in order to escape from one basin to another. We present a method based on the flooding algorithm to reveal the barriers that exist between clusters of local optima. The method is useful to obtain a coarse-grained picture of a fitness landscape. Experimental results indicate that the existence of barriers between clusters is related to search difficulty for iterated local search.

Contents

List of Papers

Chapter 2
Herrmann, Sebastian[1]; Rothlauf, Franz[1] (2015). *Predicting Heuristic Search Performance with PageRank Centrality in Local Optima Networks.* In: Proceedings of the ACM Genetic and Evolutionary Computation Conference (GECCO 2015), Madrid, July, 11-15, 2015. ACM 2015.

Chapter 3
Herrmann, Sebastian[1] (2016). *Determining the Difficulty of Landscapes by PageRank Centrality in Local Optima Networks.* In: Proceedings of the 16th European Conference on Evolutionary Computation in Combinatorial Optimisation (EvoCOP 2016), Porto, March 30 - April 01, 2016. Lecture Notes in Computer Science 9595, Springer 2016.
The paper has been invited for submission to a special issue on evolutionary computation in combinatorial optimization in the *Journal of Heuristics* (currently in second round revision).

Chapter 4
Herrmann, Sebastian[1]; Ochoa, Gabriela[2]; Rothlauf, Franz[1] (2016). *Communities of Local Optima as Funnels in Fitness Landscapes.* In: Proceedings of the ACM Genetic and Evolutionary Computation Conference (GECCO 2016), Denver, Colorado, July 20-24, 2016. ACM 2016.

Chapter 5
Herrmann, Sebastian[1]; Ochoa, Gabriela[2]; Rothlauf, Franz[1] (2016). *Coarse-Grained Barrier Trees of Fitness Landscapes.* In: The 14th International Conference on Parallel Problem Solving from Nature - PPSN XIV, Edinburgh, Scotland, September 17-21, 2016. Lecture Notes in Computer Science, Springer 2016.

[1]Johannes Gutenberg-Universität Mainz, Lehrstuhl für Wirtschaftsinformatik und BWL, Jakob-Welder-Weg 9, D-55128 Mainz, Germany
[2]University of Stirling, Department of Computing Science and Mathematics, Stirling FK9 4LA, Scotland, UK

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

"I don't have any solution, but I certainly admire the problem."

Ashleigh Brilliant

The concept of fitness landscapes originated from theoretical biology. Richter (2014) considers it to address "some of the most intriguing and fundamental questions in natural and artificial evolution: what way is evolution going, to what extend is it predictable, what can be realistically expected to be the outcome of a certain period of evolutionary development?" With his seminal work, Wright (1932) was the first to present the idea of a fitness function over a genotype space. The combinations of alleles for the gene loci shape a landscape in which the height represents the fitness of the different genomes or species, respectively. The term "fitness landscape" was coined later by Kauffman and Levin (1987), who introduced the NK model of combinatorial fitness landscapes. The NK model is a canonical example of fitness landscapes and was the first computational approach introduced to study evolutionary dynamics. Evolutionary progress in fitness landscapes is achieved by "adaptive walks" across the surface, triggered by genetic mutation or crossbreeding (Kauffman and Weinberger, 1989). In analogy to combinatorial optimization, evolution strives to maximize the fitness by the adaptation of genes. Based on this metaphor, the landscape structure reflects the environment, and the paths engraved on the landscape surface determine the stochastic evolutionary dynamics. For instance, the existence of multiple peaks or niches may explain the diversity and distribution of genotypes in a population. A "rugged" landscape with many local optima provides numerous niches. Consequently in such landscapes, it is also more difficult to locate the global optimum[1] with an adaptive walk. Figure 1.1 visualizes a hypothetical, two-dimensional[2] fitness landscape.

[1] Fitness landscapes often contain multiple global optima in terms of their fitness. In the Kauffman NK model, there is "a single sequence which is the only and global optimum" (Kauffman and Weinberger, 1989, p.218), which is why we speak of *the* global optimum.

[2] This is a rather exceptional case for the purpose of illustration. Usually, a fitness landscape is an n-dimensional hypercube, where n is the number of optimization variables.

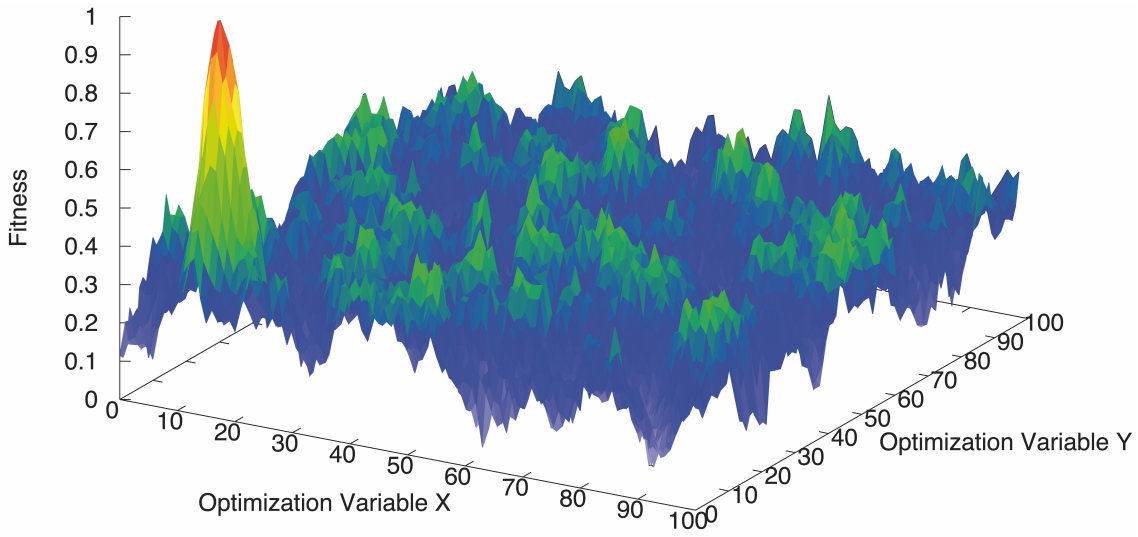

Figure 1.1: Depiction of a hypothetical, two-dimensional, static fitness landscape as a mountainous region with peaks, valleys and ridges. The points are arranged by some neighbor relationship; the height is the fitness. The highest peak is the global optimum in terms of fitness. A local optimum is a point where all neighbors are of lower fitness.

The metaphor of fitness landscapes has been adapted beyond evolutionary biology and is relevant to a wide number of disciplines, e.g. as energy landscapes in chemical physics (Stillinger, 1995), in complex adaptive systems (Holland, 1992; Miller and Page, 2007) and organizational theory (Billinger et al., 2014; Levinthal, 1997; Rivkin, 2000). Most notably is the role of fitness landscapes in metaheuristics for solving combinatorial optimization problems.

Metaheuristics are general-purpose techniques applicable to a wide range of optimization problems. Unsurprisingly, many metaheuristics are nature-inspired methods as well and mimic biological or physical optimization principles. Examples are evolutionary computation (Holland, 1975; Rechenberg, 1973; Schwefel, 1977), ant colony optimization (Middendorf et al., 2002), and simulated annealing (Kirkpatrick et al., 1983). Metaheuristics have a black-box nature and do not guarantee high performance for every kind of problem, as formally proved by the no-free-lunch theorem (Wolpert and Macready, 1997). However, some are known to work surprisingly well on a number of NP-hard problems.

In metaheuristics for combinatorial optimization, a fitness landscape is defined by a triplet: (i) a finite set of feasible solutions, (ii) a fitness function defined over all solutions, and (iii) a neighborhood structure, usually as given by a distance function. The topological structure of fitness landscapes is frequently used to portray the

structure of optimization problems (Reeves and Rowe, 2002). This is relevant for two reasons:

1. Structural features are useful to predict the "search difficulty" of a problem (Lu et al., 2014; Naudts and Kallel, 2000; Ochoa et al., 2014). For instance, a common assumption is that landscape ruggedness leads to higher search difficulty for many metaheuristics (Kallel et al., 2001). The question of which landscape properties induce search difficulty is a topic of ongoing discussion.
2. Problem-specific knowledge is necessary to build customized heuristics (Rothlauf, 2011) and to select an appropriate algorithm to solve a given problem (Malan and Engelbrecht, 2014). Fitness landscape analysis is a method to support this process (Pitzer and Affenzeller, 2012). It is hoped that new analytical methods reveal yet unknown patterns in landscape structure exploitable to construct better problem-specific heuristics.

To study problem structure and predict search difficulty, it is common to apply reductionist concepts. Reductionism means the attempt to predict emerging phenomena, such as the behavior of a system, from the properties of its constituents. An example is to predict search difficulty by ruggedness, which is usually measured by the correlation of fitness between randomly sampled pairs of neighboring solutions. It is possible to draw meaningful conclusions from this correlation, but two landscapes with the same level of ruggedness may still vary significantly in their difficulty. The same holds for other statistical measures, e.g. deceptiveness (Jones and Forrest, 1995). Probably none of them captures all structural aspects of landscapes.

In the last decade, we observe a widespread interest in the epistemological paradigm of complexity, which takes into account the interconnections of systems, and asks how the emerging behavior is encoded in the system's topology. Many systems are complex since the interconnections are often ambiguous and non-trivial, i.e. not randomly distributed or in form of a simple pattern. Complex network analysis is a framework to capture a system's underlying organizing principles in a quantitative way (Albert and Barabási, 2002). Basic elements of a network are its entities, called nodes, and the connections between the nodes, called ties (Easley and Kleinberg, 2010). Examples are social networks (Borgatti, Mehra, et al., 2009), collaborative agent networks (Lazer and Friedman, 2007; Mason and Watts, 2012) and recommendation or product networks (Oestreicher-Singer and Sundararajan, 2012).

A novel approach to take a complex system perspective on fitness landscapes are "local optima networks" ("LONs"; Ochoa, Tomassini, et al., 2008). LONs have been inspired by the study of energy landscapes in chemical physics (Doye and Massen, 2005). In a LON, the nodes are the local optima of the fitness landscape. The ties connecting the nodes model the potential transitions between the local optima basins. The result is a compressed, mathematical representation which allows us to apply network analysis methods to the underlying fitness landscape.

1.1 Purpose of the Thesis

Since the introduction of local optima networks, there exist only a handful of studies in which an analysis of local optima networks has been conducted to reveal characteristic patterns in the structure of relevant problems (Ochoa, Verel, Daolio, et al., 2014). Only two of the studies have addressed the relationship between local optima network structure and the search difficulty of the underlying optimization problem. Chicano et al. (2012) as well as Daolio, Verel, et al. (2012) found that the path lengths to the global optimum seem to play a key role for explaining search difficulty. However, other important structural features (centrality of high-quality solutions, cluster structure of the landscape) have so far not been examined in the context of search difficulty. The guiding research questions of my thesis are:

1. How can we predict the search difficulty of problems for metaheuristics with the analysis of fitness landscapes by local optima networks?
2. Which structural patterns of fitness landscapes does the analysis of local optima networks reveal?

The objective of this thesis is to examine how local optima networks reflect the structure and search difficulty of combinatorial optimization problems for metaheuristics. To address the research questions, I present four studies in which various methods of network analysis are applied to local optima networks of NK fitness landscapes. To predict the search difficulty and study the problem structure with local optima networks, I applied one approach of network analysis on the micro level (node centrality) and one on the macro level (cluster analysis). Centrality is a concept to describe the influence or position of a node (Borgatti and Everett, 2006). It is a typical representative of micro level analysis, which addresses the relationship between node properties and system behavior. The macro level focuses on structural properties of the network (Boccaletti et al., 2006). Cluster analysis seeks to identify cliques of nodes with a strong cohesion (Fortunato, 2010). Beyond these network concepts, I developed an approach to analyze local optima networks with barrier trees. Barrier trees are another representation derived from chemical physics to characterize the global structure of a landscape by the barriers between local optima basins (Becker and Karplus, 1997; Flamm et al., 2002). My approach operates on the level of local optima networks and draws a more coarsely grained picture of landscapes than traditional barrier trees.

The experimental results of my work confirm that both the centrality of high-quality solutions and the clustering pattern of solutions in landscapes are highly reliable predictors of search difficulty. Since these measures capture the structural properties of landscapes in an encompassing way, my results provide interesting new insights on general landscape features. In particular, they complement the well-known big valley hypothesis (Boese et al., 1994; Hains et al., 2011).

1.2 Structure of the Thesis

My thesis comprises four research papers, ordered by their date of publication. Structure and contents of my thesis are as follows:

In Chapter 2, I present a study on the relationship between the PageRank centrality (Brin and Page, 1998) of the global optimum in a local optima network and the performance of local search-based methods (hill climbing, simulated annealing). We used multiple instances of the Kauffman NK model and the traveling salesman problem as fitness landscapes. To model the local optima networks, we referred to the initial concept by Ochoa et. al (2010), containing *edges with basin transition probabilities*. The PageRank predicts the empirical success rate with high accuracy.

I conducted a further study on the relationship between the PageRank centrality of the global optimum and the performance of metaheuristics, which is presented in **Chapter 3**. Here, the objective was to predict the performance of *iterated local search*. For this purpose, I used the recently introduced variant of local optima networks with *escape edges* (Verel et al., 2012). The experiment with numerous instances of the NK model shows that PageRank is a reliable predictor in this case, too. Furthermore, the study demonstrates how to predict the expected solution quality and average running time with PageRank and local optima networks.

In Chapter 4, a study is shown on *community detection* in local optima networks of the Kauffman NK model. For community detection, we used the *Markov cluster algorithm* (van Dongen, 2001). The analysis reveals a structure of multiple clusters. This result supplements the big valley hypothesis (Boese et al., 1994; Hains et al., 2011), which states that good solutions are often contained in a single, giant cluster. The existence of multiple clusters offers a new explanation for the search difficulty of landscapes for metaheuristics implementing the principle of iterated local search.

In Chapter 5, I present a new approach to study fitness landscapes on a coarse level of granularity. Our method computes a *coarse-grained barrier tree* of a landscape which retains the global structure and allows the eventual visualization of larger landscapes. Barrier trees (Flamm et al., 2002) or disconnectivity graphs (Becker and Karplus, 1997) have so far been used in fitness landscape analysis to study the barriers for an algorithm to escape from one basin to another (Hallam and Prügel-Bennett, 2005). The core of our method is a new variant of the *flooding algorithm* by van Stein et al. (2013) to detect the barriers between clusters of local optima (instead of basins), as introduced in Chapter 4. We demonstrate our method with the NK model. The results indicate that the existence of barriers between clusters is related to search difficulty for iterated local search.

A brief summary of the results and main contributions of my thesis is given in **Chapter 6**. Furthermore, I discuss limitations of my work, implications and potential areas of future research.

References

Albert, Réka and Albert-László Barabási (2002). Statistical mechanics of complex networks. ***Reviews of Modern Physics,*** 74(1): 47–97.

Becker, Oren M. and Martin Karplus (1997). The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. ***The Journal of chemical physics,*** 106(4): 1495–1517.

Billinger, Stephan, Nils Stieglitz, and Terry R. Schumacher (2014). Search on Rugged Landscapes: An Experimental Study. ***Organization Science,*** 25(1): 93–108.

Boccaletti, Stefano, Vito Latora, Yamir Moreno, Martín Chávez Hoffmeister, and Dong-Uk Hwang (2006). Complex networks: Structure and dynamics. ***Physics Reports,*** 424(4-5): 175–308.

Boese, Kenneth D., Andrew B. Kahng, and Sudhakar Muddu (1994). A new adaptive multi-start technique for combinatorial global optimizations. ***Operations Research Letters,*** 16(2): 101–113.

Borgatti, Stephen P. and Martin G. Everett (2006). A Graph-theoretic perspective on centrality. ***Social Networks,*** 28(4): 466–484.

Borgatti, Stephen P., Ajay Mehra, Daniel J. Brass, and Giuseppe Labianca (2009). Network analysis in the social sciences. ***Science,*** 323(5916): 892–895.

Brin, Sergey and Lawrence Page (1998). The anatomy of a large-scale hypertextual web search engine. ***Computer Networks and ISDN Systems,*** 30(1-7): 107–117.

Chicano, Francisco, Fabio Daolio, Gabriela Ochoa, Sébastien Verel, Marco Tomassini, and Enrique Alba (2012). Local Optima Networks, Landscape Autocorrelation and Heuristic Search Performance. In: ***Parallel Problem Solving from Nature - PPSN XII: 12th International Conference.*** Ed. by Carlos A. Coello Coello, Vincenzo Cutello, Kalyanmoy Deb, Stephanie Forrest, Giuseppe Nicosia, and Mario Pavone. Vol. 7492 LNCS. Berlin and Heidelberg: Springer: 337–347.

Daolio, Fabio, Sébastien Verel, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks and the performance of iterated local search. In: ***Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference - GECCO '12.*** Ed. by Terence Soule. Philadelphia, Pennsylvania, USA: ACM Press: 369.

Doye, Jonathan P. K. and Claire P. Massen (2005). Characterizing the network topology of the energy landscapes of atomic clusters. ***Journal of Chemical Physics,*** 122(8): 084105.

Easley, David and Jon Kleinberg (2010). ***Networks, crowds, and markets: Reasoning about a highly connected world.*** 1st ed. Cambridge: Cambridge Univ. Press.

Flamm, Christoph, Ivo L. Hofacker, Peter F. Stadler, and Michael T. Wolfinger (2002). Barrier Trees of Degenerate Landscapes. ***Zeitschrift für Physikalische Chemie,*** 216: 155–173.

Fortunato, Santo (2010). Community detection in graphs. ***Physics Reports,*** 486(3-5): 75–174.

Hains, Doug R., Darrel L. Whitley, and Adele E. Howe (2011). Revisiting the big valley search space structure in the TSP. ***Journal of the Operational Research Society,*** 62(2): 305–312.

Hallam, Jonathan and Adam Prügel-Bennett (2005). Large barrier trees for studying search. ***IEEE Transactions on Evolutionary Computation,*** 9(4): 385–397.

Holland, John H. (1975). ***Adaptation in Natural and Artificial Systems.*** MIT Press Cambridge.

Holland, John H. (1992). Complex Adaptive Systems. ***Daedalus,*** 121(1): 17–30.

Jones, Terry and Stephanie Forrest (1995). Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms. In: ***Proceedings of the Sixth International Conference on Genetic Algorithms.*** San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.: 184–192.

Kallel, L., B. Naudts, and Colin R. Reeves (2001). Properties of Fitness Functions and Search Landscapes. In: ***Theoretical Aspects of Evolutionary Computing.*** Ed. by Leila Kallel, Bart Naudts, and Alex Rogers. Natural Computing Series. Berlin and Heidelberg: Springer: 175–206.

Kauffman, Stuart A. (1993). ***The origins of order.*** Oxford University Press.

Kauffman, Stuart A. and Simon Levin (1987). Towards a General Theory of Adaptive Walks on Rugged Landscapes. ***Journal of Theoretical Biology,*** 128(1): 11–45.

Kauffman, Stuart A. and Edward D. Weinberger (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. ***Journal of Theoretical Biology,*** 141(2): 211–245.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi (1983). Optimization by simulated annealing. ***Science,*** 220(4598): 671–80.

Lazer, David and Allan Friedman (2007). The Network Structure of Exploration and Exploitation. ***Administrative Science Quarterly,*** 52: 667–694.

Levinthal, Daniel A. (1997). Adaptation on Rugged Landscapes. ***Management Science,*** 43(7): 934–950.

Lu, Guanzhou, Jinlong Li, and Xin Yao (2014). Fitness Landscapes and Problem Difficulty in Evolutionary Algorithms: From Theory to Applications. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Emergence, Complexity and Computation. Berlin and Heidelberg: Springer: 133–152.

Malan, Katherine M. and Andries P. Engelbrecht (2014). Fitness Landscape Analysis for Metaheuristic Performance Prediction. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Berlin and Heidelberg: Springer: 103–129.

Mason, Winter and Duncan J. Watts (2012). Collaborative learning in networks. ***Proceedings of the National Academy of Sciences,*** 109(3): 764–769.

Middendorf, Martin, Frank Reischle, and Hartmut Schmeck (2002). Multi Colony Ant Algorithms. ***Journal of Heuristics,*** 8(3): 305–320.

Miller, John H. and Scott E. Page (2007). ***Complex Adaptive Systems: An Introduction to Computational Models of Social Life.*** Princeton. Princeton University Press.

Naudts, B. and L. Kallel (2000). A comparison of predictive measures of problem difficulty in evolutionary algorithms. ***IEEE Transactions on Evolutionary Computation,*** 4(1): 1–15.

Ochoa, Gabriela, Marco Tomassini, Sébastien Verel, and Christian Darabos (2008). A study of NK landscapes' basins and local optima networks. In: ***Proceedings of the 10th annual conference on Genetic and evolutionary computation - GECCO '08.*** Ed. by Maarten Keijzer. Atlanta, GA, USA: ACM Press: 555–562.

Ochoa, Gabriela, Sébastien Verel, Fabio Daolio, and Marco Tomassini (2014). Local Optima Networks: A New Model of Combinatorial Fitness Landscapes. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Vol. 6. Emergence, Complexity and Computation. Berlin and Heidelberg: Springer: 233–262.

Ochoa, Gabriela, Sébastien Verel, and Marco Tomassini (2010). First-improvement vs. best-improvement local optima networks of nk landscapes. In: ***PPSN'10: Proceedings of the 11th International Conference on Parallel Problem Solving from Nature.*** Ed. by Robert Schaefer, Carlos Cotta, Joanna Kolodziej, and Günter Rudolph. Vol. I. Kraków, Poland: Springer: 104–113.

Oestreicher-Singer, Gal, Barak Libai, Liron Sivan, Eyal Carmi, and Ohad Yassin (2013). The Network Value of Products. ***Journal of Marketing,*** 77(3): 1–14.

Oestreicher-Singer, Gal and Arun Sundararajan (2012). Recommendation Networks and the Long Tail of Electronic Commerce. ***MIS Quarterly,*** 36(1): 65–83.

Pitzer, Erik and Michael Affenzeller (2012). A Comprehensive Survey on Fitness Landscape Analysis. In: ***Recent Advances in Intelligent Engineering Systems.*** Ed. by János Fodor, Ryszard Klempous, and Carmen Paz Suárez Araujo. Vol. 378. Studies in Computational Intelligence. Berlin and Heidelberg: Springer: 161–191.

Rechenberg, Ingo (1973). ***Evolutionsstrategie Optimierung technischer Systeme nach Prinzipien der biologischen Evolution.*** Stuttgart: Frommann-Holzboog.

Reeves, Colin R. and Jonathan E. Rowe (2002). ***Genetic Algorithms: Principles and Perspectives.*** Vol. 20. Operations Research/Computer Science Interfaces Series. Boston: Kluwer Academic Publishers.

Richter, Hendrik (2014). Fitness Landscapes: From Evolutionary Biology to Evolutionary Computation. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Berlin and Heidelberg: Springer: 3–31.

Rivkin, Jan W (2000). Imitation of Complex Strategies. ***Management Science,*** 46(6): 824–844.

Rothlauf, Franz (2011). ***Design of modern heuristics: Principles and application.*** Berlin and Heidelberg: Springer.

Schwefel, Hans-Paul (1977). ***Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie.*** Basel: Birkhäuser.

Stillinger, Frank H. (1995). A Topographic View of Supercooled Liquids and Glass Formation. ***Science,*** 267(5206): 1935–1939.

Van Dongen, Stijn (2001). "Graph clustering by flow simulation". PhD thesis. Utrecht University.

Van Stein, Bas, Michael Emmerich, and Zhiwei Yang (2013). Fitness Landscape Analysis of NK Landscapes and Vehicle Routing Problems by Expanded Barrier Trees. In: ***EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation.*** Ed. by Alexandru-Adrian Tantar, Emilia Tantar, Jian-Qiao Sun, Wei Zhang, Qian Ding, Oliver Schütze, Michael Emmerich, Pierrick Legrand, Pierre Del Moral, and Carlos A. Coello Coello. Vol. 227. Advances in Intelligent Systems and Computing. Springer: 75–89.

Verel, Sébastien, Fabio Daolio, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks with escape edges. In: ***Artificial Evolution.*** Angers, France: Springer: 49–60.

Wolpert, David H. and William G. Macready (1997). No free lunch theorems for optimization. ***IEEE Transactions on Evolutionary Computation,*** 1(1): 67–82.

Wright, Sewall (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: ***Proceedings of the 6th International Congress of Genetics.*** Ed. by Donald F. Jones. Ithaca, New York: Morgan Kaufmann Publishers Inc.: 356–366.

Chapter 2

Predicting Heuristic Search Performance with PageRank Centrality in Local Optima Networks

Sebastian Herrmann, Franz Rothlauf

Abstract

Previous studies have used statistical properties of fitness landscapes such as ruggedness and deceptiveness in order to predict the expected quality of heuristic search methods. Novel approaches for predicting the performance of heuristic search are based on the analysis of local optima networks (LONs). A LON is a compressed stochastic model of a fitness landscape's basin transitions. Recent literature has suggested using various LON network measurements as predictors for local search performance. In this study, we suggest PageRank centrality as a new measure to predict the performance of heuristic search methods using local search. PageRank centrality is a variant of Eigenvector centrality and reflects the probability that a node in a network is visited by a random walk. Since the centrality of high-quality solutions in LONs determines the search difficulty of the underlying fitness landscape and since the big valley property suggests that local optima are not randomly distributed in the search space but rather clustered and close to one another, PageRank centrality can serve as a good predictor for local search performance. In our experiments with the Kauffman NK model and the traveling salesman problem, we found that the PageRank centrality is a very good predictor for the performance of first-improvement local search as well as simulated annealing, since it explains more than 90% of the variance of search performance. Furthermore, we found that PageRank centrality is a better predictor of performance than traditional approaches such as ruggedness, deceptiveness, and the length of the shortest path to the optimum.

2.1 Introduction

Most evolutionary search concepts use intensification during search. A common concept for intensification is to apply some kind of local search that applies incremental changes to a solution candidate and prefers solutions with a higher solution quality. Standard examples in evolutionary search that rely on local search are hill climbing, simulated annealing (Kirkpatrick et al., 1983), and genetic algorithms, in particular evolutionary strategies or evolutionary programming. A downside of local search is its tendency to get stuck in local optima (Glover, 1986). Thus, using only local search alone usually does not guarantee finding the global optimum[1]. As local search is a core of many heuristic optimization approaches, predicting the performance of local search (He et al., 2007) is relevant for the process of selecting an appropriate algorithm for a given problem and also for the design of search techniques (Malan and Engelbrecht, 2014; Rothlauf, 2011).

Most approaches for predicting the performance of local search use the concept of fitness landscapes (Wright, 1932). The fitness landscape of a problem instance is defined by the solution candidates, their neighborhood structure (defined by the used search operator), and the fitness of the solution candidates. In general, two solutions are adjacent if they can be transformed into each other by an incremental change, that is, a single local search step. Often, structural properties of the resulting landscape are used for predicting algorithm performance. For instance, the more rugged a fitness landscape is, the higher the probability that an algorithm gets stuck in a local optimum (Stadler, 1996).

Local optima networks (LONs) can be used for the statistical network analysis of fitness landscapes (Ochoa, Tomassini, et al., 2008). In general, a LON is a stochastic and compressed representation of a fitness landscape, where the vertex set (nodes) of the LON represents the landscape's local optima. In a LON, there is a directed edge between two nodes A and B if a local search step directly transforms one solution in node A's basin of attraction into another solution that lies in node B's basin of attraction. The basin of attraction of a local optimum is the set of solution candidates from which the focal local optimum can be reached using a number of local search steps (no diversification steps are allowed). Each edge of a LON is weighted by the number of possible transitions between the two basins of attraction.

LONs have already been used successfully to predict the performance of heuristic optimization methods using local search. For example, Daolio, Verel, et al. (2012) predicted the run-time performance of iterated local search in Kauffman's NK landscapes (Kauffman and Levin, 1987). The authors proposed four different network metrics as predictors, of which the average length of the shortest paths to the optimum had the highest accuracy for predicting performance ($R^2 \approx 0.5$). These results

[1] We assume optimization problems with a single global optimum.

are promising and inspired us to conduct further research on predicting search performance by network analysis of LONs.

In network analysis, the concept of centrality describes how important or influential a node is—depending on the network's linkage structure and the focal node's location (Freeman, 2004). The Eigenvector centrality is a measurement of the centrality of a node in a network. It is based on random walks on the network graph, where nodes and edges can be revisited multiple times (Borgatti, 2005). A prominent implementation of the Eigenvector centrality is Google's PageRank algorithm (Brin and Page, 1998). The PageRank value of a node (originally defined for a network of web pages) is the probability that a random surfer would visit this node/web page (Franceschet, 2011).

Thus, in this paper, we study whether the PageRank centrality of the global optimum in the LON is a valuable predictor of performance for local-search based metaheuristics. As a performance measure, we use the probability of finding the global optimum (success rate) and the average number of function evaluations (runtime). We performed a series of experiments on scalable test problems (Kauffman's NK landscapes) and a representative of a combinatorial optimization problem (symmetric traveling salesman problem). We found that the PageRank centrality of an instance's global optimum is a good predictor of the probability that the global optimum can be located by (i) stochastic hill climbing and (ii) simulated annealing.

The high accuracy of PageRank centrality at predicting the performance of local search can be explained by the structure of the LON that models the problem's fitness landscape. The weighted edges in a LON describe possible transition paths of local search. Along the edges, local search can switch between two basins of attraction. Thus, the performance of a local search process highly depends on whether the path between the local optima in a LON leads to the global optimum. If many paths lead to the global optimum, the performance of the local search is high; if there are only a few paths, local search can hardly make its way to the global optimum, leading to low performance. For instance, if the number of edges (transition probabilities) are equal between all local basins of attraction, there is no inherent trajectory in such a landscape that leads local search towards the optimum and we expect a low probability of finding the global optimum. In contrast, the probability to find the global optimum is high if the LON's linkage structure supports the convergence into the basin around the optimum. This explanation is in line with the findings for many combinatorial optimization problems, where a "big valley structure" (Boese et al., 1994) can be observed. In many of such problems, local optima are not randomly distributed in the search space; rather they are clustered. Thus, most local optima are close to each other and the global optimum is usually not isolated in the search space; rather it is surrounded by all the local optima. Big valley properties have been observed for the TSP (ibid.) and other problem types.

Our paper is structured as follows: In Section 2.2, we describe the local search-based algorithms used for our experiments (hill climbing and simulated annealing). In Section 2.3, we introduce a formal definition of a fitness landscape, the neighborhood structure, basins of attraction and some commonly used landscape features. In Section 2.4, we give a formal definition of the concept of local optima networks. Section 2.5 deals with the calculation of PageRank centrality and its application in LONs. In Section 2.6, we describe our experiment setup and the problem types used (NK model and TSP). We present our results in Section 2.7 and draw our conclusions in Section 2.8.

2.2 Local Search Algorithms

This section briefly reviews the two variants of heuristic optimization methods used as representatives for local search-based methods: hill climbing and simulated annealing.

2.2.1 Hill Climbing

Hill climbing is a straightforward and rather simple application of local search. Let S be the solution space, i.e., the set of all valid problem solutions. The function $f : S \to \mathbb{R}_{\geq 0}$ assigns a fitness value to each $s \in S$. Hill climbing starts with a randomly selected solution $s_0 \in S$. Then, in iterative steps, the algorithm selects a random solution x from the neighborhood of s_i. The neighborhood $N(s)$ is the set of solutions that can be reached by performing an incremental change to s. In first improvement hill climbing (fi-hc), a random solution is selected from the neighborhood; if it is better than the existing solution, it replaces the existing solution (Algorithm 2.1). Hill climbing performs incremental changes until no further improvement can be achieved or a fixed limit of iterations has been reached. We use first improvement hill climbing as a representative of a basic local search method since it is well-understood and often used.

2.2.2 Simulated Annealing

The major drawback of hill climbing is that it can get stuck in a local optimum (Glover, 1986). Simulated annealing (SA) is a method inspired by statistical mechanics that has the ability to overcome local optima (Kirkpatrick et al., 1983). It is based on an analogy with cooling down a liquid to a solid substance. Algorithm 2.2 shows the principle of SA using a fixed cooling schedule for a maximization problem.

Algorithm 2.1: First Improvement Hill Climbing

Require: Solution space S,
Fitness function $f(S)$,
Neighborhood $N(S)$

1: i $\leftarrow 0$
2: Choose initial random solution $s_0 \in S$
3: **repeat**
4: choose random $x \in N(s_i)$
5: **if** $f(x) > f(s_i)$ **then**
6: $s_{i+1} \leftarrow x$
7: **else**
8: $s_{i+1} \leftarrow s_i$
9: **end if**
10: $i \leftarrow i + 1$
11: **until** s_i is local optimum
12: **return** s_i

Algorithm 2.2: Simulated Annealing

Require: Solution space S,
Fitness function $f(S)$,
Neighborhood $N(S)$,
Initial Temperature T_0,
Cooling Rate α

1: $i \leftarrow 0$
2: Choose random initial solution $s_0 \in S$
3: **repeat**
4: choose random $x \in N(s_i)$
5: **if** $f(x) > f(s_i)$ **then**
6: $s_{i+1} \leftarrow x$
7: **else if** $\frac{e^{-|f(s_i)-f(x)|}}{T_i} \geq rand(0,1)$ **then**
8: $s_{i+1} \leftarrow x$
9: **else**
10: $s_{i+1} \leftarrow s_i$
11: **end if**
12: $T_{i+1} \leftarrow \alpha \times T_i$
13: $i \leftarrow i + 1$
14: **until** s_i is global optimum or i $>$ iteration limit
15: **return** s_i

SA always accepts better solutions; solutions with a lower fitness are accepted with probability $\frac{e^{-|f(s_i)-f(x)|}}{T_i}$. Thus, the probability of accepting worse solutions decreases with a higher fitness difference $|f(s_i) - f(x)|$ and a lower temperature T_i. For the temperature $T \to 0$, SA becomes first improvement hill climbing. The parameters for the initial temperature T_0 and cooling rate α must be set problem-specific. A rule of thumb for a proper starting temperature is to randomly select a number of solutions and to set $T_0 \approx \sigma(f(s))...\, 2\sigma(f(s))$ (where σ is the standard deviation). A proper setting for the cooling is $\alpha \in (0.9, 0.999)$ (Laarhoven and Aarts, 1988). For our experiments, we randomly sampled 1000 solutions s_i before each SA run and set the initial temperature to $T_0 = 1.5 \times \sigma(f(s_i))$. For all runs, we set $\alpha = 0.97$.

2.3 Fitness Landscape Analysis

2.3.1 Neighborhood Structure

A fitness landscape (Wright, 1932) is a triplet of a search space S, a fitness function f, and a neighborhood structure $N(S)$. There is a set of valid solution candidates $s \in S$. The evaluation function $f : S \to \mathbb{R}_{\geq 0}$ assigns a fitness value to each $s \in S$. The neighborhood function $N : S \to \mathcal{P}(S)$ assigns a set of neighbors $N(s)$ (which is a subset of S) to every $s \in S$ [2]. The neighborhood determines the position of each s in the landscape (Reidys and Stadler, 2002). Furthermore, we assume a distance function between two solutions s_0 and s_1 as

$$d : (s_0, s_1) \to \mathbb{N}_0; s_0, s_1 \in S. \tag{2.1}$$

The distance function depends on the search operator used. Usually, the application of a local search operator creates a new solution s_1 with distance $d_{s_0,s_1} = 1$ to the original solution s_0. Using distances, we can define the neighborhood function as

$$N : s_0 \to \{s_1 \in S : \, s_1 \neq s_0 \wedge 0 < d(s_0, s_1) \leq d_{\max}\}. \tag{2.2}$$

Most local search uses a low constant value for $d_{\max}$ (like $d_{\max} = 1$). Other heuristic methods (like variable neighborhood search) vary $d_{\max}$ during run-time to obtain higher diversification and to escape from a local optimum by using perturbation steps. This results in changes in the landscape during the run of the algorithm, and makes static analyses more difficult. Since most operators of local-search based techniques are based on simple metrics—e.g. Hamming distance—it is generally accepted to study a landscape defined by a fitness function and one or several induced distances (Pitzer and Affenzeller, 2012).

[2] The reader should be aware that different definitions of $N(s)$ are possible. For example, variable neighborhood search iteratively switches between different neighborhoods.

2.3.2 Definition of Local Optima

A fitness landscape can have one or more local optima. A local optimum is a solution that has no superior neighbors ("top of the hill"). For a maximization problem, we define a function

$$N_{\text{sup}}(s) = \{n \in N(s) : f(n) > f(s)\} \tag{2.3}$$

which returns the neighbors of a solution $s \in S$ with superior fitness. Thus, the set

$$LO = \{lo \in S : N_{\text{sup}}(lo) = \emptyset\}. \tag{2.4}$$

contains all solutions that have no neighbors with superior fitness, which is equivalent to the set of local optima. LO also contains all global optima.

2.3.3 Landscape Features

Many concepts that can be used for the prediction of performance focus on the structural properties of fitness landscapes. Kallel et al. (2001) provide a well-elaborated collection of such features. In the literature, two of the frequently used concepts to help predict performance are ruggedness (Weinberger, 1990) and deceptiveness (Jones and Forrest, 1995). The idea of ruggedness is that the smoother the landscape is, the more easily heuristic search methods can find the global optimum. Ruggedness is usually measured as the correlation of fitness values between neighboring solutions (nearest-neighbor-correlation).

A landscape is deceptive if the structure of the search space leads local search approaches away from the global optimum. A measurement for deceptiveness is the correlation between the fitness of the solutions and their distance to the global optimum (Jones and Forrest, 1995). For the calculation of fitness-distance correlation, the global optimum must be known in advance.

2.3.4 Basins of Attraction

Basins of attraction are an important concept for the analysis of fitness landscapes. Each local (or global) optimum is surrounded by a basin of attraction, which is defined as the set of solution candidates from which local search converges to the local optimum ("way-points that lead to the top of the hill"). Identifying the basins is a prerequisite for the calculation of LONs. They depend on the selection-rule of the hill climbing algorithm (Ochoa, Verel, and Tomassini, 2010). The function

$$B : lo \rightarrow \mathcal{P}(S \backslash LO) \tag{2.5}$$

assigns a set of solutions P (which represent the basin) to each local optimum $lo \in LO$. First improvement hill climbing uses a stochastic selection rule and accepts

all better neighboring solutions. Hence a solution can belong to more than one basin. The attractor function $Attr : s \rightarrow s < LO$ returns the set of local optima for each solution candidate. A local optimum is solely attracted by itself, such that $|Attr(lo)| = 1 \ \ \forall \ lo \in LO$. To determine the probability that a solution s belongs to the basin around the local optimum lo, we assume that all optima are equally strong attractors. The probability then depends on the number of attractors:

$$p_{\text{bas}}(s, lo) = \begin{cases} \frac{1}{|Attr(s)|} & \text{if } s \in B(lo) \wedge lo \in LO \\ 0 & \text{else.} \end{cases} \tag{2.6}$$

2.4 Local Optima Networks

The concept of local optima networks (LONs) has been introduced for the purpose of conducting a statistical network analysis on fitness landscapes (Ochoa, Tomassini, et al., 2008). LONs have been inspired by the study of energy landscapes in chemical physics (Stillinger, 1995). The basic idea is to construct a network in which a landscape's local optima represent the nodes in the network graph. Let our graph representation be $G = (V, A)$. The vertex set $V = LO$ contains all local optima (including the global optima) of the fitness landscape.

A is the set of arcs. There is a weighted arc from a local optimum lo_0 to a local optimum lo_1 if local search can pass from the basin of attraction $B(lo_0)$ to the basin $B(lo_1)$. For such a transition, both basins must be connected such that a solution in $B(lo_1)$ with a high fitness is a neighbor to a solution in $B(lo_0)$ with a lower fitness. The more such connections between two basins exist, the higher is the transition probability between their surrounded local optima, and the higher is the arc weight.

We use the following rule-set for the calculation of arcs: the probability to perform a move from any solution s_0 to another solution s_1 depends on the number of superior neighbors of s_0. First improvement hill climbing (fi-hc) selects a random neighbor with superior fitness. The probability to move from s_0 to s_1 can be calculated as

$$p(s_0, s_1) = \frac{1}{|N_{\text{sup}}(s_0)|}. \tag{2.7}$$

The probability to move from any solution s_0 to the basin of attraction of any local optimum $lo_1 \in LO$ depends on (i) the probability to move from s_0 to a new solution in $B(lo_1)$ and (ii) the probability that the new solution belongs to $B(lo_1)$. Thus, the conditional probability to select any superior neighbor s_1 that belongs to a different basin $B(lo_1)$ can be calculated as

$$p(s_0, s_1 | s_1 \in B(lo_1)) = \frac{1}{|N_{\text{sup}}(s_0)|} \times p_{bas}(s_1, lo_1), \tag{2.8}$$

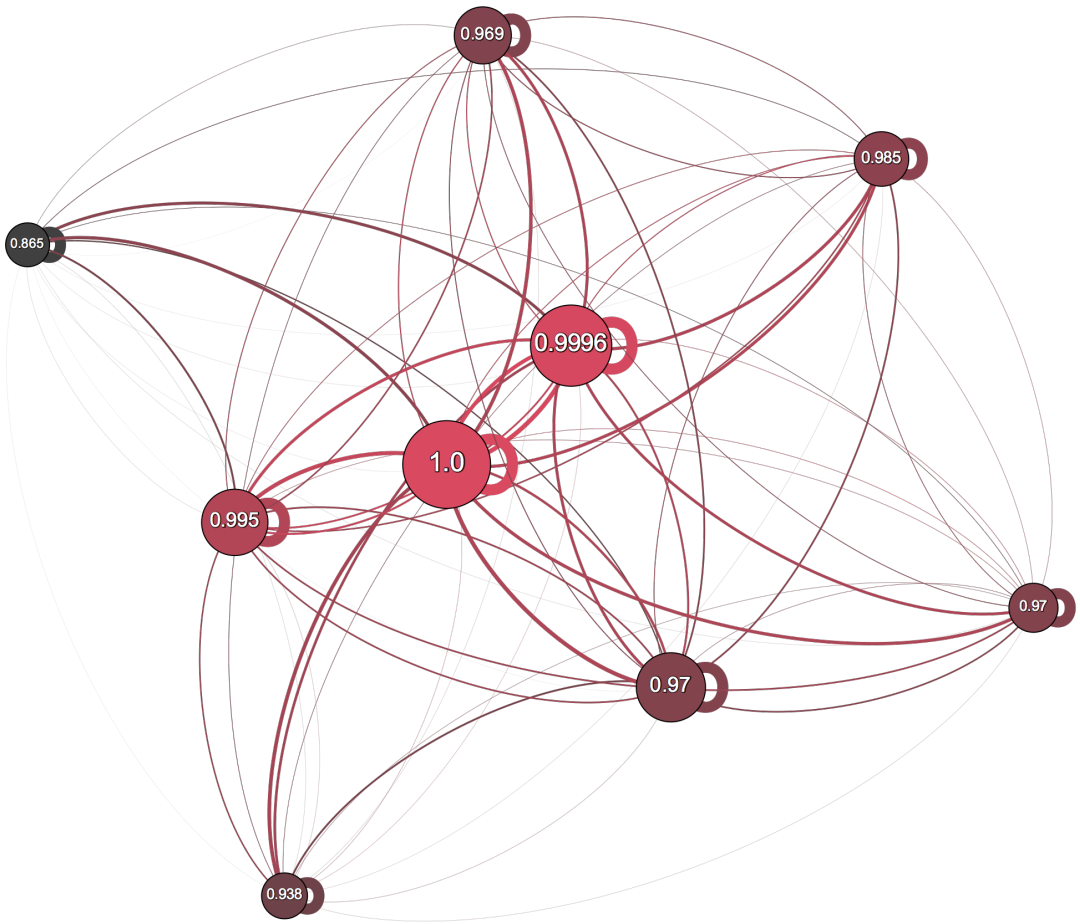

Figure 2.1: The LON of a TSP instance with eight cities. The node label and color indicate the fitness value. The node thickness and arrangement (the higher the PageRank, the more central the node position) represent the PageRank centrality. The thickness of arcs indicates their weight, i.e., the transition probability between the basins of attraction around the local optima.

To determine the probability that the algorithm moves from s_0 to any solution in the basin around lo_1, we iterate over all superior neighbors of s_0 that can be members of $B(lo_1)$ and sum up the conditional probabilities that the algorithm moves from s_0 to one of these solutions:

$$p(s_0, B(lo_1)) = \sum_{s_i:\{s|s\in N_{\text{sup}}(s_0)\wedge s\in B(lo_1)\}} [p(s_0, s_1|s_1 \in B(lo_1))]. \tag{2.9}$$

To calculate the arc weights of the LON, which is the overall probability to move from one basin $B(lo_0)$ to another basin $B(lo_1)$, we sum up the probabilities of all $s_i \in B(lo_0)$ to move from s_i to $B(lo_1)$. Each probability is weighted by the probability that s_i belongs to $B(lo_0)$. The total probability is given by the ratio between this sum and the sum of probabilities of all solutions in the basin $B(lo_0)$:

$$p(B(lo_0), B(lo_1)) = \frac{\sum_{s_i\in B(lo_0)}[p_{\text{bas}}(s_i, lo_0) \times p(s_i, lo_1)]}{\sum_{s_i\in B(lo_0)}[p_{\text{bas}}(s_i, lo_0)]}. \tag{2.10}$$

Using this function, we can calculate the transition probabilities between all basins, and thus the weights of arcs between all local optima in the LON:

$$A_{i,j} = p(B(i), B(j)) \quad \forall\, i, j \in LO. \tag{2.11}$$

A is a stochastic matrix. The total probability of the outgoing links, i.e., the sum of the arc weights is

$$\sum_{i:LO} A_{i,j} = 1 \quad \forall\, j \in LO. \tag{2.12}$$

Local Optima are self-connected by the probability that the search algorithm remains in the current basin. Figure 2.1 shows a LON for a TSP instance with eight cities. The visualization was created with *Gephi* (Bastian et al., 2009).

2.5 PageRank Centrality

The centrality of nodes is a meaningful concept in graph theory and network analysis. With the help of centrality, we can identify important or influential nodes in a network (Borgatti, 2005). Modeling the *world wide web* as a network, where web pages are nodes and links to other web pages are (directed) edges, Google was the first search engine that used the centrality of nodes for assessing the relevance and importance of web sites (instead of simple keyword based search).

To calculate a website's centrality, Google have been using the concept of PageRank (Brin and Page, 1998), which is a variant of the Eigenvector centrality (Bonacich, 2007). The PageRank is based on the model of a surfer who is randomly visiting

webpages. The surfer starts at a random page and then follows one of its links to another page and so forth. After following a number of links, the surfer randomly selects a new page (e.g., by typing the address into the browser's address field). Based on the model of the random surfer, the PageRank value of a website indicates the probability that a page is visited by the surfer. Incoming links from other websites increase the centrality of a page. However, PageRank also takes into account the importance of the linking sources. An incoming link from an important page can increase the relevance of a website more than numerous links coming from rather unimportant sites. Moreover, a website with many outgoing links contributes less to the importance of a linked page than a website with fewer outgoing links (given that both linking pages are equally important). Thus, three factors determine the PageRank of a web page: the number of links a page receives, the number of outgoing links of the linking pages and the PageRank of the linking pages. For detailed information on the calculation of the PageRank, we refer to Franceschet (2011).

To calculate the PageRank of websites, it is first necessary to represent the linkage structure as an adjacency matrix A that models the existence of links between two nodes. Then, A is normalized into a transition matrix Π that contains the probabilities of moving from one node (page) to another. Π is a stochastic matrix, since all rows and columns sum up to 1. To calculate the PageRank of the nodes of a LON, we use the transition matrix A as defined in (2.11). Since A is already a normalized, stochastic matrix, we set $\pi = A$.

A problem of the PageRank model is that a random surfer could get trapped in a node that has no outgoing links. The solution to this problem is that a surfer—having stayed on a page for some time—loses interest and types in a random address in the browser field. The transition matrix for this behavior is the teleportation matrix E that contains identical rows of uniform probability vectors. The convex combination of Π and E results in the transition matrix (also called *Google Matrix*)

$$G = \alpha \times \Pi + (1 - \alpha) \times E. \tag{2.13}$$

α is a damping factor, where $1 - \alpha$ is the probability that a random surfer stops following links and visits a random page instead. A typical value is $\alpha = 0.85$, which says that a surfer chooses a random page after about five link clicks. Since local search uses no perturbation operators (e.g., by randomly selecting a new solution from a wider neighborhood), we ignore this behavior and set $\alpha = 1$.

Then, the PageRank centrality of all nodes (local optima) is given by the vector P, such that

$$P = G \times P. \tag{2.14}$$

P is the Eigenvector of Π, and there is a solution for P if the real square matrix Π has positive components and is irreducible, i.e., it is a strongly connected graph (Frobenius, 1912; Perron, 1907). The transition matrix of a LON has positive transition

probabilities and the nodes are—by our empirical observation—strongly connected for instances of the NK model and the TSP. P contains the PageRank centralities of local optima in the graph. Consequently, we define P_{opt} as the PageRank value of the global optimum s_{opt}.

2.6 Experimental Setting

For our study, we assessed the correlations (R^2 values in linear, univariate regression models) between the performance of local search methods (first improvement hill climbing (fi-hc) and simulated annealing (SA)) and the PageRank centrality of the global optimum for the NK model and the traveling salesman problem (TSP).

2.6.1 Kauffman NK Model

The NK model (Kauffman and Levin, 1987) is a combinatorial optimization problem that is frequently used in organizational theory, complex systems and evolutionary computation. Each instance of the model can be generated by the two parameters N and K. Each solution $s \in S$ consists of N binary decision variables, forming a search space of $|S| = 2^N$ possible states. Every combination of bits is a valid solution and has an assigned score. Each problem instance has one unique global optimum.

Instances of the NK model are tunably difficult as there are adjustable interdependencies between the decision variables: for a given solution, each decision variable contributes to the overall score. The level of contribution depends on the variable's own state and on the state of some other pre-selected variables. These co-variables are randomly assigned in the process of generating a problem instance. The parameter K determines the number of co-variables per decision variable and thus the complexity of an instance. A value of $K = 0$ results in a problem solvable in linear time. $K = N-1$ leads to a maximally difficult problem, where each decision variable can only be set to the optimal value if all other $N - 1$ co-variables are considered. In general, problem difficulty increases with higher values of K. An instance of the NK model contains a single global optimum (Kauffman and Weinberger, 1989).

The distance between two binary solutions $x, y \in S$ is calculated as the Hamming distance $d(x, y) = \sum_{i=0}^{n} |x_i - y_i|$, i.e., the number of bits that are set to different values when comparing two solutions. For both search algorithms (fi-hc as well as SA), we assumed that two solutions x, y are neighbors if their Hamming distance equals one ($d_{\max} = 1$). Then, a local search step flips exactly one bit of the current solution.

2.6.2 Traveling Salesman Problem (TSP)

The TSP is a well-studied and well-understood $\mathcal{NP}$-hard combinatorial optimization problem (Lenstra and Kan, 1981). Given a set of n cities C and a pairwise symmetric distance matrix D_{ij}, the objective of the symmetric TSP is to find the Hamiltonian tour with minimum length. The Hamiltonian tour visits each city exactly once.

We studied instances of a symmetric TSP with Euclidean distances and used an order-based representation for the solutions. This encoding assigns to each city an index from the set $C = \{0, 1, .., n-1\}$ and each tour is a permutation of the indices. For a tour of n cities, there are $(n-1)!$ different permutations. This encoding is redundant since each solution tour is represented by $2N$ genotypes (N rotations and two directions; cf. Choi and Moon, 2008). We reduced the size of the search space by a factor N by removing all rotated solutions. For the TSP, there are two copies of the optimal solution, since we did not remove the symmetric solutions. If there were multiple global optima, we selected one of them randomly as *the* global optimum for the purpose of our study. We normalized and transformed all instances to a maximization problem such that $f(s_{\text{opt}}) = 1.0$ and $f(s_{\text{min}}) = 0.0$.

To be able to represent an instance of the TSP as a fitness landscape, we had to assign a unique integer number to each solution in the search space. In order to number the permutations, we used a scheme introduced by Lehmer (1960). We used standard 2-opt moves as the local search operator. Thus, a single move "deletes two nonadjacent edges of the current tour and then reconnects the two resulting paths into a new tour" (Boese et al., 1994). The distance between two solutions x and y is the minimum number of 2-opt moves necessary to transform x into y.

2.6.3 Experiments

For each problem (NK model and TSP), we generated 500 problem instances. For the NK model, we used $N = 12$ decision variables. The parameter K was randomly chosen with $K \in (2, ...10)$. The considered TSP instances were of size $n = 8$. The size of our problem instances was relatively low, since the computational effort for the experiments grows factorially by the problem size n.

To be able to measure the PageRank centrality of the global optimum in the LON, we extracted the fitness landscape and the LON for each problem instance. Overall, we calculated the following measures:

1. ρ_{nn}: the ruggedness of the fitness landscape measured by the Pearson correlation between the fitness of nearest neighbors (Weinberger, 1990),
2. ρ_{fd}: the deceptiveness of the landscape measured by the Pearson correlation between fitness and distance to the global optimum (Jones and Forrest, 1995),
3. P_{opt}: the PageRank of the global optimum in the LON.

ρ_{nn} and ρ_{fd} were determined by a random sample of 1,000 solutions per problem instance. The PageRank centrality of the peak P_{opt} was calculated using the NetworkX Library (Hagberg et al., 2008). Since we assumed that the algorithms used for our experiments do not perform any random moves through the search space, we set the damping factor to $\alpha = 1$.

For each problem instance, we performed 1,000 independent runs of first improvement hill climbing and simulated annealing (both with random initial solutions). For each instance, we measured search performance by

1. the success rate (p_{s}), which is defined as the percentage of runs that find the global optimum, and
2. the average number of fitness evaluations T_{s} necessary to find the global optimum. For the calculation of $avg(T_{\text{s}})$, we considered only the runs that found the global optimum.

We assessed the quality of the different predictor metrics by the determination coefficient (R^2) of a linear regression model (which is identical to the squared Pearson correlation) between each predictor metric and each measure of empirical performance over all problem instances.

2.7 Results

2.7.1 Performance of Local Search Methods

In a first step, we studied the performance of the two local search heuristics, first improvement hill climbing (fi-hc) and simulated annealing (SA). We measured their performance using success rate p_{s} and average running time $avg(T_{\text{s}})$. For the NK model, the average success rate was $p_{\text{s}} = 0.099$ for fi-hc and $p_{\text{s}} = 0.187$ for SA, resp. For the TSP, $p_{\text{s}} = 0.154$ for fi-hc and $p_{\text{s}} = 0.442$ for SA. Figure 2.2 compares the performance of fi-hc and SA. In this figure, we show the average performance (p_{s} and $avg(T_{\text{s}})$) of SA over the average performance of fi-hc for the NK model (top) and TSP (bottom). Each data point represents the performance of both algorithms for a single problem instance. In general, SA shows a better performance than fi-hc. This is expected, since SA can escape from a local optimum. Along with a higher p_{s}, SA needs more fitness evaluations than fi-hc.

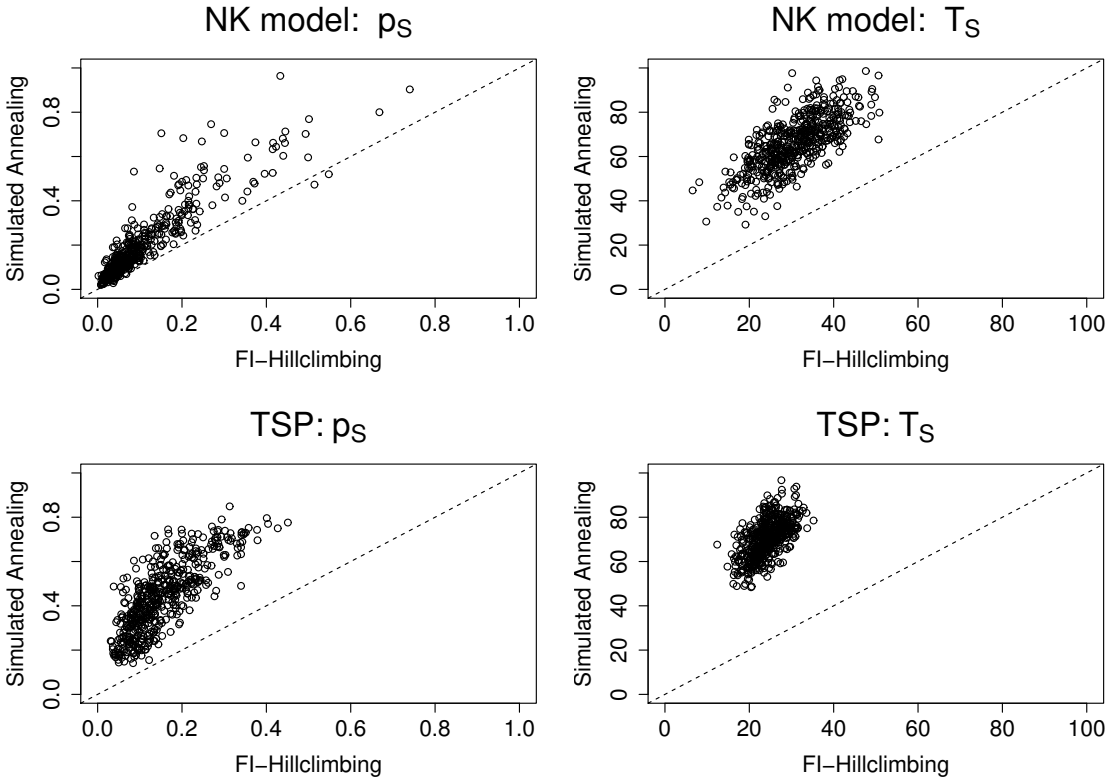

Figure 2.2: The performance of first improvement hill climbing over the performance of simulated annealing for 500 instances of the NK model (top) and the TSP (bottom). In this figure, we show the success rate (left) and the running time (right). Each dot is a single problem instance.

	FI Hill Climbing		**SA**	
Performance	p_s	$avg(T_s)$	p_s	$avg(T_s)$
NK Model				
ρ_{nn}	0.475	0.139	0.587	0.374
ρ_{fd}	0.37	0.101	0.541	0.306
P_{opt}	0.906	0.307	0.916	0.537
TSP				
ρ_{nn}	0.006	0.003	0.001	0.001
ρ_{fd}	0.11	0.043	0.273	0.001
P_{opt}	0.757	0.605	0.646	0.338

Table 2.1: R^2 values for the NK model and the TSP.

2.7.2 Evaluation of Predictive Quality

We studied the quality of our metrics to predict the performance of the two search algorithms in both problem types. Table 2.1 gives an overview of the results of our experiments. We determined the coefficient of determination (R^2) for each combination of predictor metric, performance metric, search method and problem type. For instances of the NK model, the PageRank P_{opt} of the global optimum explains more than 90% of the variance of the success rates of both fi-hc and SA. The correlation between running time and P_{opt} is much lower since only about 30%-50% of the variance is explained. The benchmark predictors nearest-neighbor correlation ρ_{nn} and fitness-distance correlation ρ_{fd} have a much lower predictive power since the $R^2 \approx [0.4, 0.6]$ for success rate and $R^2 \approx [0.1, 0.3]$ for running time.

For TSP, we observe similar numbers: the explanatory power of the standard metrics ρ_{nn} and ρ_{fd} ranges from 0% to 20%. In contrast, the PageRank explains around 65%-75% of the variance of p_s and between 35%-60% of $avg(T_s)$. In summary, the PageRank is a good predictor for the expected success probability of local search approaches like fi-hc and SA.

We will now take a closer look at the predictive quality of PageRank for the different search methods and problem types. Figure 2.3 plots p_s (left) and $avg(T_s)$ (right) over the PageRank of the global optimum for fi-hc (top) and SA (bottom). All results are for the NK model. Each dot represents the average performance

of the particular search method in a single problem instance over the PageRank centrality of the global optimum in the corresponding LON. The plots reveal a high correlation between PageRank and success rate. For running time (right), the correlation is lower. Figure 2.4 plots the same numbers for the TSP instances. We observed the following characteristics of the results:

1. For both the NK model and TSP, the PageRank explains almost the total variance of success rate. This also holds for different values of K in the NK model. This result is due to the fact that the transition matrix in the LON is a stochastic approximation of the moves of local search algorithms in the fitness landscape. The LON graph induces a finite-state Markov Chain, and the PageRank vector corresponds to the chain's stationary distribution (Franceschet, 2011). Since the success rate is the probability of finding the global optimum, the scalar value of the global optimum in the PageRank Vector of the LON approximates the success rate.
2. The PageRank explains more variance of success rate in the NK model than in the TSP. We explain this by the implementation of our TSP model: due to the nature of the symmetric TSP, there are two global optima, reducing the significance of prediction.
3. The prediction of success rate is similar for fi-hc and SA in the NK model. For the TSP, the correlation between PageRank and success rate of fi-hc is slightly higher than for SA. An explanation for this effect could be that the success rate and its variance are in general higher for SA than for fi-hc, leading to a loss of predictive power.
4. In the NK model, the PageRank explains less of the variance of the running time of fi-hc than of SA. We explain this by the higher success rate of SA. Since only successful runs were considered in the run-time evaluation, the sample size of SA is higher than that of fi-hc. Furthermore, the plots indicate that the relationship between PageRank and $avg(T_s)$ is non-linear. In such a case, our linear regression model has only limited significance.
5. In the TSP, the PageRank explains more of the variance of the running time of fi-hc than SA. This could be due to the stochastic nature of SA, which accepts fitness deterioration with a decreasing probability. This behavior is not modeled in the transition probabilities of LONs.

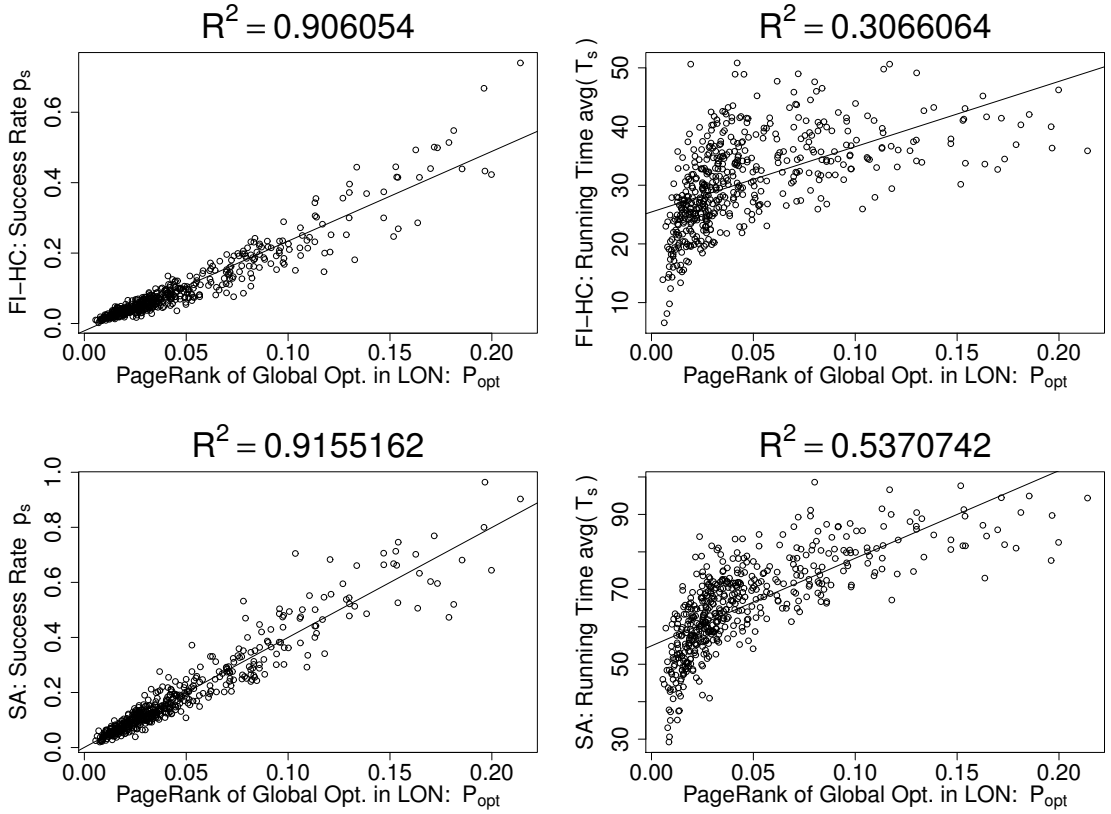

Figure 2.3: The PageRank of the global optimum over the performance of first improvement hill climbing (top) and simulated annealing (bottom) for the *NK model*. We plot success rate (left) and running time (right). Each dot is a single problem instance.

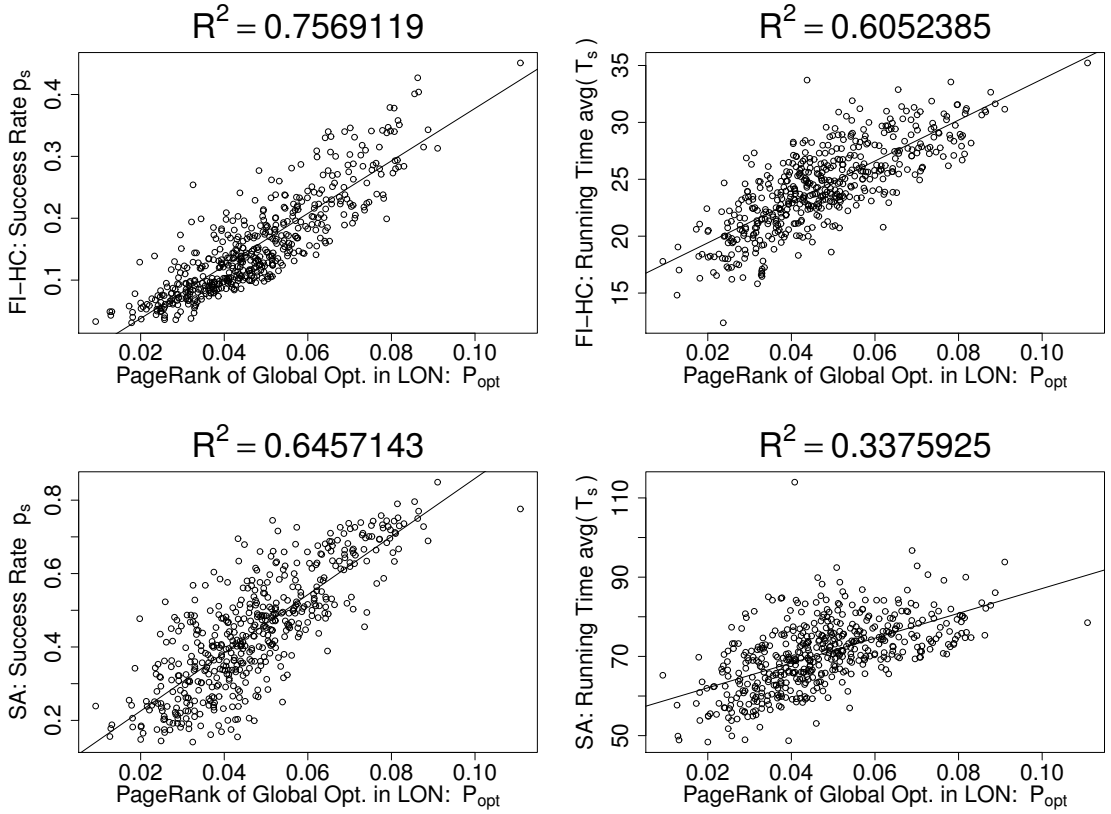

Figure 2.4: The PageRank of the global optimum over the performance of first improvement hill climbing (top) and simulated annealing (bottom) for the *TSP*. We plot success rate (left) and running time (right). Each dot is a single problem instance.

2.8 Conclusions

We suggest that the PageRank of the global optimum in a local optima network is a predictor of the expected performance of local search heuristics like first improvement local search and simulated annealing. A comparison of PageRank with standard predictors like ruggedness and deceptiveness shows a much higher predictive power. For the NK model and TSP, the PageRank explains around 90% of the variance of performance. The explanatory power of the PageRank is also higher than the average length of the shortest paths to the optimum, which has been suggested by Daolio, Verel, et al., 2012 and explains around 50% of the variance.

The PageRank centrality is a good predictor of local search performance since LONs are a good approximation for the computationally expensive extraction of Markov Chain models from fitness landscapes. The PageRank vector provides the stationary distribution of all local optima and, thus, can be used as a proper predictor for finding the high-quality solutions. Furthermore, many combinatorial optimization problems have a "big valley structure" (Boese et al., 1994), where local optima are not randomly distributed in the search space, but rather, are clustered. Thus, most local optima are close to each other and the global optimum is usually not isolated in the search space; rather, it is surrounded by all the local optima. Thus, the PageRank of the global optimum in a LON is a good proxy of performance for many simple search methods based on local search.

A limitation of the PageRank is that it works well if local search is the main search operator; however, with an increasing amount of diversification, its predictive power becomes lower since it does not take into account non-intensifying search steps. Another limitation is the computational effort for the extraction of basins. Even though we used only small problem instances for our experiments, we hope that our findings can be extrapolated onto larger problem instances. In addition, we suggest that future research should focus on studying other metrics which can serve as approximations for PageRank centrality. Finally, we suggest as future work to test our results with the alternative model of LONs with escape edges.

References

Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. In: ***Third International AAAI Conference on Weblogs and Social Media.***: 361–362.

Boese, Kenneth D., Andrew B. Kahng, and Sudhakar Muddu (1994). A new adaptive multi-start technique for combinatorial global optimizations. ***Operations Research Letters,*** 16(2): 101–113.

Bonacich, Phillip (2007). Some unique properties of eigenvector centrality. ***Social Networks,*** 29(4): 555–564.

Borgatti, Stephen P. (2005). Centrality and network flow. ***Social Networks,*** 27(1): 55–71.

Brin, Sergey and Lawrence Page (1998). The anatomy of a large-scale hypertextual web search engine. ***Computer Networks and ISDN Systems,*** 30(1-7): 107–117.

Choi, Sung-Soon and Byung-Ro Moon (2008). Normalization for Genetic Algorithms With Nonsynonymously Redundant Encodings. ***IEEE Transactions on Evolutionary Computation,*** 12(5): 604–616.

Daolio, Fabio, Sébastien Verel, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks and the performance of iterated local search. In: ***Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference - GECCO '12.*** Ed. by Terence Soule. Philadelphia, Pennsylvania, USA: ACM Press: 369.

Franceschet, Massimo (2011). PageRank: standing on the shoulders of giants. ***Communications of the ACM,*** 54(6): 92–101.

Freeman, Linton C. (2004). ***The Development of Social Network Analysis: A Study in the Sociology of Science.*** Empirical Press.

Frobenius, Ferdinand Georg (1912). Ueber Matrizen aus nicht negativen Elementen. ***Sitzungsberichte Preussische Akademie der Wissenschaft, Berlin,*** 456–477.

Glover, Fred (1986). Future paths for integer programming and links to artificial intelligence. ***Computers & Operations Research,*** 13(5): 533–549.

Hagberg, Aric A., Daniel A. Schult, and Pieter J. Swart (2008). Exploring network structure, dynamics, and function using NetworkX. ***Proceedings of the 7th Python in Science Conference (SciPy 2008),*** 11–15.

He, Jun, Colin Reeves, Carsten Witt, and Xin Yao (2007). A note on problem difficulty measures in black-box optimization: classification, realizations and predictability. ***IEEE Transactions on Evolutionary Computation,*** 15(4): 435–43.

Jones, Terry and Stephanie Forrest (1995). Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms. In: ***Proceedings of the Sixth International Conference on Genetic Algorithms.*** San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.: 184–192.

Kallel, L., B. Naudts, and Colin R. Reeves (2001). Properties of Fitness Functions and Search Landscapes. In: ***Theoretical Aspects of Evolutionary Computing.*** Ed. by Leila Kallel, Bart Naudts, and Alex Rogers. Natural Computing Series. Berlin and Heidelberg: Springer: 175–206.

Kauffman, Stuart A. and Simon Levin (1987). Towards a General Theory of Adaptive Walks on Rugged Landscapes. ***Journal of Theoretical Biology,*** 128(1): 11–45.

Kauffman, Stuart A. and Edward D. Weinberger (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. ***Journal of Theoretical Biology,*** 141(2): 211–245.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi (1983). Optimization by simulated annealing. ***Science,*** 220(4598): 671–80.

Laarhoven, P. J. M. and E. H. L. Aarts (1988). ***Simulated Annealing: Theory and Applications.*** Norwell, MA, USA: Kluwer Academic Publishers.

Lehmer, Derrick H. (1960). Teaching combinatorial tricks to a computer. In: ***Proceedings of the Symposium on Applied Mathematics and Combinatorial Analysis.***

Lenstra, Jan Karel and A. H. G. Rinnooy Kan (1981). Complexity of vehicle routing and scheduling problems. ***Networks,*** 11(2): 221–227.

Malan, Katherine M. and Andries P. Engelbrecht (2014). Fitness Landscape Analysis for Metaheuristic Performance Prediction. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Berlin and Heidelberg: Springer: 103–129.

Ochoa, Gabriela, Marco Tomassini, Sébastien Verel, and Christian Darabos (2008). A study of NK landscapes' basins and local optima networks. In: ***Proceedings of the 10th annual conference on Genetic and evolutionary computation - GECCO '08.*** Ed. by Maarten Keijzer. Atlanta, GA, USA: ACM Press: 555–562.

Ochoa, Gabriela, Sébastien Verel, and Marco Tomassini (2010). First-improvement vs. best-improvement local optima networks of nk landscapes. In: ***PPSN'10: Proceedings of the 11th International Conference on Parallel Prob-***

lem Solving from Nature. Ed. by Robert Schaefer, Carlos Cotta, Joanna Kolodziej, and Günter Rudolph. Vol. I. Kraków, Poland: Springer: 104–113.

Perron, Oskar (1907). Zur Theorie der Matrices. ***Mathematische Annalen 1,*** 64(1): 248–263.

Pitzer, Erik and Michael Affenzeller (2012). A Comprehensive Survey on Fitness Landscape Analysis. In: ***Recent Advances in Intelligent Engineering Systems.*** Ed. by János Fodor, Ryszard Klempous, and Carmen Paz Suárez Araujo. Vol. 378. Studies in Computational Intelligence. Berlin and Heidelberg: Springer: 161–191.

Reidys, Christian M. and Peter F. Stadler (2002). Combinatorial Landscapes. ***SIAM Review,*** 44(1): 3–54.

Rothlauf, Franz (2011). ***Design of modern heuristics: Principles and application.*** Berlin and Heidelberg: Springer.

Stadler, Peter F. (1996). Landscapes and their correlation functions. ***Journal of Mathematical Chemistry,*** 20(1): 1–45.

Stillinger, Frank H. (1995). A Topographic View of Supercooled Liquids and Glass Formation. ***Science,*** 267(5206): 1935–1939.

Weinberger, Edward D. (1990). Correlated and uncorrelated fitness landscapes and how to tell the difference. ***Biological cybernetics,*** 336: 325–336.

Wright, Sewall (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: ***Proceedings of the 6th International Congress of Genetics.*** Ed. by Donald F. Jones. Ithaca, New York: Morgan Kaufmann Publishers Inc.: 356–366.

Chapter 3

Determining the Difficulty of Landscapes by PageRank Centrality in Local Optima Networks

Sebastian Herrmann

Abstract

The contribution of this study is twofold: First, we show that we can predict the performance of iterated local search (ILS) in different landscapes with the help of local optima networks (LONs) with escape edges. As a predictor, we use the PageRank centrality of the global optimum. Escape edges can be extracted with lower effort than the edges used in a previous study. Second, we show that the PageRank vector of a LON can be used to predict the solution quality (average fitness) achievable by ILS in different landscapes.

3.1 Introduction

Local optima networks (Ochoa, Tomassini, et al., 2008) are a novel approach to study the structure of optimization problems by using complex network analysis. A local optima network (LON) is a compressed representation of a combinatorial fitness landscape. Mathematically, a LON is a graph in which the vertices are the search space's local optima. The edges are modeled to reflect the transitions between the local optima and are weighted by transition probabilities. Three types of LON models have been introduced so far in order to represent different search operators: edges with basin transition probabilities for the trajectory of hill climbing algorithms (Ochoa, Verel, and Tomassini, 2010), escape edges for iterated local search (Verel et al., 2012) and LONs for partition crossover (Ochoa, Chicano, et al., 2015).

An application of fitness landscape analysis is to predict the performance of a search algorithm in a particular problem instance (search difficulty; Lu et al., 2014; Malan and Engelbrecht, 2014). Network features of LONs can in particular capture the search difficulty of landscapes (Ochoa, Verel, Daolio, et al., 2014). Among the different network metrics, it was shown that the shortest path to the global optimum (Daolio, Verel, et al., 2012) and its PageRank centrality (Chapter 2) are good predictors for local search-based methods. The PageRank of the global optimum predicts the empirical success rate of hill climbing with approx. 90% accuracy. Success rate is the probability that a search algorithm hits the global optimum. The explanation for this high correlation is that LONs are an approximate Markov Chain representation of fitness landscapes. As the PageRank vector of the nodes represents their stationary distribution in the stochastic process, the PageRank of the global optimum approximates the probability that a search algorithm finds it.

A major limitation of previous studies is that success rate is just one alternative of measuring search performance. Most heuristics are—in the first place—not designed to solve a problem exactly. Instead, they make use of an existing trade-off between computation time and solution quality, with the goal to generate a yet suboptimal, but acceptable solution. Thus, predicting the expected solution quality is a relevant issue. Another limitation of the PageRank study (Chapter 2) is that predicting success rates has only been tested for the LON model with basin transition probabilities. However, the extraction procedure for this model is computationally expensive since it requires to compute all the local optima basins.

In this paper, we take up previous efforts on determining search difficulty of fitness landscapes with local optima networks. We present a method to predict both success rate and solution quality (average fitness) with LONs with escape edges (LON_{ee}; Verel et al., 2012) using PageRank centrality. The escape edges can be extracted with much lower computational effort than the basin transition probabilities. To predict the average fitness, we make use of the fact that the calculation of the Page-

Rank results in a vector which covers the stationary distribution of the whole search space. We combine these probabilities with the distribution of fitness in the search space to calculate an expected value for the fitness. Using this value, we can predict the fitness that is achieved on average by iterated local search in different instances of the Kauffman NK family of landscapes.

Our paper is structured as follows: In Section 3.2, we describe the search heuristic of which we aimed to predict its performance in our experiments, i.e. iterated local search. In Section 3.3, we give a short introduction to fitness landscapes and provide a formal definition. In Section 3.4, we define LONs with escape edges. In Section 3.5, we shortly describe the concept of PageRank centrality. In Section 3.6, we describe our experimental design and the search space used (NK family of landscapes). We present our results in Section 3.7 and draw our conclusions in Section 3.8.

3.2 Iterated Local Search

Iterated local search (ILS; Lourenço et al., 2003) has so far been used in a variety of studies on local optima networks (Daolio, Verel, et al., 2012; Ochoa, Veerapen, et al., 2016; Verel et al., 2012) . The concept of ILS is used in many practically relevant search methods, e.g. the Chained Lin Kernighan heuristic (Applegate et al., 2003; Lin and Kernighan, 1973). This Section gives a brief review on the algorithm.

ILS combines the concept of intensification by local search with a number of perturbation steps to obtain some diversification. During intensification, heuristics focus their search on promising areas of the search space, whereas during diversification, new areas are explored (Rothlauf, 2011). Algorithm 3.1 describes the search method in pseudo code. Given a search space of valid solutions S for an optimization problem, we assign a fitness value to each $s \in S$ by the function $f : S \to \mathbb{R}_{\geq 0}$. ILS starts with a randomly selected solution $s_0 \in S$. Then, the algorithm performs a hill climbing procedure with best improvement as selection rule (Algorithm 3.2): from the neighborhood $N(s)$, the best solution with higher fitness is selected. This requires a scan of the whole neighborhood of s. The neighborhood $N(s)$ is the set of solutions that can be reached by performing an incremental change to s. This hill climbing procedure is then repeated until it reaches a local optimum s^*, i.e no further improvement is possible. Then, ILS performs a diversification step by applying a limited perturbation to the local optimum, resulting in s'. As a next step, hill climbing is applied from s', until the next local optimum $s^{*'}$ is reached. If the new local optimum $s^{*'}$ has higher fitness, the algorithm has "escaped" to a new local optimum, and the change is accepted. Otherwise, another perturbation is applied to s^*. This procedure is repeated until a termination condition is met, e.g. a fixed number of escapes without any further improvement.

Algorithm 3.1: Iterated Local Search (ILS)

Require: Solution space S,
Fitness function $f(S)$,
Neighborhood function $N(S)$,
Stopping Threshold t

1: $i \leftarrow 0$
2: Choose initial random solution $s_0 \in S$
3: $s^* \leftarrow$ hillClimbBI(s_0)
4: **repeat**
5: $s' \leftarrow$ perturbation(s^*)
6: $s^{*'} \leftarrow$ hillClimbBI(s')
7: **if** $f(s^{*'}) > f(s^*)$ **then**
8: $s^* \leftarrow s^{*'}$
9: $i \leftarrow 0$
10: **end if**
11: $i \leftarrow i + 1$
12: **until** i $\geq$ t
13: **return** s^*

Algorithm 3.2: Best Improvement Hill Climbing (hillClimbBI)

Require: Solution space S,
Fitness function $f(S)$,
Neighborhood function $N(S)$,
Initial solution s_0

1: i $\leftarrow 0$
2: **repeat**
3: choose x s.t. $f(x) = max_{x \in N(s_i)}(f(x))$
4: **if** $f(x) > f(s_i)$ **then**
5: $s_{i+1} \leftarrow x$
6: **else**
7: $s_{i+1} \leftarrow s_i$
8: **end if**
9: $i \leftarrow i + 1$
10: **until** s_i is local optimum: $\{s \in N(s_i) \mid f(s) < f(s_i)\} = \emptyset$
11: **return** s_i

3.3 Fitness Landscape Analysis

3.3.1 Concept

The notion of fitness landscapes originated from evolutionary biology (Wright, 1932). The idea is that there is a fitness for each genome of the different species, and by the distances between the genomes a landscape is shaped in which the fitness is the height. In combinatorial optimization, a motivation to analyze fitness landscapes is to gain a better understanding of algorithm performance on a related set of problem instances. Landscape characteristics reflect the search difficulty for a variety of heuristics (Lu et al., 2014; Malan and Engelbrecht, 2014), thus problem specific knowledge can help construct better search methods (Pitzer and Affenzeller, 2012). In this Section, we provide a short explanation of important fundamentals of fitness landscape analysis.

3.3.2 Neighborhood Structure

In combinatorial optimization, a fitness landscape is a triplet of the search space S, the fitness function f, and the neighborhood structure $N(S)$. S contains all valid solution candidates. The fitness function $f : S \to \mathbb{R}_{\geq 0}$ assigns a fitness value to each $s \in S$. The neighborhood function $N : S \to \mathcal{P}(S)$ assigns a set of neighbors $N(s)$ to every $s \in S$ [1]. The neighborhood structure determines the position of each s in the landscape (Reidys and Stadler, 2002). To determine the neighbors, we assume a distance function between all pairs of solutions s_0 and s_1 as

$$d : (s_0, s_1) \to \mathbb{N}_0, s_0 \wedge s_1 \in S. \tag{3.1}$$

The distance function depends on the search operator used. Starting from a solution s_0, local search uses a small distance $d_{max} = d_{s_0,s_1}$ to choose a new solution s_1. We define the neighborhood function as

$$N : s_0 \to \{s_1 \in S : \ s_1 \neq s_0 \wedge 0 < d(s_0, s_1) \leq d_{\max}\}. \tag{3.2}$$

Iterated local search varies $d_{\max}$ during run-time to obtain higher diversification and to escape from a local optimum by using perturbation steps. This results in changes in the landscape during the run of the algorithm, and makes static analyses more difficult. Despite that, it is generally accepted to study a landscape defined by a fitness function and one or several induced distances (Pitzer and Affenzeller, 2012).

[1] The reader should be aware that different definitions of $N(s)$ are possible. For example, variable neighborhood search iteratively switches between different neighborhoods.

3.3.3 Definition of Local Optima

A fitness landscape can have one or more local optima. A local optimum is a solution that has no superior neighbors. For a maximization problem, we define a function

$$N_{\text{sup}}(s) = \{n \in N(s) : f(n) > f(s)\} \tag{3.3}$$

which returns the neighbors of a solution $s \in S$ that have a superior fitness. As local optima have no superior neighbors, the set

$$LO = \{lo \in S : N_{\text{sup}}(lo) = \emptyset\}. \tag{3.4}$$

contains all the local optima, which also includes the global optimum.

3.3.4 Basins of Attraction

The basin of attraction is the set of solution candidates from which local search converges to a particular local or global optimum. The definition of a basin depends on the selection rule of the hill climbing algorithm (Ochoa, Verel, and Tomassini, 2010). Our implementation of ILS used best improvement hill climbing, (Algorithm 3.2) which accepts only the best of all superior neighbor solutions. Consequently, each solution in the search space belongs to the basin around exactly one local optimum and the basins form a partition set of the search space. These basins are referred to as unconditional basins (Pitzer and Affenzeller, 2012). The function

$$B : lo \rightarrow \mathcal{P}(S \backslash LO) \tag{3.5}$$

assigns a subset from the power set over the solutions in the search space to each local optimum $lo \in LO$, which is the basin around lo. In the following, we use B as a function which returns the unconditional basins.

3.3.5 Landscape Features

Structural features of fitness landscapes are often used to predict the performance of algorithms. A well-elaborated collection of such features is given by Kallel et al. (2001). Two of the frequently used features are ruggedness (Weinberger, 1990) and deceptiveness (Jones and Forrest, 1995). The idea of ruggedness is that the smoother the landscape is, the easier it is to search the landscape in order to find the global optimum. Ruggedness is a consequence of modality, i.e., the presence of local optima. The higher the number of local optima, the more rugged is the landscape. Ruggedness is usually measured as the correlation of fitness values between pairs of neighboring solutions ρ_{nn} (nearest-neighbor correlation). The usual way to calculate

ρ_{nn} is to perform a random walk across the search space and draw samples of the fitness of solution pairs that are neighbors.

A landscape is deceptive if the structure of the search space leads away from the global optimum. A measurement for deceptiveness is the correlation between the fitness of the solutions and their distance to the global optimum (Jones and Forrest, 1995). For the calculation of the fitness-distance correlation ρ_{fd}, the global optimum must be known in advance. A random sample of solutions is drawn and ρ_{fd} is determined between their fitness and their distance to the global optimum. A strongly misleading landscape with $\rho_{fd} \approx -1$ is often referred to as a trap.

3.4 Local Optima Networks with Escape Edges

LONs have been inspired by the study of energy landscapes in chemical physics (Ochoa, Tomassini, et al., 2008; Stillinger, 1995). A LON is a graph representation of a fitness landscape. A graph G consists of vertices and edges $G = (V, E)$. The vertex set V contains all the local optima of the fitness landscape. E contains the edges that model transitions between the local optima. In the case of LONs, the edges are directed and weighted. The existence and weight of edges depend on the trajectory of the search algorithm. To model the dynamics of iterated local search, Verel et al. (2012) introduced the concept of escape edges.

Escape edges are defined according to the distance function d of the fitness landscape (minimal number of moves between two solutions). There is an integer $D > 0$ that is depicted as the distance that the ILS search applies to perform a perturbation step. There is a directed edge $E_{xy} > 0$ from local optimum lo_x to lo_y if there exists a solution s such that

$$d(s, lo_x) \leq D \wedge s \in B(lo_y). \tag{3.6}$$

The weight of this edge is the probability that ILS escapes from lo_x to lo_y. It is the number of solutions within reach of the perturbation step and which belong to the basin around lo_y. Since our implementation used best improvement hill climbing, the function B here returns the unconditional basin of attraction[2]. The number of solutions with an opportunity to escape is normalized by the total number of solutions within the distance D:

$$E_{xy} = \frac{|\{s \in S \,|\, d(s, lo_x) \leq D \wedge s \in B(lo_y)\}|}{|\{s \in S \,|\, d(s, lo_x) \leq D\}|}. \tag{3.7}$$

[2]Unlike in the case of edges for basin transition probabilities (cf. Chapter 2), it is not necessary to evolve all the basins, which makes the extraction of escape edges less expensive. The local optimum for any solution can be easily determined by applying the hill climbing algorithm, starting at the solution at hand.

3.5 PageRank Centrality

The centrality of nodes is a concept of network analysis to identify important or influential nodes (Borgatti, 2005). Google were the first to assess the relevance and importance of web sites by their centrality in the linkage structure of the web (as an enhancement to keyword based search). To this purpose, they have been using PageRank centrality (Brin and Page, 1998), which is a variant of the Eigenvector centrality (Bonacich, 2007). It is based on the model of a user who surfs the web by randomly clicking links. The PageRank value of a website reflects the probability that the surfer currently is on this website. There are three factors determining the PageRank of a web page: the number of links a page receives, the number of outgoing links of the linking pages, and the PageRank of the linking pages. Thus, PageRank is a recursively defined concept. For detailed information on the notion and application of PageRank, we refer to Franceschet (2011).

To calculate the PageRank of websites (here: local optima), we need a transition matrix Π, which is a stochastic matrix of the linkage structure matrix E with all rows and columns normalized to sum up to 1. The transition matrix E of a LON is defined by the edge weights, as shown in statement 3.7. E is a normalized, stochastic matrix, and we can set $\Pi = E$.

A parameter of PageRank is the damping factor α, which reflects the fact that a random surfer may—instead of following links—visit a totally random page at some point. This probability is represented by $1 - \alpha$. A typical value is $\alpha = 0.85$, which says that a surfer chooses a random page after about five link clicks. Hill climbing algorithms do not make any jumps in the search space, thus $\alpha = 1.0$ was set for analyzing the LONs with basin transition probabilities in previous work (Chapter 2). In the case of ILS, there is a perturbation operator. However, the escape edges in the corresponding LON model already reflect this behavior. Consequently, we set $\alpha = 1.0$ for our analysis.

Then, the PageRank centrality of all nodes (local optima) is given by the vector P, which is the Eigenvector of Π:

$$P = \Pi \times P. \tag{3.8}$$

If Π is a strongly connected graph, there exists a solution for P (Frobenius, 1912; Perron, 1907). These conditions are fulfilled in our case, since negative probabilities are impossible by our definition of the LONs transition matrix. In addition, we did not observe any disconnected components in our LONs. The vector P contains the PageRank centralities of all the local optima in the search space's LON. Consequently, we define P_{opt} as the PageRank value of the global optimum.

3.6 Experiment

3.6.1 Search Space: NK Model

For our experiment, we used the well-known Kauffman NK model (Kauffman and Levin, 1987), which is a family of combinatorial optimization problems from the class of pseudo-boolean functions. Each instance of the model can be generated by the two parameters N and K. Each solution $s \in S$ consists of N binary decision variables, forming a search space of $|S| = 2^N$ possible states. The fitness function

$$f_{NK} : [0,1]^N \rightarrow [0,1] \tag{3.9}$$

assigns a score to every combination of bits. It is the sum of N sub-functions, which assign a fitness for each bit i, depending on the state of bit i and the states of K other bits

$$f_i : [0,1]^{K+1} \rightarrow [0,1]. \tag{3.10}$$

The total fitness $f_{NK}(s)$ is the average of the values of the N sub-functions. All function values are normalized between 0 and 1, with 1.0 as the fitness of the global optimum. The parameter K determines the number of co-variables per decision variable and thus the complexity of an instance (epistasis). A value of $K = 0$ results in a problem solvable in linear time. $K = N - 1$ leads to a problem where each decision variable can only be set to the optimal value if all other $N - 1$ co-variables are considered. Even though it is commonly accepted that a higher level of epistasis lead to higher search difficulty of landscapes, it is only a rough measure for difficulty. Landscapes with an identical level of epistasis can have a significant variety of search difficulty. Our results on the performance of ILS in Section 3.7.1 underpin this assumption.

The distance between two binary solutions $x, y \in S$ is calculated by the Hamming distance $d(x,y) = \sum_{i=0}^{n} |x_i - y_i|$, i.e., the number of bits that are set to different values when comparing two solutions. For the hill climbing procedure in ILS, we assumed that two solutions x, y are neighbors if their Hamming distance equals one ($d_{\max} = 1$). Thus, a local search step flips exactly one bit of the current solution. As perturbation operator in ILS, we flipped two bits in one step.

3.6.2 Implementation

The objective of our experiment is to predict the success rate p_s and the average fitness $avg(f)$ achieved by ILS in a variety of different search spaces. We generated 300 instances in total of the NK model with $N = 15$ bits. To test different levels of epistasis, we used 100 instances each for $K \in \{2, 7, 12\}$. A search space contains

$2^{15} = 32,768$ solutions, which is small, but manageable for our analysis. For each instance, we extracted the fitness landscape and the LON with escape edges and calculated the following features:

1. P: the PageRank Vector, and P_{opt}: the PageRank of the global optimum,
2. F: the vector containing the fitness values of all the local optima,
3. ρ_{nn}: the ruggedness of the fitness landscape, measured by the Pearson correlation between the fitness of nearest neighbors (Weinberger, 1990) and
4. ρ_{fd}: the deceptiveness of the landscape, measured by the Pearson correlation between fitness and distance to the global optimum (Jones and Forrest, 1995).

For each problem instance, we performed 1,000 independent runs of ILS. The initial solutions were randomly selected. As a stopping threshold for ILS, we limited the running time by a reasonable number of function evaluations, which was 1/5th of the search space size (Daolio, Verel, et al., 2012). We examined two relationships: first, we studied how the PageRank of the global optimum P_{opt} predicts the success rate of ILS in the different problem instances. The purpose of this approach is to confirm that the findings from the previous study on LONs with basin transition probabilities (Chapter 2) also hold for the LON_{ee} model. We have also compared the PageRank to the average number of function evalutions $avg(t)$ that were performed in those runs in which the global optimum was found. Second, we aimed to predict the solution quality of ILS (the average fitness) by the PageRank vectors of the LONs. To achieve this, we used the PageRank vector P and the fitness vector F of the local optima for each search space. We expect that P provides a stationary distribution over the whole search space and is a probability vector, s.t. $\sum P = 1$. We use these probabilities to calculate an average of the fitness of the local optima as given by F, weighted by their stationary probability:

$$E[f] = P \times F. \tag{3.11}$$

The result is a scalar value which we call the expected fitness $E[f]$ achieved by ILS in a distinct search space. We calculated $E[f]$ for all the problem instances. We assessed the predictive power of the different predictor metrics by the determination coefficient R^2 from a univariate, linear regression model. As a benchmark, we have also calculated the R^2 between the performance measures and the classical metrics from fitness landscapes analysis (ruggedness and deceptiveness).

We implemented our generator for NK landscapes, the extraction procedure for LONs and the ILS algorithm in the Java programming language. For computation, we utilized 20 nodes from an HPC cluster called *Mogon* with 64 cores and 256 GB of RAM each. To calculate the PageRank values of the nodes, we used the *NetworkX Library* (Hagberg et al., 2008). Our statistical analysis was conducted using the *R Framework* (R Development Core Team, 2009).

3.7 Results

3.7.1 Empirical Performance of ILS

As a pre-test of our experiments, we examined the performance of ILS by success rate p_s, average fitness $avg(f)$ and the number of fitness function evaluations to find the global optimum $avg(t)$. The results can be obtained from Figure 3.1. In the landscapes with low epistasis, we can see that ILS could easily find the global optimum in the majority of the search spaces. With increasing value of the exogenous parameter for epistasis K, the average success rate decreases, and so does the average score achieved by ILS. The number of fitness evaluations necessary to find the global optimum increases with the epistasis: more epistasis lead to a higher modality, i.e., the number of local optima. The more local optima are in a search space, the more perturbations are necessary to find the global optimum. All of these observations are as expected: higher epistasis leads to a higher search difficulty, and thus lead to a lower success rate, a lower average fitness and longer running times. We can also see that there is as high variance of all the performance measures within the different classes of K, indicating that epistasis has only limited explanatory power for search difficulty.

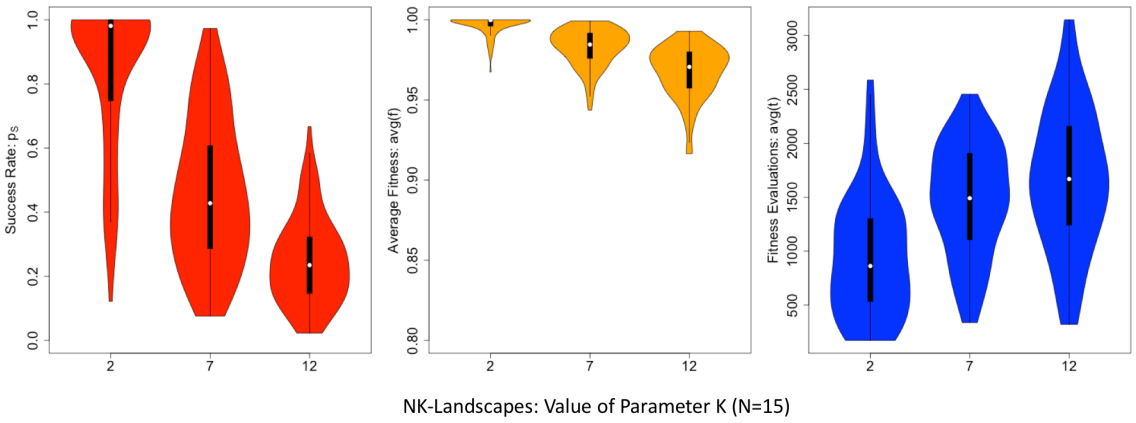

Figure 3.1: The three performance measures of ILS over the epistasis K of the NK landscapes: success rate (left), average fitness (middle) and running time by the number of fitness evaluations (right).

3.7.2 Prediction of Success Rate and Average Fitness

To assess the predictive quality, we calculated all coefficients of determination (R^2) for each combination of predictor metric and performance measure. We have also made separate calculations for different levels of epistasis. The results for all combinations can be obtained from Table 3.1. In Figure 3.2, we have plotted the performance of ILS (success rate, average fitness and running time) over the three predictor metrics. Each dot in the plot represents one search space.

As a first step, we take a look at the standard metrics that are frequently used in literature. Over all K, ruggedness and deceptiveness each can explain around 55% of the variance of success rate and 44%/35% of average fitness. This is an intermediate statistical correlation. This correlation becomes weaker the higher the level of epistasis is. In the cases where $K \in \{7, 12\}$, the traditional metrics fail to explain any variance in the performance of ILS over all metrics. An explanation for this could be that the landscapes with high epistasis have a low variance in their ruggedness ρ_{nn} and deceptiveness ρ_{fd}. A low variance in the regressor variables then results in a low R^2.

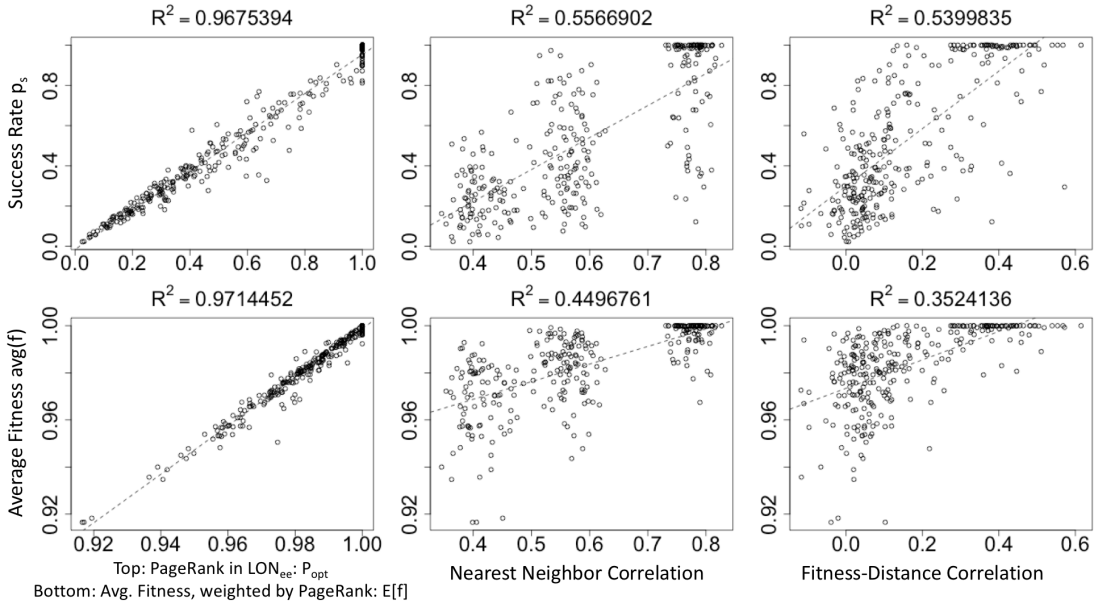

Figure 3.2: The performance of ILS (top: success rate, bottom: average fitness) over the predictor metrics (PageRank/PageRank weighted avg. fitness, ruggedness and deceptiveness, from left to right). Each of the dots represents a single problem instance.

We will now take a look at the prediction by PageRank centrality in LONs. Our expectations were a high correlation between the PageRank of the global optimum and success rate, as well as between the average score of the local optima (weighted by their PageRank) and the fitness achieved by ILS on average. We observed the following patterns in our results:

- The PageRank of the global optimum P_{opt} explains almost 97% of the success rate of ILS p_s. Obviously, the PageRank as obtained from the LON model with escape edges is a good indicator of the search difficulty for ILS.
- The expected average fitness $E[f]$ explains almost 97% of the average score (solution quality) achieved by ILS $avg(f)$. Thus, the PageRank vector of a LON P seems to reflect the dynamics of ILS in terms of the probability to achieve a certain state.
- These results are robust for different levels of epistasis K. For high values of K, the R^2 is slightly reduced for both predictors, but it is still a very strong correlation and significantly better than traditional landscape metrics. Thus, the LON with escape edges nearly approximates the dynamics of ILS.
- The number of fitness evaluations needed to locate the global optimum $avg(t)$ is only weakly correlated to all of the predictor metrics. Surprisingly, in the case of medium and high epistasis, the PageRank seems to predict 30-40% of the variance of running time. An explanation for this could be that in cases of high epistasis, the basins are very small. In landscapes with small basins, the running time of ILS is dominated by perturbation steps, and there is nearly no hill climbing. Since the escape edges in the LON_{ee} map these perturbations, the LON_{ee} perfectly matches the stochastic process of ILS in such cases. Then, the LON is more likely to reflect the running time than in cases where ILS needs to spend many function evaluations for the hill climbing procedure.

In summary, we have shown that the PageRank of the global optimum in LONs with escape edges perfectly predicts the search difficulty of landscapes for ILS, i.e., the empirical success rate. Moreover, we found that the stationary distribution of the PageRank vector over all local optima is useful to make predictions about the solution quality when running ILS in a certain search space[3]. Both predictions work for all levels of epistasis, which is a clear advantage to the concepts of ruggedness and deceptiveness.

[3]We have also replicated this result to predict the average fitness achieved by local search with LONs with basin transition probabilities. Results are available from the authors upon request.

Performance	Predictor	$\forall K$	K = 2	K = 7	K = 12
	NN Correl.: ρ_{nn}	0.5567	0.0054	0.0019	0.0134
Success Rate: p_s	FD Correl.: ρ_{fd}	0.5400	0.1076	0.0751	0.0897
	PageRank: P_{opt}	0.9675	0.9340	0.9683	0.8870
	NN Correl.: ρ_{nn}	0.4497	0.0022	0.0235	0.0012
Average Fitness: $avg(f)$	FD Correl.: ρ_{fd}	0.3523	0.0785	0.0075	0.0295
	PageRank Weightd.Fitn.: E[f]	0.9714	0.8614	0.9668	0.9554
	NN Correl.: ρ_{nn}	0.2090	0.0027	0.0019	0.0007
Running Time: $avg(t)$	FD Correl.: ρ_{fd}	0.1455	0.0006	0.0089	0.0069
	PageRank: P_{opt}	0.0006	0.1444	0.3081	0.3950

Table 3.1: R^2 values for the different performance measures and predictor metrics.

3.8 Conclusions

In this study, we have contributed to recent research on predicting search difficulty of landscapes with the help of local optima networks and the metrics from the network analysis framework. We have shown that the PageRank centrality of local optima can be used to predict the average fitness and success rate achieved by search heuristics. This works because LONs are an approximation of the fitness landscape's Markov Chain and the PageRank reflects the stationary distribution of the states in this chain. Other than classical metrics of landscape analysis, this method is robust against different levels of epistasis, i.e., the number of interdependencies between the decision variables. The PageRank of the global optimum also predicts the running time with limited accuracy in landscapes with high epistasis. Thus, LONs can be used as a tool to draw conclusions on the structure of problems. We have shown that predictions made with PageRank in a previous study are applicable with a LON model that can be computed in reasonable time. A practical application of these findings could be in the selection of problem instances for benchmark purposes. In benchmarks, test reliability is an important criterion, and the PageRank could be easily used to select instances that guarantee a uniform search difficulty or expected fitness outcome.

A limitation of our study is the size of the problem instances used. Even though we are convinced that our results extrapolate to larger instances, it would be interesting to perform further examinations on this, e.g. by sampling the local optima instead of evolving the whole search space. Another limitation is of fundamental nature: even though we have not made further tests in this assumption, we think that it is not possible to make general statements on the performance of an algorithm with an arbitrary LON model. Instead, the LON model must match the dynamics of the search method. For example, in the case of ILS, the escape edges must consider the distance of the perturbation step. However, this study provides evidence that the prediction by PageRank works across different LON models in combination with a distinct search heuristic. For future work, we suggest to conduct further analyses on problem structure by LONs. Apart from larger instances, it would also be worthwhile to study if it is possible to make assumptions on the search difficulty of landscapes for a variety of search heuristics with the help of LONs.

References

Applegate, David, William Cook, and Andre Rohe (2003). Chained Lin-Kernighan for Large Traveling Salesman Problems. ***INFORMS Journal on Computing,*** 15(1): 82–92.

Bonacich, Phillip (2007). Some unique properties of eigenvector centrality. ***Social Networks,*** 29(4): 555–564.

Borgatti, Stephen P. (2005). Centrality and network flow. ***Social Networks,*** 27(1): 55–71.

Brin, Sergey and Lawrence Page (1998). The anatomy of a large-scale hypertextual web search engine. ***Computer Networks and ISDN Systems,*** 30(1-7): 107–117.

Daolio, Fabio, Sébastien Verel, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks and the performance of iterated local search. In: ***Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference - GECCO '12.*** Ed. by Terence Soule. Philadelphia, Pennsylvania, USA: ACM Press: 369.

Franceschet, Massimo (2011). PageRank: standing on the shoulders of giants. ***Communications of the ACM,*** 54(6): 92–101.

Frobenius, Ferdinand Georg (1912). Ueber Matrizen aus nicht negativen Elementen. ***Sitzungsberichte Preussische Akademie der Wissenschaft, Berlin,*** 456–477.

Hagberg, Aric A., Daniel A. Schult, and Pieter J. Swart (2008). Exploring network structure, dynamics, and function using NetworkX. ***Proceedings of the 7th Python in Science Conference (SciPy 2008),*** 11–15.

Jones, Terry and Stephanie Forrest (1995). Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms. In: ***Proceedings of the Sixth International Conference on Genetic Algorithms.*** San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.: 184–192.

Kallel, L., B. Naudts, and Colin R. Reeves (2001). Properties of Fitness Functions and Search Landscapes. In: ***Theoretical Aspects of Evolutionary Computing.*** Ed. by Leila Kallel, Bart Naudts, and Alex Rogers. Natural Computing Series. Berlin and Heidelberg: Springer: 175–206.

Kauffman, Stuart A. and Simon Levin (1987). Towards a General Theory of Adaptive Walks on Rugged Landscapes. ***Journal of Theoretical Biology,*** 128(1): 11–45.

Lin, S. and B. W. Kernighan (1973). An Effective Heuristic Algorithm for the Traveling-Salesman Problem. ***Operations Research,*** 21(2): 498–516.

Lourenço, Helena R., Olivier C. Martin, and Thomas Stützle (2003). Iterated Local Search. In: ***Handbook of Metaheuristics.*** Boston: Kluwer Academic Publishers: 320–353.

Lu, Guanzhou, Jinlong Li, and Xin Yao (2014). Fitness Landscapes and Problem Difficulty in Evolutionary Algorithms: From Theory to Applications. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Emergence, Complexity and Computation. Berlin and Heidelberg: Springer: 133–152.

Malan, Katherine M. and Andries P. Engelbrecht (2014). Fitness Landscape Analysis for Metaheuristic Performance Prediction. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Berlin and Heidelberg: Springer: 103–129.

Ochoa, Gabriela, Francisco Chicano, Renato Tinós, and Darrell Whitley (2015). Tunnelling Crossover Networks. In: ***Proceedings of the 2015 Genetic and Evolutionary Computation Conference - GECCO '15.*** Ed. by Sara Silva. Madrid, Spain: ACM Press: 449–456.

Ochoa, Gabriela, Marco Tomassini, Sébastien Verel, and Christian Darabos (2008). A study of NK landscapes' basins and local optima networks. In: ***Proceedings of the 10th annual conference on Genetic and evolutionary computation - GECCO '08.*** Ed. by Maarten Keijzer. Atlanta, GA, USA: ACM Press: 555–562.

Ochoa, Gabriela, Nadarajen Veerapen, Darrell Whitley, and Edmund K. Burke (2016). The Multi-Funnel Structure of TSP Fitness Landscapes: A Visual Exploration. In: ***Artificial Evolution: 12th International Conference, Evolution Artificielle, EA 2015.*** Lyon: Springer International Publishing: 1–13.

Ochoa, Gabriela, Sébastien Verel, Fabio Daolio, and Marco Tomassini (2014). Local Optima Networks: A New Model of Combinatorial Fitness Landscapes. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Vol. 6. Emergence, Complexity and Computation. Berlin and Heidelberg: Springer: 233–262.

Ochoa, Gabriela, Sébastien Verel, and Marco Tomassini (2010). First-improvement vs. best-improvement local optima networks of nk landscapes. In: ***PPSN'10: Proceedings of the 11th International Conference on Parallel Problem Solving from Nature.*** Ed. by Robert Schaefer, Carlos Cotta, Joanna Kolodziej, and Günter Rudolph. Vol. I. Kraków, Poland: Springer: 104–113.

Perron, Oskar (1907). Zur Theorie der Matrices. ***Mathematische Annalen 1,*** 64(1): 248–263.

Pitzer, Erik and Michael Affenzeller (2012). A Comprehensive Survey on Fitness Landscape Analysis. In: ***Recent Advances in Intelligent Engineering Systems.*** Ed. by János Fodor, Ryszard Klempous, and Carmen Paz Suárez Araujo. Vol. 378. Studies in Computational Intelligence. Berlin and Heidelberg: Springer: 161–191.

R Development Core Team (2009). ***R: A Language and Environment for Statistical Computing.***

Reidys, Christian M. and Peter F. Stadler (2002). Combinatorial Landscapes. ***SIAM Review,*** 44(1): 3–54.

Rothlauf, Franz (2011). ***Design of modern heuristics: Principles and application.*** Berlin and Heidelberg: Springer.

Stillinger, Frank H. (1995). A Topographic View of Supercooled Liquids and Glass Formation. ***Science,*** 267(5206): 1935–1939.

Verel, Sébastien, Fabio Daolio, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks with escape edges. In: ***Artificial Evolution.*** Angers, France: Springer: 49–60.

Weinberger, Edward D. (1990). Correlated and uncorrelated fitness landscapes and how to tell the difference. ***Biological cybernetics,*** 336: 325–336.

Wright, Sewall (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: ***Proceedings of the 6th International Congress of Genetics.*** Ed. by Donald F. Jones. Ithaca, New York: Morgan Kaufmann Publishers Inc.: 356–366.

Chapter 4

Communities of Local Optima as Funnels in Fitness Landscapes

Sebastian Herrmann, Gabriela Ochoa, Franz Rothlauf

Abstract

We conduct an analysis of local optima networks extracted from fitness landscapes of the Kauffman NK model under iterated local search. Applying the Markov cluster algorithm for community detection to the local optima networks, we find that the landscapes consist of multiple clusters. This result complements recent findings in the literature that landscapes often decompose into multiple funnels, which increases their difficulty for iterated local search. Our results suggest that the number of clusters as well as the size of the cluster in which the global optimum is located are correlated to the search difficulty of landscapes. We conclude that clusters found by community detection in local optima networks offer a new way to characterize the multi-funnel structure of fitness landscapes.

4.1 Introduction

The analysis of fitness landscapes reveals that local optima are often not randomly distributed in the search space, but instead they are clustered in a "central massif" or "big valley". This so-called big valley hypothesis holds for a variety of optimization problems including the traveling salesman problem (TSP; Boese et al., 1994). Recent studies (Hains et al., 2011; Ochoa, Veerapen, et al., 2016) extend the big valley hypothesis as they find that there is a structure of multiple funnels (instead of one single cluster) in the fitness landscape, which leads to a higher search difficulty for algorithms based on the principle of iterated local search (ILS), a search strategy that combines local search with perturbation steps (Lourenço et al., 2003). Such global multi-funnel structures have been observed before in continuous optimization (Kerschke et al., 2015; Locatelli, 2005; Lunacek and Whitley, 2006), however a more detailed characterization of funnels in combinatorial spaces is still lacking.

Ochoa, Veerapen, et al. (2016) propose to characterize funnels in combinatorial spaces using local optima networks (LONs; Ochoa, Tomassini, et al., 2008). The idea of LONs was inspired by the study of energy landscapes (Stillinger, 1995), and it was found that energy landscapes often have a structure of multiple funnels as well (Massen and Doye, 2005). A local optima network is a network representation of a fitness landscape. Ochoa, Veerapen, et al. (2016) examined the LONs extracted from several TSP instances[1] under the *Chained Lin-Kernighan heuristic* (Applegate et al., 2003; Lin and Kernighan, 1973), which is an implementation of the ILS approach. The extracted LONs consisted of multiple components, and Ochoa et al. conjectured that the absence of ties between these components could be an explanation of why ILS often fails to find the global optimum: since there is no path frequently connecting the components, the algorithm may get trapped in one of them. Consequently, components in LONs could offer a way to characterize funnels. This study considers large search spaces of the TSP instances, in consequence the LONs were collected by a sampling procedure. Thus, the presence of multiple components could be a consequence of the sampling. Furthermore, the notion of components conflicts with the formal definition of a fitness landscape, which usually consists of a single component.

We argue that LONs are a promising approach for a deeper study on the extended big valley hypothesis, and it would be worthwhile to examine if the existence of funnels can be shown by a state-of-the-art method from the portfolio of complex network analysis, i.e., "community detection" (Fortunato, 2010). Furthermore, it would be interesting to examine systematically how the presence of funnels is related to search difficulty.

[1] http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/

This paper explores whether a multi-funnel structure exists for landscapes from the Kauffman NK model (Kauffman and Levin, 1987) under iterated local search. The NK model is a class of pseudo-Boolean functions that have been used frequently in studies on fitness landscapes and search heuristics performance. When studying landscapes from the NK model, we are able to generate a large number of instances, and to limit the size of the search space. As a result, the computational effort for extracting local optima networks can be adjusted. Our study applies a "community detection" approach called the Markov cluster algorithm (van Dongen, 2001) to identify the funnels in LONs or landscapes, resp. Community detection has so far only been applied once on LONs (Daolio, Tomassini, et al., 2011), but the implications for heuristic search are yet unclear.

The article is structured as follows: Section 4.2 describes the concept of fitness landscapes and local optima networks with escape edges. Section 4.3 summarizes the principle of iterated local search. Section 4.4 describes our experimental setup. Our results are presented and discussed in Section 4.5. A brief summary and our conclusions are in Section 4.6.

4.2 Fitness Landscapes & Local Optima Networks

Fitness landscapes are a concept that originated from theoretical biology (Wright, 1932). In combinatorial optimization, the concept of fitness landscapes can be used to study the structure of problems as well as the dynamics of heuristic search. A fitness landscape is defined as a triplet of the search space S, the fitness function f, and the neighborhood structure $N(S)$. The search space S contains all valid solution candidates. The fitness function $f : S \to \mathbb{R}_{\geq 0}$ assigns a fitness value[2] to each $s \in S$. The neighborhood function $N : S \to \mathcal{P}(S)$ assigns a set of neighbors $N(s)$ to every $s \in S$. Usually, the neighbors are the solutions that can be reached by a local search step.

Local search is a concept that iteratively tries to improve a solution by applying small changes (in terms of the distance function). A simple implementation of local search is the best improvement hill climber (Algorithm 4.1). The algorithm usually starts with a random solution. It scans the neighborhood of the current solution and selects the best neighbor with a superior fitness as the next solution. This procedure is repeated until no better neighbor is found. Then, the algorithm has reached a local optimum and terminates.

A *local optimum* is a solution that has a higher fitness than its neighbors. Local optima cannot be overcome by a search method moving from a solution to one of its neighbors and accepting only better solutions (Glover, 1986). A higher number of

[2]We assume that the fitness function returns non-negative values.

local optima leads to a landscape that is more "rugged", which generally indicates a higher search difficulty for local search (Weinberger, 1990).

A local optimum is surrounded by a *basin of attraction*, i.e., the set of solution candidates from which the hill climbing algorithm converges to this local optimum. The basin around a local optimum lo is defined as a function

$$B : lo \rightarrow \mathcal{P}(S \backslash LO) \tag{4.1}$$

which assigns an element from the set of all subsets (power set $\mathcal{P}$) over the solutions in the search space to each local optimum $lo \in LO$ (the set of all local optima).

Algorithm 4.1: Best Improvement Hill Climbing (hillClimb)

Require: Solution space S,
Fitness function $f(S)$,
Neighborhood function $N(S)$,
Initial solution s_0

1: i $\leftarrow 0$
2: **repeat**
3: choose $x \in N(s_i)$ s.t. $f(x) = max_{x \in N(s_i)}(f(x))$
4: **if** $f(x) > f(s_i)$ **then**
5: $s_{i+1} \leftarrow x$
6: **else**
7: $s_{i+1} \leftarrow s_{i-1}$
8: **end if**
9: $i \leftarrow i + 1$
10: **until** s_i is local optimum: $\{x \in N(s_i) \mid f(x) \geq f(s_i)\} = \emptyset$
11: **return** s_i

A *local optima network* (LON; Ochoa, Tomassini, et al., 2008) is a representation of a fitness landscape that allows the application of the complex-network analysis framework. Complex networks have been used to study the structure and dynamics of systems that consist of numerous entities which are in some way connected (Albert and Barabási, 2002; Boccaletti et al., 2006). Studies on the dynamics in networks include the influence of nodes (centrality) as well as information flow and diffusion (Borgatti, 2005; Valente, 1996). LONs are a novel way to examine the trajectory of algorithms in fitness landscapes.

A network is a graph $G = (V, E)$ with the set of vertices V and the set of edges E. For a LON, the vertex set V represents all local optima of the fitness landscape. An edge exists between two nodes (local optima), if there is a potential transition between the two local optima. The edges are directed and weighted. The edge

weights $w_{x,y}$ represent the probability that a search algorithm moves from local optimum lo_x to a solution in the basin around lo_y, assuming that the current state is lo_x. Verel et al. (2012) introduced the concept of escape edges, which are defined according to the distance function of the fitness landscape d (minimal number of moves between two solutions). An escape edge is defined as follows: there exists a directed edge e_{xy} (escape edge) from local optimum lo_x to lo_y if there is a solution s such that

$$d(s, lo_x) \leq D \wedge s \in B(lo_y). \tag{4.2}$$

The weight w_{xy} of edge e_{xy} is the probability that a search algorithm can escape from the local optimum lo_x into the basin around lo_y. The constant $D > 0$ determines the maximum distance that is allowed for the escape. A LON with escape edges is a model describing the stochastic process of iterated local search (ILS) in a fitness landscape (cf. Chapter 3).

4.3 Iterated Local Search

Iterated local search (ILS) combines the concept of intensification by local search with diversification by a number of perturbation steps. During intensification, heuristics focus on promising areas of the search space, whereas during diversification, new areas are explored (Rothlauf, 2011). Algorithm 4.2 describes the principle of ILS. Usually, ILS starts with a randomly selected solution s_0 from the search space S. Then, the algorithm performs a hill climbing procedure (algorithm 4.1).

Hill climbing stops when it reaches a local optimum s^*, i.e no further improvement is possible. Then, ILS performs a diversification step by applying a limited perturbation to the local optimum, resulting in s'. As a next step, hill climbing is again applied starting with s', until the next local optimum $s^{*'}$ is reached. If the new local optimum $s^{*'}$ is different from the previous s^* and has higher fitness, the algorithm has "escaped" to a new local optimum, and the change is accepted. Otherwise, another perturbation is applied to s^*. This procedure is repeated until a termination condition is met, e.g. a fixed number of perturbation steps without any further improvement.

Algorithm 4.2: Iterated Local Search (ILS)

Require: Solution space S,
Fitness function $f(S)$,
Neighborhood function $N(S)$,
Stopping threshold t

Choose initial random solution $s_0 \in S$
$s^* \leftarrow \text{hillClimb}(s_0)$
$i \leftarrow 0$
repeat
  $s' \leftarrow \text{perturbation}(s^*)$
  $s^{*'} \leftarrow \text{hillClimb}(s')$
  if $f(s^{*'}) > f(s^*)$ **then**
    $s^* \leftarrow s^{*'}$
    $i \leftarrow 0$
  end if
  $i \leftarrow i + 1$
until i $\geq$ t
return s^*

4.4 Experimental Setup

For our experiments, we calculated the local optima networks for 300 instances of the Kauffman NK model (Kauffman and Levin, 1987). The NK model is a combinatorial optimization problem from the class of pseudo-Boolean functions. An instance is defined by the two parameters N and K, where N is the number of binary variables and K is the number of variables interacting with each other. The size of the search space S is $|S| = 2^N$. The fitness function

$$f_{NK} : [0,1]^N \rightarrow [0,1] \tag{4.3}$$

assigns a score to every combination of bits. The fitness $f_{NK}(s)$ of a solution s is the average of the values of N sub-functions (one for each bit). Each sub-function f_i assigns a fitness contribution for each bit i, depending on the value of bit i and K other bits that were randomly selected before instantiation:

$$f_i : [0,1]^{K+1} \rightarrow [0,1]. \tag{4.4}$$

The parameter K determines the number of co-variables per decision variable (epistasis). All values of the fitness functionf_{NK} are normalized to values between 0 and 1, with $f_{NK}(s_{opt}) = 1$ as the fitness of the global optimum s_{opt}. In general, a higher

value leads to a higher search difficulty (Weinberger, 1990). The distance between two solutions $x, y \in S$ is calculated by the Hamming distance $d(x, y) = \sum_{i=0}^{n} |x_i - y_i|$, i.e., the number of bits that are set to different values when comparing two solutions.

We randomly generated 300 NK fitness landscapes with $N = 20$ decision variables and different values of $K \in \{5, 10, 15\}$. Thus, we have 100 problems instances each for three levels of epistasis K. The size N of our problem instances is relatively low, since the computational effort for the experiments grows factorially by the problem size N (especially calculating the LON is time-consuming). For each instance, we extracted the local optima network and applied the **M**arkov **cl**uster algorithm (MCL) to the networks in order to detect funnels in the landscapes.

Furthermore, we applied ILS to each problem instance and calculated the percentage of runs that are able to find the optimal solution. Results are averaged for 100 independent ILS runs for each problem instance. For the hill climbing steps procedure in ILS, we assumed that two solutions x, y are neighbors if their Hamming distance is equal to one ($d_{\max} = 1$). Thus, a local search step flips exactly one bit of the current solution. For the perturbation operator in ILS, we flip two random bits in one step. When extracting the LONs, we set the parameter D for the maximum escape distance to $D = 2$.

We calculated the following measures for each fitness landscape or LON, resp.:

- $\#lo$: the number of local optima, i.e., the number of nodes in the LON,
- $\#c$: the number of clusters that are found by the MCL algorithm,
- $\#br$: the number of bridges (nodes that are in-between two clusters) that are found by the MCL algorithm,
- fo: the size (number of nodes) of the global optimum's cluster over the total number of local optima,
- mod and mod_w: two measurements of modularity to assess the quality of the clustering as proposed by the MCL algorithm (the concept of modularity will be introduced in Section 4.5.3),
- $SRate_{ILS}$: the success rate of ILS to find the global optimum (averaged over 100 independent runs).

We used Java to generate the NK landscapes, to extract the LONs, and to apply ILS. We ran the experiments on a cluster using 64 cores with 256 GB of RAM per node. The running time per fitness landscape was approx. 20 minutes in the case of low epistasis, and 2 hours in the case of high epistasis. We implemented the Markov cluster algorithm using *Numerical Python* (Oliphant, 2007). Statistical analysis was done with *R* (R Development Core Team, 2009), visualizations of the graphs with *Gephi* (Bastian et al., 2009).

4.5 Community Detection Analysis

4.5.1 Markov Cluster Algorithm

Community detection is method of graph partitioning. The objective of a graph partitioning problem is to search for a partition of a graph's nodes which optimizes a given cost function. A typical cost function is the number of links that connect between the partitions. Mostly, there are also several constraints, e.g. limits for the allowed number of partitions, nodes per group, etc. Many graph partitioning problems are NP-hard (Talbi and Bessière, 1991).

Community detection is a rather exploratory method in the sense that there are no pre-formulated constraints to the problem of choosing a partition of a graph (or network, respectively). Instead, a community detection algorithm is free in determining the number of communities or the number of nodes per community. A very general definition of a community is a group of nodes that have more links among each other than to nodes in other communities. However, the definition of a community depends on the discipline applied and there exists a variety of algorithms that have been validated for different purposes (Fortunato, 2010; Porter et al., 2009).

To select an algorithm for community detection in LONs, we took into consideration that a LON represents the stochastic process of an algorithm in the fitness landscape. An algorithm for detecting communities in graphs of stochastic flows is the **M**arkov **cl**uster algorithm (MCL; van Dongen, 2001). The MCL algorithm has been successfully used in various domains, e.g. protein folding networks (Satuluri et al., 2010), and also in a study on local optima networks of the quadratic assignment problem (Daolio, Tomassini, et al., 2011) to identify clusters of local optima.

A description of MCL is given in Algorithm 4.3. Let G be the graph of a LON. E is the adjacency matrix containing the weights of the directed edges in G. Since the edge weights are the non-negative probabilities to move from a given local optimum into the basin around another local optimum, the probabilities of all columns in E sum up to 1. Thus, E is a stochastic matrix, which can be interpreted as a transition matrix in a discrete-time Markov chain.

To identify communities in G, MCL applies two mechanisms: expansion and inflation. The expansion operator raises the adjacency matrix E to the non-negative power p. Expansion ensures that different regions of the graph stay connected. The second mechanism is the inflation operator. Inflation raises each column E_i from the adjacency matrix E to a non-negative power r, and then re-normalizes the column. The re-normalization ensures that each column again sums up to 1, which is a constraint for a stochastic matrix of a Markovian process. As a result, inflation increases the weights of heavy-weighted edges, whereas the weights of low-weighted edges are reduced. Both mechanisms are repeated until the algorithm converges,

Algorithm 4.3: Markov Cluster Algorithm (MCL)

Require: Graph G,
Power Parameter p,
Inflation Parameter r

$E = AdjacencyMatrix(G)$
repeat
  {Expand Matrix}
  $E = E^p$
  for all $A_i \in E$ **do**
    {Inflate Columns}
    $E_i = A_i^r$
    $E_i = normalizeVector(E_i)$
  end for
until E has converged
return E

i.e., the transition matrix E reaches a steady state. For our NK landscapes with $N = 20$, this state is usually reached after ≈ 30 iterations.

Applying the Markov cluster algorithm on a *LON* of an NK landscape results in a new graph (see Figure 4.2 for an example), where the different clusters can be identified. As a result of the clustering algorithm, we obtain several unconnected subgraphs, where each subgraph is a star: in each star, there is an *attractor* in its center and a periphery around it. Each star (sub-graph) represents a community (or cluster) of the original graph (Figure 4.1), and thus each node in the original graph belongs to one of these partitions, as indicated by the colors. Sometimes, we find nodes that belong to two clusters (for example, the purple node in Figure 4.2). In our analysis, we call such nodes *bridges*.

4.5.2 Community Structure of the LONs

Figure 4.1 plots an example of a LON of an NK landscape with low epistasis ($K = 5$). An edge between two nodes indicates the existence of an escape edge. All nodes that are assigned to the same cluster by the MCL algorithm have the same color. The size of a node indicates the fitness (larger size means better fitness). We only plot the best 10% of the nodes. In the plot, the nodes have been positioned by a force-directed layout ("ForceAtlas2"). This algorithm arranges the nodes in an aesthetically pleasing way by simulating that the edges between nodes are springs, and then tries to minimize the tension of the springs as well as the number of intersections.

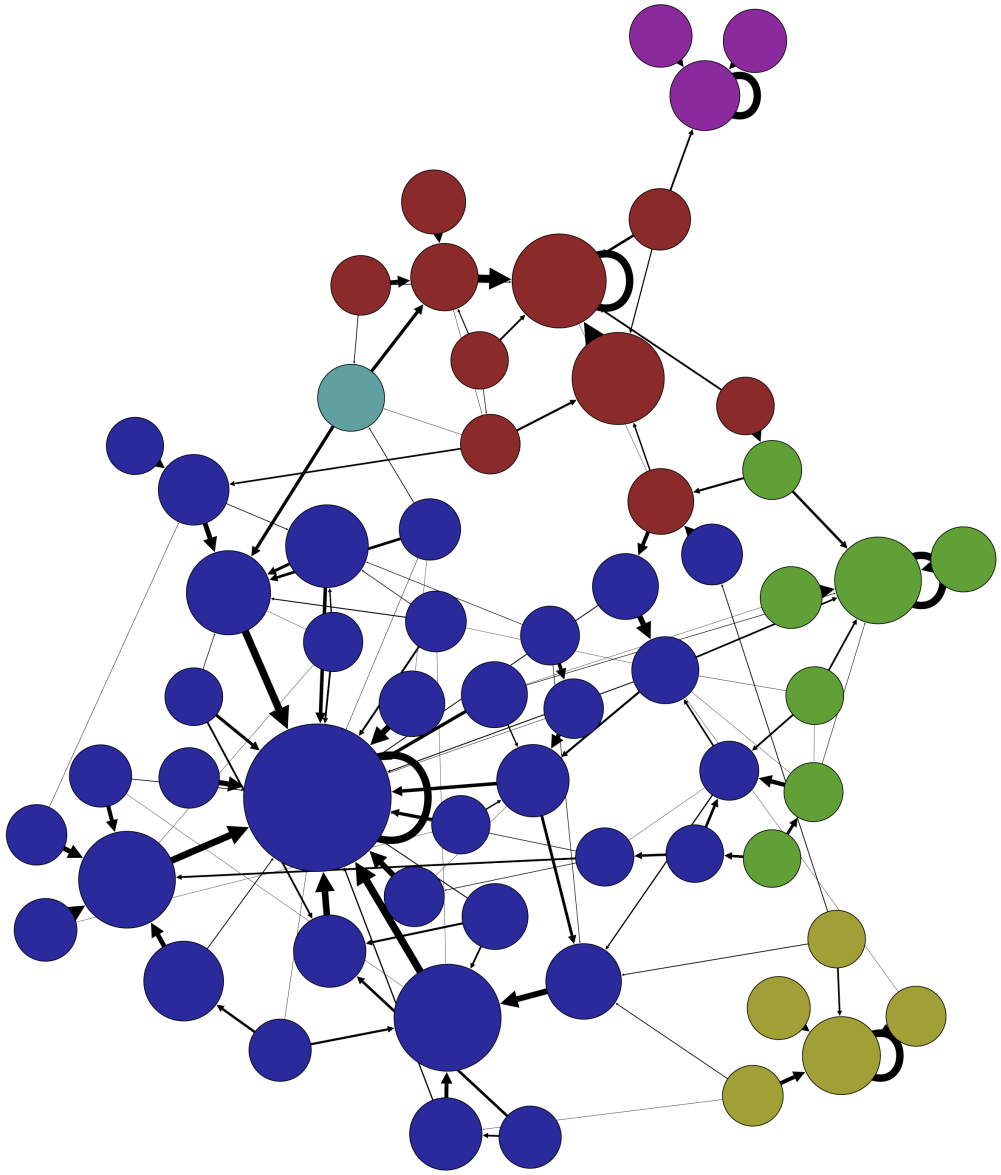

Figure 4.1: An example of a local optima network of an NK landscape with $N = 20$, $K = 5$. Each node is a local optimum. The color of the nodes represents the cluster assignment as obtained from the MCL algorithm. The size of a node indicates its fitness. The edge thickness indicates the edge weight, which is the probability to move from the outbound local optimum into the basin around the inbound local optimum. To highlight the structure of the network, we plot only the best 10% of nodes (by fitness).

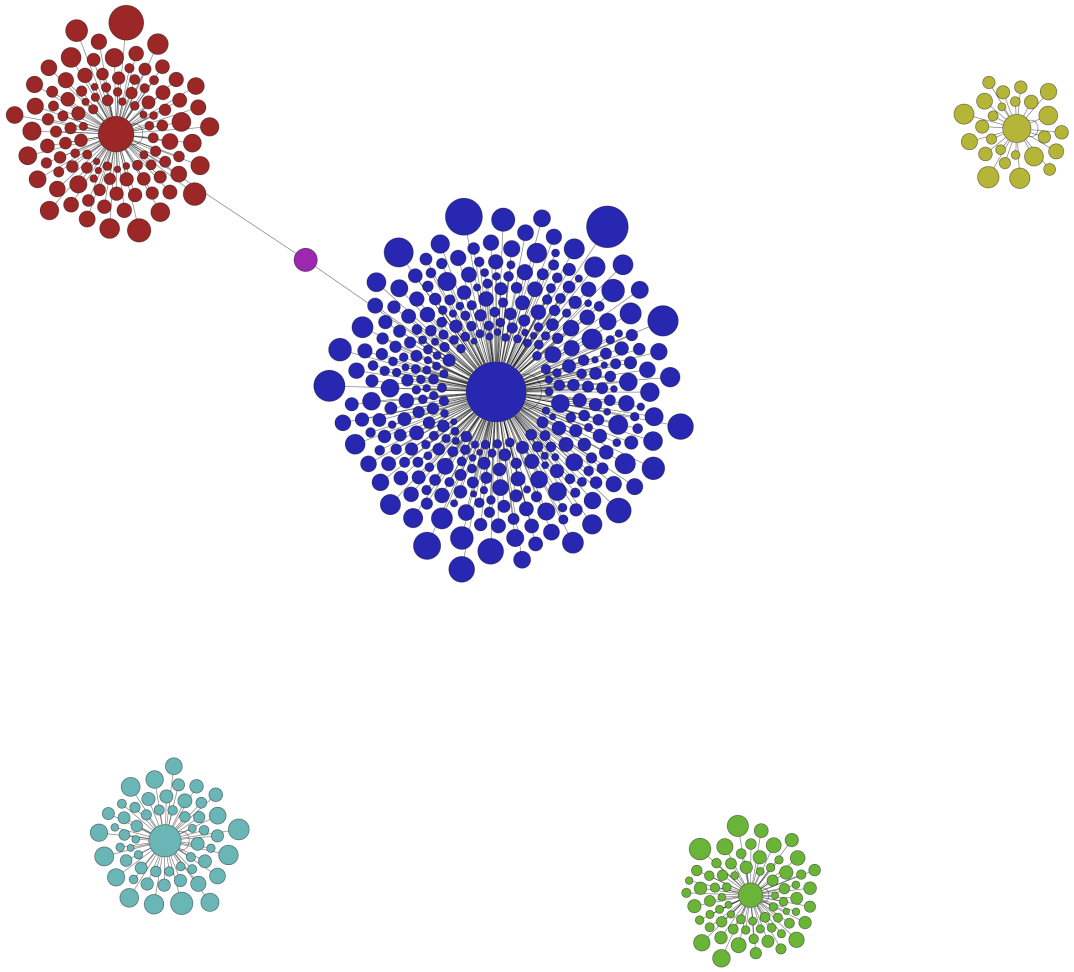

Figure 4.2: The output of the Markov cluster algorithm applied to the LON from Figure 4.1. As before, each node is a local optimum and the node size indicates fitness. Each subgraph represents a community, and nodes in the centers are their attractors. There is a bridge node connecting the blue and the red cluster.

K	$\overline{\#lo}$	$\overline{\#c}$	$\overline{\#br}$	$\overline{fo}$	$\overline{SRate_{ILS}}$
5	546.9	5.285	0.25	0.412	0.407
10	3384.5	19.985	2.303	0.147	0.142
15	7957	42.101	7.621	0.074	0.069

Table 4.1: Characteristics of networks obtained by the Markov cluster algorithm: average number $\overline{\#lo}$ of local optima, clusters $\overline{\#c}$, and bridges $\overline{\#br}$, average size $\overline{fo}$ of global optimum's cluster over the total number of local optima, and average success rate $\overline{SRate_{ILS}}$ of ILS.

Applying the MCL algorithm to this LON yields the network plotted in Figure 4.2: MCL iteratively increases the heavy weights and decreases the low weights until a new graph, consisting of several unconnected stars emerges. In the figure, we plot all nodes (and not only the best 10%).

The nodes in both figures have been colored according to the clusters that result from applying the MCL algorithm on this instance. In Figure 4.1, we see that nodes which are identified to be in the same cluster are not randomly distributed across the graph, but clustered. The nodes that are closely positioned to each other have the same color in most cases and thus belong to the same community. Even though we should be careful about intuitive conclusions from a visual inspection, the detection of the communities as proposed by the MCL algorithm matches the visual topological structure of the LON.

As a next step, we study the characteristics of the networks obtained by the Markov cluster algorithm. In particular, we examine how the characteristics of LONs depend on the epistasis K. Table 4.1 lists the results. As expected, the average number of local optima increases with K. This explains the lower value of modularity with growing K: a higher number of nodes makes it more difficult to find a good partition of a graph. The total number of clusters also grows with K, and so does the average number of bridges. The fraction of the global optimum's cluster becomes lower for higher K. This effect also holds for the empirical success rate of ILS, which is a measurement for the search difficulty of the landscapes.

An additional finding we made is on the characteristics of the bridge nodes. Studying our data, we found that the bridges are among those nodes that have the highest closeness centrality in the LONs. Closeness centrality takes into account the geodesic (shortest-path) distances between all the nodes. The shorter a node's paths to all other nodes are, the higher is its closeness centrality (Freeman, 1979). An interpretation of this is that these nodes are by average "close" to the other nodes; they

K	$\overline{Q}$	$\overline{Q_w}$
5	0.3721	0.4789
10	0.2629	0.3966
15	0.1947	0.3019

Table 4.2: The average modularity (with and without considering edge weights) for different values of K as achieved by applying the Markov cluster algorithm for community detection to our LONs.

connect different areas of the fitness landscape. It would be worthwhile to study in future research whether typical characteristics of these nodes can be identified. This possibility might offer new opportunities for better search operators.

4.5.3 Quality of Community Structure

We want to quantify the quality of the community structure, i.e., the partition of the network. A common approach to quantify the strength of a community structure is the modularity Q as proposed by Newman and Girvan (Newman and Girvan, 2004). Given a certain partition of a graph, the modularity of this partition is "the number of edges falling within groups minus the expected number in an equivalent network with edges placed at random" (Newman, 2006). Thus, $Q = 0$ indicates a partition of the network by which the number of within-community edges is not better than random, whereas $Q = 1$ is a perfect partition of the network. In practice, values between $[0.3, 0.7]$ indicate a strong community structure.

The modularity of a graph can be calculated by either ignoring or considering the edge weights w_{ij}. We calculated both variants of modularity, where Q ignores the edge weights and Q_w considers the edge weights. Table 4.2 presents the results for the community structures revealed by the MCL algorithm in our LONs. When ignoring weights, we observe a value of Q between 0.19 and 0.37. Taking the edge weights into account, we observe higher values for Q_w between 0.48 and 0.30. In general, the modularity decreases with increasing epistatis K (see below). A possible explanation for the difference between Q and Q_w is that in LONs, the edge weights represent the transition probabilities between the local optima. Due to a non-linear distribution of these weights, the calculation of Q is biased towards lower values. In general, considering edge weights for the modularity returns more accurate results; values of $Q_w > 0.3$ indicate that the partition of the networks as proposed by the MCL algorithm is satisfactory and we obtain meaningful clusters.

We conclude from the high quality of the community structure that the presence of several clusters in the LONs can be confirmed. Assuming that these cluster (communities) are an alternative way of characterizing "funnels", this finding underpins the hypothesis that there is a structure of multiple funnels in fitness landscapes of the NK model under ILS.

4.5.4 Community Structure and Search Difficulty

Finally, we examine the relationship between the community structure and search difficulty. We calculate the squared Pearson correlations (R^2 in a univariate, linear regression model) between the number of clusters as well as the size of the cluster containing the global optimum and the success rate of ILS. For the number of clusters, we observe a medium correlation ($R^2 \approx 0.46$, Figure 4.3). Thus, the presence of multiple clusters has an effect on search difficulty, even though it explains only a limited fraction of the variance in the success rate of ILS.

Figure 4.4 plots the R^2 between the relative size fo of the cluster containing the global optimum and ILS success rate. We find a strong correlation between fo and search difficulty ($R^2 \approx 0.94$). Thus, the size of the cluster containing the global optimum is a strong predictor for the performance of ILS. A possible explanation for this correlation is that ILS can more easily find the global optimum if it is surrounded by many other local optima. In contrast, if the global optimum's cluster is very small, it is very likely that ILS gets stuck in another cluster and a high number of perturbations are necessary to reach the cluster of the global optimum.

Interestingly, we found that the centers of the stars are the final local optima that are returned by ILS. Thus, the global optimum is always a center of a star and among the group of attractors. We already found that the PageRank centrality of a LON reflects the stationary distribution of a local search algorithm (cf. Chapter 2). Our results indicate that the relative cluster size of the attractors constructed by the MCL algorithm is nearly identical to the PageRank value of the absorbing-state local optima in the landscape. This observation could offer an alternative way for performance prediction, however, we leave this for further research.

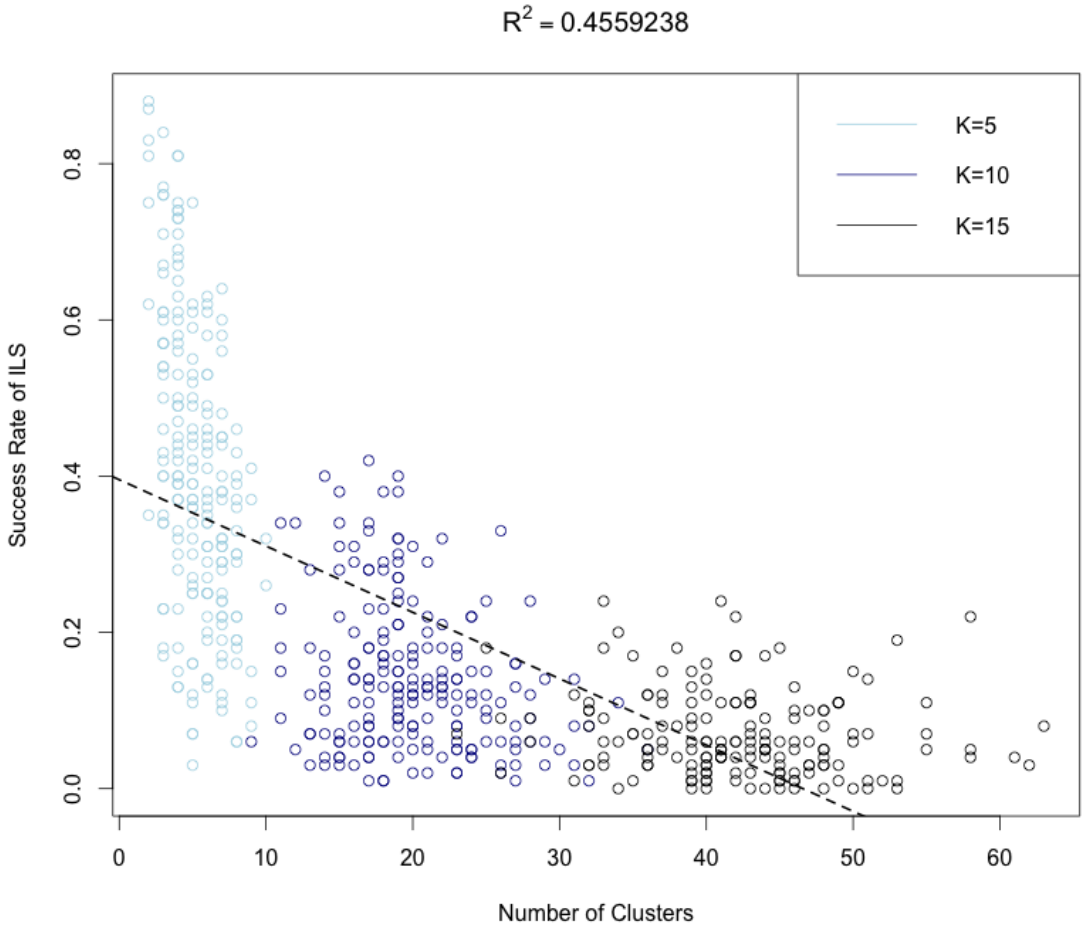

Figure 4.3: The success rate of ILS over the number of clusters. Each dot represents a particular NK landscape. The dashed line is a univariate linear regression model, which is identical to the squared Pearson correlation.

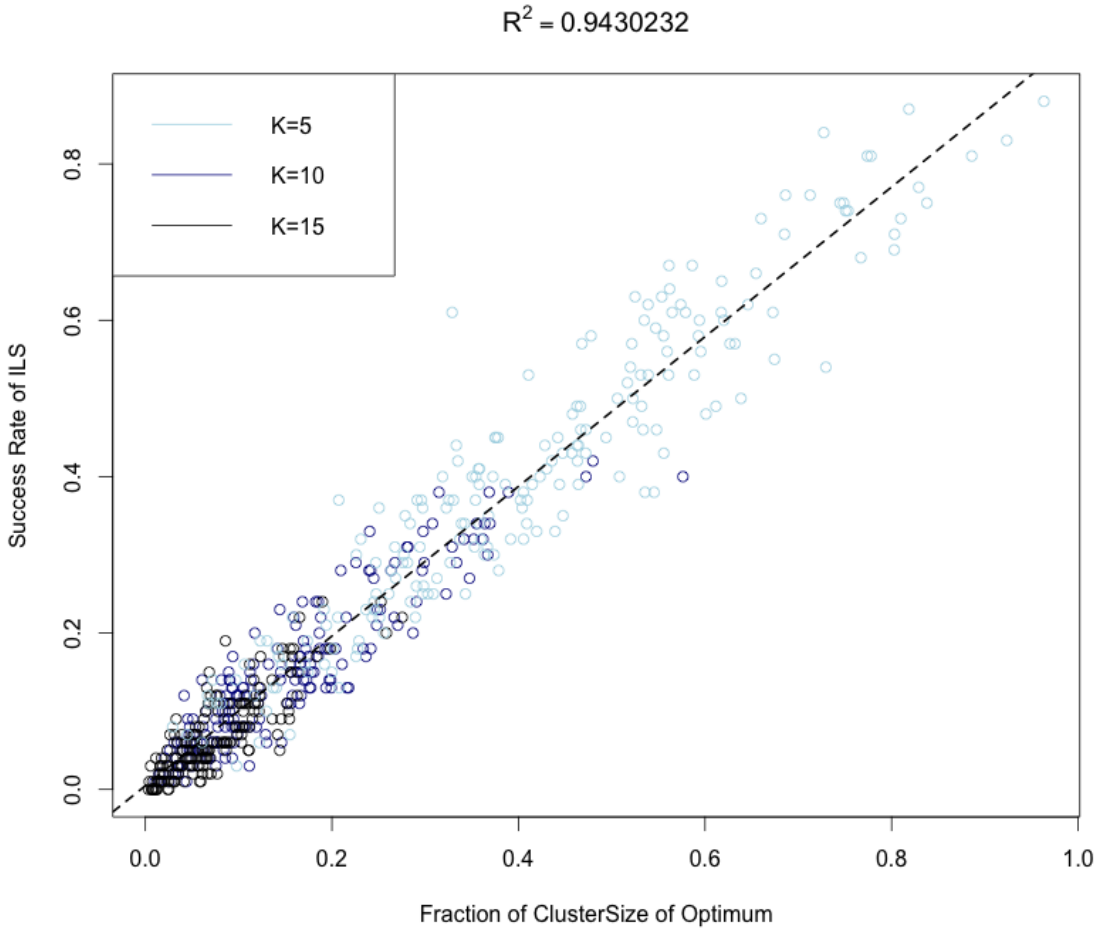

Figure 4.4: The success rate of ILS over the relative size of the cluster which contains the global optimum. Each dot represents a particular NK landscape. The dashed line is a univariate linear regression model, which is identical to the squared Pearson correlation.

4.6 Conclusion

We conducted an experimental study on the characterization of a multi-funnel structure in fitness landscapes emerging from the Kauffman NK model under iterated local search. To analyze the presence of funnels in the landscapes, we used local optima networks with escape edges and applied an algorithm for community detection, i.e., the Markov cluster algorithm.

The results confirm that the landscapes consist of several clusters (communities), and the number of clusters grows with the number of interdependencies between the decision variables (epistasis). A higher number of clusters leads to a higher search difficulty, measured by the empirical success rate of ILS. An explanation for this observation is that ILS gets stuck due to the presence of funnels, which cannot be overcome by the perturbation operator applied. We estimate that a stronger perturbation operator could be used in such a case to overcome this situation.

Furthermore, the size of the cluster which contains the global optimum is strongly correlated to the success rate of ILS. The probability to find the global optimum's cluster decreases with lower size of this cluster. On the other hand, the global optimum can be found with higher probability if it is surrounded by many other local optima. This is no surprise when considering that—given a fixed number of local optima—a higher number of clusters should in general lead to smaller clusters (by the number of nodes) and thus to higher search difficulty. Another explanation is that smaller clusters are probably harder to locate, since they require ILS to perform more perturbation steps.

The size of a cluster returned by the Markov cluster algorithm also offers a new possibility to predict the performance of ILS. We conjecture that the effect of applying the MCL algorithm is closely related to the PageRank centrality, which we leave for further research. Another finding that could be interesting for further research is that there are nodes connecting different areas in the fitness landscape. It would be interesting to see if this could be exploited by a search operator to help escape from one cluster to another.

In summary, our analysis has shown by community detection in LONs that the underlying landscapes are clustered, and that the presence and shape of these clusters are related to search difficulty. Furthermore, the results obtained by community detection have sufficient quality in terms of the modularity measure. We conclude that communities in LONs are a novel way to characterize the notion of funnels in fitness landscapes.

References

Albert, Réka and Albert-László Barabási (2002). Statistical mechanics of complex networks. ***Reviews of Modern Physics,*** 74(1): 47–97.

Applegate, David, William Cook, and Andre Rohe (2003). Chained Lin-Kernighan for Large Traveling Salesman Problems. ***INFORMS Journal on Computing,*** 15(1): 82–92.

Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. In: ***Third International AAAI Conference on Weblogs and Social Media.***: 361–362.

Boccaletti, Stefano, Vito Latora, Yamir Moreno, Martín Chávez Hoffmeister, and Dong-Uk Hwang (2006). Complex networks: Structure and dynamics. ***Physics Reports,*** 424(4-5): 175–308.

Boese, Kenneth D., Andrew B. Kahng, and Sudhakar Muddu (1994). A new adaptive multi-start technique for combinatorial global optimizations. ***Operations Research Letters,*** 16(2): 101–113.

Borgatti, Stephen P. (2005). Centrality and network flow. ***Social Networks,*** 27(1): 55–71.

Daolio, Fabio, Marco Tomassini, Sébastien Verel, and Gabriela Ochoa (2011). Communities of minima in local optima networks of combinatorial spaces. ***Physica A: Statistical Mechanics and its Applications,*** 390(9): 1684–1694.

Fortunato, Santo (2010). Community detection in graphs. ***Physics Reports,*** 486(3-5): 75–174.

Freeman, Linton C. (1979). Centrality in social networks conceptual clarification. ***Social Networks,*** 1(3): 215–239.

Glover, Fred (1986). Future paths for integer programming and links to artificial intelligence. ***Computers & Operations Research,*** 13(5): 533–549.

Hains, Doug R., Darrel L. Whitley, and Adele E. Howe (2011). Revisiting the big valley search space structure in the TSP. ***Journal of the Operational Research Society,*** 62(2): 305–312.

Kauffman, Stuart A. and Simon Levin (1987). Towards a General Theory of Adaptive Walks on Rugged Landscapes. ***Journal of Theoretical Biology,*** 128(1): 11–45.

Kerschke, Pascal, Mike Preuss, Simon Wessing, and Heike Trautmann (2015). Detecting Funnel Structures by Means of Exploratory Landscape Analysis. In: ***Proceedings of the 2015 Genetic and Evolutionary Computation Conference - GECCO '15.*** Ed. by Sara Silva. Madrid, Spain: ACM Press: 265–272.

Lin, S. and B. W. Kernighan (1973). An Effective Heuristic Algorithm for the Traveling-Salesman Problem. ***Operations Research,*** 21(2): 498–516.

Locatelli, M. (2005). On the Multilevel Structure of Global optimization problems. ***Computational Optimization and Applications,*** 30(1): 5–22.

Lourenço, Helena R., Olivier C. Martin, and Thomas Stützle (2003). Iterated Local Search. In: ***Handbook of Metaheuristics.*** Boston: Kluwer Academic Publishers: 320–353.

Lunacek, Monte and Darrell Whitley (2006). The dispersion metric and the CMA evolution strategy. In: ***Proceedings of the 8th annual conference on Genetic and evolutionary computation - GECCO '06.*** New York, New York, USA: ACM Press: 477.

Massen, Claire P. and Jonathan P. K. Doye (2005). Identifying communities within energy landscapes. ***Physical Review E - Statistical, Nonlinear, and Soft Matter Physics,*** 71(4): 1–13.

Newman, Mark E. J. (2006). Modularity and community structure in networks. ***Proceedings of the National Academy of Sciences,*** 103(23): 8577–8582.

Newman, Mark E. J. and Michelle Girvan (2004). Finding and evaluating community structure in networks. ***Physical Review E - Statistical, Nonlinear, and Soft Matter Physics,*** 69(2): 026113.

Ochoa, Gabriela, Marco Tomassini, Sébastien Verel, and Christian Darabos (2008). A study of NK landscapes' basins and local optima networks. In: ***Proceedings of the 10th annual conference on Genetic and evolutionary computation - GECCO '08.*** Ed. by Maarten Keijzer. Atlanta, GA, USA: ACM Press: 555–562.

Ochoa, Gabriela, Nadarajen Veerapen, Darrell Whitley, and Edmund K. Burke (2016). The Multi-Funnel Structure of TSP Fitness Landscapes: A Visual Exploration. In: ***Artificial Evolution: 12th International Conference, Evolution Artificielle, EA 2015.*** Lyon: Springer International Publishing: 1–13.

Oliphant, Travis E. (2007). Python for scientific computing. ***Computing in Science and Engineering,*** 9(3): 10–20.

Porter, Mason a., Jukka-Pekka Onnela, and Peter J. Mucha (2009). Communities in Networks. ***Notices of the AMS,*** 486(3-5): 1082–1097.

R Development Core Team (2009). ***R: A Language and Environment for Statistical Computing.***

Rothlauf, Franz (2011). ***Design of modern heuristics: Principles and application.*** Berlin and Heidelberg: Springer.

Satuluri, Venu, Srinivasan Parthasarathy, and Duygu Ucar (2010). Markov clustering of protein interaction networks with improved balance and scalability. In: ***Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology - BCB '10.*** New York, USA: ACM Press: 247.

Stillinger, Frank H. (1995). A Topographic View of Supercooled Liquids and Glass Formation. ***Science,*** 267(5206): 1935–1939.

Talbi, E.G. and P. Bessière (1991). A parallel genetic algorithm for the graph partitioning problem. In: ***Proceedings of the 5th international conference on Supercomputing - ICS '91.*** New York, USA: ACM Press: 312–320.

Valente, Thomas W. (1996). Network models of the diffusion of innovations. ***Computational and Mathematical Organization Theory,*** 2(2): 134.

Van Dongen, Stijn (2001). "Graph clustering by flow simulation". PhD thesis. Utrecht University.

Verel, Sébastien, Fabio Daolio, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks with escape edges. In: ***Artificial Evolution.*** Angers, France: Springer: 49–60.

Weinberger, Edward D. (1990). Correlated and uncorrelated fitness landscapes and how to tell the difference. ***Biological cybernetics,*** 336: 325–336.

Wright, Sewall (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: ***Proceedings of the 6th International Congress of Genetics.*** Ed. by Donald F. Jones. Ithaca, New York: Morgan Kaufmann Publishers Inc.: 356–366.

Chapter 5

Coarse-Grained Barrier Trees of Fitness Landscapes

Sebastian Herrmann, Gabriela Ochoa, Franz Rothlauf

Abstract

Recent literature suggests that local optima in fitness landscapes are clustered, which offers an explanation of why perturbation-based metaheuristics often fail to find the global optimum: they become trapped in a sub-optimal cluster. We introduce a method to extract and visualize the global organization of these clusters in form of a barrier tree. Barrier trees have been used to visualize the barriers between local optima basins in fitness landscapes. Our method computes a more coarse-grained tree to reveal the barriers between clusters of local optima. The core element is a new variant of the flooding algorithm, applicable to local optima networks. We use local optima networks as a compressed representation of fitness landscapes. To identify the clusters, we apply a community detection algorithm. A sample of 200 NK fitness landscapes suggests that the depth of their coarse-grained barrier tree is related to their search difficulty for perturbation-based metaheuristics.

5.1 Introduction

To overcome the problem of getting stuck in a local optimum, many metaheuristics based on local search apply a perturbation operator. The perturbation is supposed to "kick" an algorithm away from the current region of the search space. This principle is known as iterated local search (ILS; Lourenço et al., 2003), e.g. as implemented in the Chained Lin-Kernighan heuristic (Applegate et al., 2003; Lin and Kernighan, 1973). The "big valley" hypothesis (Hains et al., 2011) states that the local optima in many fitness landscapes are not randomly distributed, but clustered and surrounding the global optimum. Consequently, one might assume that once a local optimum has been reached, ILS-based algorithms should easily find the global optimum after a limited number of perturbations. However, we know that this is by no means the case in practice. An approach to explain this observation is given in the most recent literature (Hains et al., 2011; Ochoa and Veerapen, 2016b; Ochoa, Veerapen, et al., 2016; cf. also Chapter 4): instead of one big valley, fitness landscapes consist of multiple clusters (or funnels). The existence of such a structure offers a new explanation for the search difficulty of landscapes: since the connections between clusters are sparse, perturbation steps fail to escape from sub-optimal clusters to the cluster of the global optimum.

The objective of this paper is to complement the recent literature on the multi-cluster structure of landscapes with a new approach to study this structure, and to draw conclusions on search difficulty. A method that has been used to characterize the structure of fitness landscapes are barrier trees (Hallam and Prügel-Bennett, 2005). A barrier tree shows in a hierarchical structure how the local optima basins are connected in the landscape. The leaf nodes are the local optima and the branching nodes are the saddle points connecting the basins (van Stein et al., 2013). Due to the ability of ILS to easily move from local optimum to local optimum, we are primarily not interested in the barriers between their basins. The core issue for ILS is that local optima are clustered. Thus, we need to study which barriers exist between these clusters. The method we introduce here addresses this purpose. It allows us to compute a coarse-grained barrier tree and to characterize the landscape on the level of clusters. To reveal the clustering structure of landscapes, local optima networks (LONs; Ochoa, Tomassini, et al., 2008) have been used. A LON is a compressed representation of a fitness landscape. In a LON, each node is a local optimum, and the edges represent the transitions of an algorithm between the basins around the local optima. A problem with LONs is that it can be difficult to visualize their structure when they consist of a large number of nodes and edges. To identify clusters in fitness landscapes, statistical measures have been applied to LONs, e.g. counting the network graph's connected components (Ochoa, Veerapen, et al., 2016) or community detection (cf. Chapter 4).

Our contribution is a modified version of the "flooding algorithm", which accepts as an input (i) a LON of a fitness landscape and (ii) a pre-computed clustering structure of the LON. The output is a coarse-grained picture of the landscape which retains the global structure and allows the eventual visualization of larger landscapes. We demonstrate our method with instances of the Kauffman NK model. For each instance, we computed the LON and the clusters. We obtained the clusters from community detection with the Markov cluster algorithm (van Dongen, 2001), as proposed in an earlier study (Chapter 4). We analyze the resulting barrier trees by visual inspection and a statistical approach. We provide an indication how the depth of the barrier tree is related to the search difficulty of a landscape.

The article is structured as follows: Section 5.2 introduces the concept of fitness landscapes for the study of problems and heuristic search. In Section 5.3, we explain how to construct a standard barrier tree for fitness landscape analysis. In order to construct a coarse-grained barrier tree (based on the local optima clusters), we need a method to identify the clustering structure. In Section 5.4, we introduce local optima networks as a compressed representation of fitness landscapes, and the Markov cluster algorithm to reveal the clustering structure of a fitness landscape. In Section 5.5, we present the algorithm to calculate the coarse-grained barrier tree of a fitness landscape. We visualize two instances and examine their search difficulty. A brief summary and conclusions are in Section 5.6.

5.2 Fitness Landscapes

The concept fitness landscapes was introduced to study the reproductive success of genotypes in theoretical biology (Wright, 1932). Fitness landscapes have been adopted in combinatorial optimization to study the structure of problems and the dynamics of heuristic search. A fitness landscape is defined as a triplet of the search space S, the fitness function f, and the neighborhood structure $N(S)$. The search space S contains all valid solutions. The fitness function $f : S \to \mathbb{R}_{\geq 0}$ assigns a fitness value to each $s \in S$ (we assume non-negative values and a maximization problem). The neighborhood function $N : S \to \mathcal{P}(S)$ assigns a set of neighbors $N(s)$ to every $s \in S$. A solution s_2 is a neighbor of s_1 if it can be reached by applying one step of local search, starting from s_1.

A local optimum is a solution that has a higher fitness than its neighbors (Glover, 1986). A higher number of local optima (modality) leads to a landscape that is more "rugged", which increases the search difficulty for local search-based algorithms (Weinberger, 1990). A local optimum is surrounded by a *basin of attraction*. The basin around a local optimum is the set of solutions from which the optimum attracts

a local search algorithm. We define a function for the basin around a local optimum lo as $B : lo \rightarrow \mathcal{P}(S \backslash LO)$. B assigns an element from the set of all subsets (power set $\mathcal{P}$) over all solutions in the search space to each local optimum $lo \in LO$ (the set of all the local optima in the fitness landscape).

The Kauffman NK model of landscapes (Kauffman and Weinberger, 1989) is frequently used for the study of fitness landscapes. The NK model is a combinatorial optimization problem from the class of pseudo-Boolean functions. An instance is defined by the two parameters N and K, where N is the number of binary variables. The size of the search space S is $|S| = 2^N$. K is the number of variables interacting with each other (epistasis). To instantiate the model, the co-variables are randomly selected. A higher value of K leads to a higher search difficulty (Weinberger, 1990). The distance between two solutions $x, y \in S$ is the number of differing bits (Hamming distance).

5.3 Barrier Trees of Fitness Landscapes

Barrier trees were introduced in computational chemistry to study the structure of potential energy landscapes (Becker and Karplus, 1997; Flamm et al., 2002), i.e., to examine the barriers that exist between the optima basins. Barrier trees are sometimes referred to as *disconnectivity graphs* (Doye et al., 1999a; Doye et al., 1999b). Even though Barrier trees have been used to study heuristic search (Hallam and Prügel-Bennett, 2005; van Stein et al., 2013), the literature on this topic is rather sparse. To construct the barrier tree of a fitness landscape, the set of the local optima (we assume local maxima in this paper), and the solutions connecting at least two basins around different local optima, is required. The connecting solutions are also called *s*addle points. In a two-dimensional landscape, a saddle point is a local minimum. In a higher dimensional landscape, multiple local minima connecting two basins may exist. In such a case, the saddle point is the local minimum with maximal fitness. Since the fitness of the saddle point is lower than the fitness of the two connected local optima, it can be interpreted as a barrier between them: to move from one of the local optima to the other, an algorithm has to accept a fitness deterioration down to the level of the local minimum. To visualize the barrier tree, local optima are identified with leaves, while the branching nodes represent saddle points separating groups of local optima.

A method to compute the barrier tree of a fitness landscape is the so-called "flooding algorithm" (van Stein et al., 2013). We think that a comprehensive understanding of this method is essential; hence we depict the mechanism in Figure 5.1. For a maximization problem, the algorithm iterates over all solutions in the search space

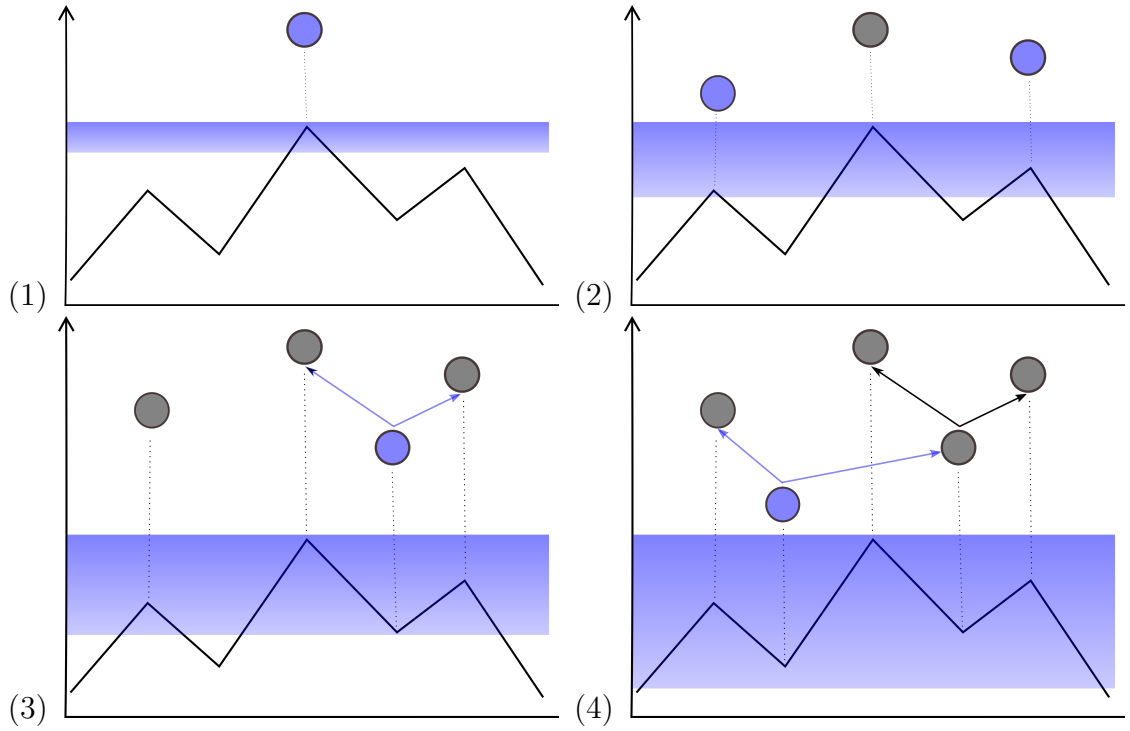

Figure 5.1: Four steps of the flooding algorithm, constructing the barrier tree of a fitness landscape. The vertical axis is the fitness, the horizontal axis is the landscape. Since we use a maximization problem, the space is "flooded" from the top to the bottom.

in a descending order (in terms of fitness): the landscape is "flooded". When a local maximum is found, a node is added to the barrier tree (steps 1 and 2). When a saddle point is found, a branching node is added to the tree. Then, edges are added to connect the saddle point to the local optima representing the basins in which solutions are neighbors of the saddle point. From here, the saddle point now represents the basins of all adjacent local optima (step 3, the basins are merged by the flooding). This procedure is repeated until the last local optimum or saddle point has been found (step 4).

Since we are interested in the barriers that exist between the clusters of local optima in a landscape, we present a variant of the flooding algorithm suitable for this purpose in Section 5.5. Before, we need to explain how to characterize funnels in fitness landscapes. For this purpose, we introduce a special representation of fitness landscapes known as local optima networks (LONs) and a method using this representation to characterize funnels in the next Section 5.4.

5.4 Clusters of Local Optima in Fitness Landscapes

Local optima networks (LONs) are a novel approach to study the structure of fitness landscapes (Ochoa, Tomassini, et al., 2008) and have recently been used to reveal the structure of multiple clusters (Ochoa and Veerapen, 2016a; Ochoa and Veerapen, 2016b; Ochoa, Veerapen, et al., 2016; cf. also Chapter 4). LONs were originally inspired by the study of energy landscapes (Stillinger, 1995). A LON is a complex network in which the nodes represent the local optima in a landscape (and their basins, resp.). The edges reflect an algorithm's transition between the basins. The concept of LONs allows the study of fitness landscapes from a network perspective and has the potential to deepen our understanding of metaheuristics and problems.

A network is a graph $G = (V, E)$ with the set of vertices V and the set of edges E. In a LON, the vertex set V consists of the local optima of the fitness landscape. There exists an edge between two local optima if their basins are connected, leading to a potential transition between the two local optima. In particular, an escape edge (Verel et al., 2012) is defined by the distance function of the fitness landscape d (minimal number of moves between two solutions): there exists a directed edge e_{xy} from local optimum lo_x to lo_y if there is a solution s such that $d(s, lo_x) \leq D \wedge s \in B(lo_y)$. The weight w_{xy} of edge e_{xy} is the probability that a search algorithm can escape from the local optimum lo_x into the basin around lo_y. The constant $D > 0$ determines the maximum distance an algorithm uses during a perturbation step.

To reveal the clustering structure of fitness landscapes, we proposed to apply "community detection" to local optima networks (cf. Chapter 4). Community detection is an exploratory variant of graph partitioning (Talbi and Bessière, 1991). The objective of this method is to partition the network graph in a discipline-related, meaningful way. A very general definition of a community is a group of nodes that have more links among each other than to nodes in other communities. However, the definition of a community depends on the discipline applied and there exists a variety of algorithms that have been validated for different purposes (Fortunato, 2010; Porter et al., 2009).

Community detection in LONs has been done in earlier studies (Daolio, Tomassini, et al., 2011; Iclanzan et al., 2014). However, we found in Chapter 4 that in particular, the **M**arkov **Cl**uster algorithm (MCL; van Dongen, 2001) is an appropriate method of community detection to detect clusters in LONs and characterize the clustering structure of fitness landscapes. An explanation for this is that the MCL algorithm is based on stochastic flows. LONs model the stochastic process of an algorithm in a fitness landscape. For this reason, the application of the MCL algorithm matches the network model and produces meaningful results.

5.5 Coarse-Grained Barrier Trees of Fitness Landscapes

In order to escape from a cluster of local optima to another cluster, ILS needs to pass a barrier by a deterioration of the fitness. To visualize the structure of the barriers between the clusters in the landscape, we present a variant of the flooding algorithm (van Stein et al., 2013) as introduced in Section 5.3 and Figure 5.1. The pseudo code can be obtained from Algorithm 5.1.

Algorithm 5.1: Flooding Algorithm for LONs (Maximization Problem)

Require: Local Optima Network $G = (V, E)$, Partition $\mathcal{P}$ (cluster sets) over V
Let $R = \emptyset$
for all $P \in \mathcal{P}$ **do**
  Add the local maximum with max. fitness of P to R
end for{R contains one representing local optimum per cluster in $\mathcal{P}$}
Let $T = (V_{Tree}, E_{Tree})$ with $V_{Tree} = \emptyset, E_{Tree} = \emptyset$
Order V by f in descending order
for all $v \in V$ **do**
  if $v \in R$ **then**
    Add Node v to V_{Tree}
  else
    $\mathcal{C} = \{P \in \mathcal{P} : \exists\, n \in P : \ [(v, n) \in E \vee (n, v) \in E]\}$
    {Select those partition sets (clusters) which contain a local optimum adjacent to v in the LON graph}
    if $|\mathcal{C}| > 1$ **then** {v connects at least two clusters: v is a saddle point}
      Add Node v to V_{Tree}
      for all $C \in \mathcal{C}$ **do** {For each cluster set C connected to v}
        $r = C \cap R$ {Choose node r representing connected cluster set C}
        Add Edge (v, r) to E_{Tree}
        Update $\mathcal{P}$: Merge Partition sets containing v and c
        Remove r from R {Flood the connected cluster}
      end for
    end if
  end if
end for
return T

As an input, the algorithm accepts a LON and a partition of the LON's vertex set, i.e., a set with the clustering structure of the landscape. To obtain the clusters,

we propose to apply the Markov cluster algorithm to the LON. As a first step, the algorithm selects the best local optimum for each cluster (set R). Then, the set of local optima nodes V is ordered by fitness in descending order. The algorithm iterates over each node. If the node is a representing node (in R), it is added to the barrier tree. Else, the algorithm determines the number of clusters adjacent to the current node in the LON. If the number is higher than one, the node is a saddle point and is also added to the tree. Then, the algorithm connects the saddle point to the nodes representing the adjacent clusters in the tree. From here, the saddle point represents all adjacent clusters ("flooding"): the clusters of the current and all the adjacent nodes are merged in the partition set, and the representers of the adjacent clusters are removed from R. This process is repeated until the whole LON is flooded (merged into one partition).

To demonstrate our method, we selected an easy and a hard instance of the Kauffman NK model ($N = 20$, $K = 5$). To determine their difficulty, we performed 1000 independent runs of ILS per instance and measured the success rates (0.76 and 0.22). The ILS stopped after a limited number of fitness function evaluations ($1/5th$ of the search space), or when the global optimum was found. We extracted the LONs and computed the clusters in both LONs with the MCL algorithm. We used the LONs and the clusters to construct the coarse-grained barrier trees with our variant of the flooding algorithm. Figures 5.2 and 5.3 plot the LON and the corresponding tree. Visual inspection of the LONs (left) confirms that the clustering as obtained from the MCL algorithm is meaningful: nodes of the same color have a higher proximity to their own cluster than to those of a different cluster. Comparing both barrier trees (right), we observe a much deeper tree and thus a higher number of barriers in the case of the hard instance.

Even though a deeper study on the search difficulty is out of the scope of this paper, we conducted a first systematic approach towards this observation. We generated 200 instances of NK landscapes ($N = 20$, $K = 5$). We grouped the landscapes by the depth of the coarse-grained barrier tree and compared their difficulties for ILS. The results can be obtained from Figure 5.4. For landscapes with a very small tree, we observe that the difficulty has a high variety, even though the median indicates a low difficulty (≈ 0.6). The median success rates get lower with a deeper tree, which means that their difficulty increases. This is not surprising: a deeper tree means that a traversal to the global optimum has—by average—a longer path. A search algorithm needs to pass more barriers then, and the difficulty is higher. This finding is consistent with the previous literature on regular barrier trees (van Stein et al., 2013), however the observation that many landscapes with a low number of barriers can be difficult is counter-intuitive. We suggest that in these cases, additional factors, like the cluster size of the global optimum (Chapter 4) need to be considered. We plan to conduct more research towards this direction.

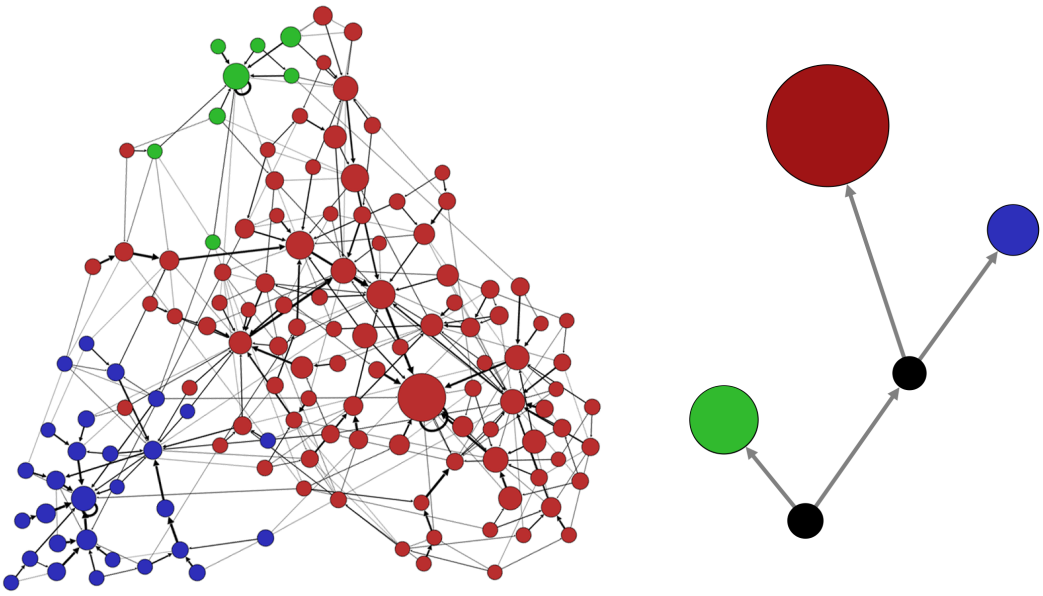

Figure 5.2: The Local optima network (left) and the coarse-grained barrier tree (right) of an NK landscape ($N = 20$, $K = 5$) with low search difficulty (success rate of ILS: 0.76). In the local optima network (left), the size of the node represents the fitness, whereas the node size in the tree (right) is the size of the cluster by the number of local optima. The color of the nodes indicates the cluster (global optimum cluster is red in both graph types). The black nodes in the tree are the saddle points. In the tree, the fitness is visualized by the node height (higher distance to the root means higher fitness). The layout of the local optima network is based on the ForceAtlas2 algorithm (Jacomy et al., 2014). The local optima network shows only the best 20% of the nodes (all clusters still visible).

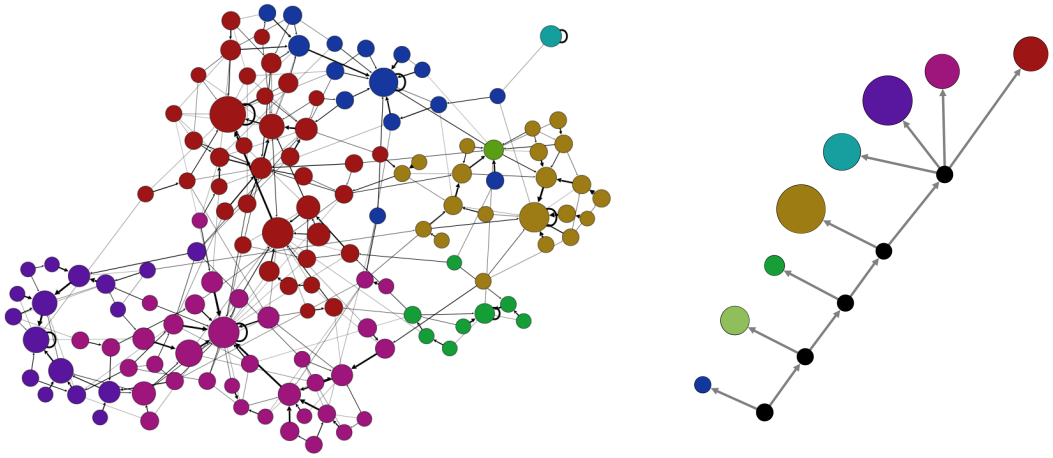

Figure 5.3: The LON (left) and the coarse-grained barrier tree (right) of an NK landscape ($N = 20$, $K = 5$) with high search difficulty (success rate of ILS: 0.22). Please cf. Figure 5.2 for further explanations.

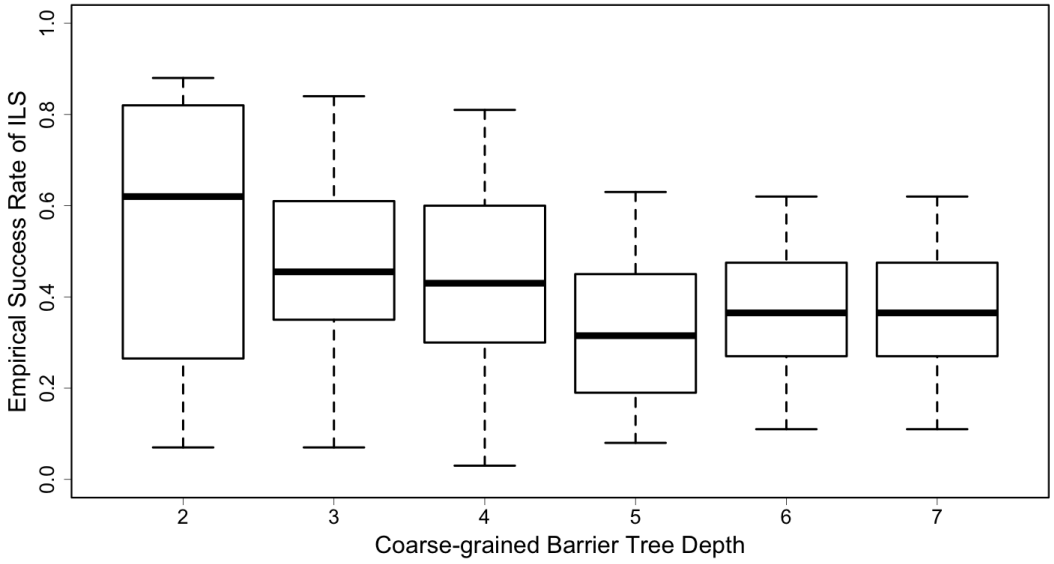

Figure 5.4: The success rate of ILS (search difficulty) for different values of tree depth. The median success rate initially declines (search difficulty increases) with a higher depth of the tree.

5.6 Summary and Conclusion

As our main contribution, we presented a new method to visualize fitness landscapes and characterize them by the barriers between clusters of local optima. The existence of a multiple-cluster structure has recently emerged (Ochoa and Veerapen, 2016b) as a refinement of the big valley hypothesis. We applied our method to a limited set of instances of the Kauffman NK model. Our results suggest that the tree depth might be related to the search difficulty of the landscapes for iterated local search. This is consistent with previous findings on difficulty in the literature (van Stein et al., 2013). A possible explanation is that the existence of barriers prevents iterated local search from escaping local optima clusters. This finding is rather preliminary and needs further investigation. Other structural properties of the landscapes must be taken into consideration, too. For further research, it would be interesting to see how the coarse-grained trees look for NK landscapes with higher levels of epistasis. It is also unclear whether or not there are differences between the NK model with random and adjacent co-variables. The adjacent NK model is often considered to be solvable with less effort. It would be worthwhile to examine if the tree depths are different between both models. We think that the method introduced here points to a new direction in studies of fitness landscapes.

References

Applegate, David, William Cook, and Andre Rohe (2003). Chained Lin-Kernighan for Large Traveling Salesman Problems. ***INFORMS Journal on Computing,*** 15(1): 82–92.

Becker, Oren M. and Martin Karplus (1997). The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. ***The Journal of chemical physics,*** 106(4): 1495–1517.

Daolio, Fabio, Marco Tomassini, Sébastien Verel, and Gabriela Ochoa (2011). Communities of minima in local optima networks of combinatorial spaces. ***Physica A: Statistical Mechanics and its Applications,*** 390(9): 1684–1694.

Doye, Jonathan P. K., Mark A. Miller, and David J. Wales (1999a). Evolution of the potential energy surface with size for Lennard-Jones clusters. ***The Journal of Chemical Physics,*** 111(18): 8417–8428.

Doye, Jonathan P. K., Mark A. Miller, and David J. Wales (1999b). The double-funnel energy landscape of the 38-atom Lennard-Jones cluster. ***The Journal of Chemical Physics,*** 110(14): 6896–6906.

Flamm, Christoph, Ivo L. Hofacker, Peter F. Stadler, and Michael T. Wolfinger (2002). Barrier Trees of Degenerate Landscapes. ***Zeitschrift für Physikalische Chemie,*** 216: 155–173.

Fortunato, Santo (2010). Community detection in graphs. ***Physics Reports,*** 486(3-5): 75–174.

Glover, Fred (1986). Future paths for integer programming and links to artificial intelligence. ***Computers & Operations Research,*** 13(5): 533–549.

Hains, Doug R., Darrel L. Whitley, and Adele E. Howe (2011). Revisiting the big valley search space structure in the TSP. ***Journal of the Operational Research Society,*** 62(2): 305–312.

Hallam, Jonathan and Adam Prügel-Bennett (2005). Large barrier trees for studying search. ***IEEE Transactions on Evolutionary Computation,*** 9(4): 385–397.

Iclanzan, David, Fabio Daolio, and Marco Tomassini (2014). Data-driven local optima network characterization of QAPLIB instances. In: ***Proceedings of the 2014 conference on Genetic and evolutionary computation - GECCO '14.*** Ed. by Christian Igel. Vancouver, BC, Canada: ACM Press: 453–460.

Jacomy, Mathieu, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. ***PLoS ONE,*** 9(6). Ed. by Mark R. Muldoon: e98679.

Kauffman, Stuart A. and Edward D. Weinberger (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. ***Journal of Theoretical Biology,*** 141(2): 211–245.

Lin, S. and B. W. Kernighan (1973). An Effective Heuristic Algorithm for the Traveling-Salesman Problem. ***Operations Research,*** 21(2): 498–516.

Lourenço, Helena R., Olivier C. Martin, and Thomas Stützle (2003). Iterated Local Search. In: ***Handbook of Metaheuristics.*** Boston: Kluwer Academic Publishers: 320–353.

Ochoa, Gabriela, Marco Tomassini, Sébastien Verel, and Christian Darabos (2008). A study of NK landscapes' basins and local optima networks. In: ***Proceedings of the 10th annual conference on Genetic and evolutionary computation - GECCO '08.*** Ed. by Maarten Keijzer. Atlanta, GA, USA: ACM Press: 555–562.

Ochoa, Gabriela and Nadarajen Veerapen (2016a). Additional Dimensions to the Study of Funnels in Combinatorial Landscapes. In: ***Proceedings of the 2016 Genetic and Evolutionary Computation Conference - GECCO '16.***

Ochoa, Gabriela and Nadarajen Veerapen (2016b). Deconstructing the Big Valley Search Space Hypothesis. In: ***Evolutionary Computation in Combinatorial Optimization: 16th European Conference, EvoCOP 2016, Porto, Portugal, March 30 – April 1, 2016, Proceedings.*** Ed. by Francisco Chicano, Bin Hu, and Pablo Garcia-Sanchez. Porto: Springer: 58–73.

Ochoa, Gabriela, Nadarajen Veerapen, Darrell Whitley, and Edmund K. Burke (2016). The Multi-Funnel Structure of TSP Fitness Landscapes: A Visual Exploration. In: ***Artificial Evolution: 12th International Conference, Evolution Artificielle, EA 2015.*** Lyon: Springer International Publishing: 1–13.

Porter, Mason a., Jukka-Pekka Onnela, and Peter J. Mucha (2009). Communities in Networks. ***Notices of the AMS,*** 486(3-5): 1082–1097.

Stillinger, Frank H. (1995). A Topographic View of Supercooled Liquids and Glass Formation. ***Science,*** 267(5206): 1935–1939.

Talbi, E.G. and P. Bessière (1991). A parallel genetic algorithm for the graph partitioning problem. In: ***Proceedings of the 5th international conference on Supercomputing - ICS '91.*** New York, USA: ACM Press: 312–320.

Van Dongen, Stijn (2001). "Graph clustering by flow simulation". PhD thesis. Utrecht University.

Van Stein, Bas, Michael Emmerich, and Zhiwei Yang (2013). Fitness Landscape Analysis of NK Landscapes and Vehicle Routing Problems by Expanded Barrier Trees. In: ***EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation.*** Ed. by Alexandru-Adrian Tantar, Emilia Tantar, Jian-Qiao Sun, Wei Zhang, Qian Ding, Oliver Schütze, Michael Emmerich, Pierrick Legrand, Pierre Del Moral, and Carlos A. Coello Coello. Vol. 227. Advances in Intelligent Systems and Computing. Springer: 75–89.

Verel, Sébastien, Fabio Daolio, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks with escape edges. In: ***Artificial Evolution.*** Angers, France: Springer: 49–60.

Weinberger, Edward D. (1990). Correlated and uncorrelated fitness landscapes and how to tell the difference. ***Biological cybernetics,*** 336: 325–336.

Wright, Sewall (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: ***Proceedings of the 6th International Congress of Genetics.*** Ed. by Donald F. Jones. Ithaca, New York: Morgan Kaufmann Publishers Inc.: 356–366.

Chapter 6

Summary and Conclusions

6.1 Summary

The purpose of my thesis is to examine how local optima networks (LONs) of fitness landscapes reflect the structure and search difficulty of combinatorial optimization problems for metaheuristics. Four papers were presented towards this objective. The first two articles applied a micro-level approach of complex network analysis. Their focus was on predicting the search difficulty of numerous problem instances with the concept of PageRank centrality. The third paper applied community detection, which is a macro-level concept of network analysis, to local optima networks. This revealed new, yet unknown structural properties of fitness landscapes. The last paper introduced a novel method to take a macroscopic perspective on the global structure of fitness landscapes. For this purpose, the concept of barrier trees was adopted and applied to local optima networks. This Section provides a brief summary of my work and the most relevant results.

PageRank centrality in LONs predicts difficulty. In the first and second study on LONs (Chapters 2 and 3), I showed that the *PageRank centrality of the global optimum* in a LON is a reliable predictor ($R^2 \gtrsim 90\%$) of the performance of metaheuristics. The PageRank is a significantly better predictor than standard metrics frequently used in the literature, such as ruggedness and deceptiveness. Thus, PageRank is a useful metric for the search difficulty of landscapes. In Chapter 2, we showed that the PageRank in LONs with edges of basin transition probabilities predicts the success rate of local search-based methods (simple hill climbing and simulated annealing). This is possible because PageRank is a variant of Eigenvector centrality. A LON approximates the stochastic process of an algorithm in the fitness landscape. The LON graph's matrix of edge weights is equivalent to the transition matrix of a finite-state Markov chain. The Eigenvector of the transition matrix reflects the stationary distribution of a random walk across the Markov chain. Hence, the scalar value of the global optimum approximates the probability to pass it.

In the study outlined in Chapter 3, we used a similar experimental setup to predict the success rate of iterated local search. As predictor, we tested the PageRank in LONs with escape edges, extracted from multiple instances of the NK model. The results were in accordance with the first study and showed a high predictive power of our measure for search difficulty. Again, the explanation is that the escape edges model the stochastic process of iterated local search: the potential transitions between local optima basins reflect the behavior of the perturbation operator. Another finding is that we could use the PageRank to predict the expected solution quality: we computed a weighted average score of the fitness of all local optima, using the PageRank vector of the local optima as weights. This measure was strongly correlated to the average fitness achieved by empirical runs of iterated local search. An explanation is that the PageRank vector describes the stationary distribution of the search algorithm in the fitness landscape. Hence, it is reasonable to use the probabilities of the distribution as weights. In addition, we used the PageRank of the global optimum to predict the average running time of iterated local search to locate the global optimum. This worked especially well when the landscapes had a high level of epistasis (interdependencies between optimization variables). This is because in such landscapes, the basins around the local optima are small. Then, iterated local search uses most of the running time to apply the perturbation operator instead of local search.

Clusters of local optima. Another aspect of my thesis was the study of landscape structure by network analysis. An important property of the landscape structure is how the local optima are distributed. In the paper presented in Chapter 4, we applied the *Markov cluster algorithm for community detection* to LONs. This revealed a structure of multiple clusters. The existence of multiple clusters complements the big valley hypothesis, which states that good solutions are often contained in a single, giant cluster. Our findings are consistent with a similar study recently conducted by Ochoa and Veerapen (2016b) on LONs, sampled from large instances of the traveling salesman problem. The results also show that the existence of multiple clusters explains search difficulty for iterated local search: the size of the cluster containing the global optimum is strongly correlated to the empirical success rate. This offers an explanation for the observation that iterated local search often finds high-quality solutions, but fails to locate the global optimum. The clusters of local optima are strongly connected within themselves, but connections between clusters are sparse. Since the perturbation operator moves the algorithm from one local optimum basin to another, it is unlikely to move along the sparse connections and escape from one cluster to another. Since our implementation of iterated local search used random starting points, a larger cluster containing the global optimum increases the probability to start in this cluster and find the global optimum.

Coarse-grained barrier trees. In the last paper (Chapter 5), I presented a new approach to study fitness landscapes on a coarse level of granularity. Our method computes a *coarse-grained barrier tree* of a landscape which retains the global structure and allows the eventual visualization of larger landscapes. Barrier trees (Flamm et al., 2002) or disconnectivity graphs (Becker and Karplus, 1997) have so far been used to study the barriers which an algorithm needs to pass in order to escape from one basin to another (Hallam and Prügel-Bennett, 2005). The method proposed here calculates the barriers that exist between clusters of local optima (instead of basins). Our algorithm to calculate the barriers is a new variant of the *flooding algorithm* by van Stein et al. (2013). As an input, it accepts a LON and a cluster structure. A method to reveal the cluster structure was already introduced in our previous study (Chapter 4). The output is a coarse-grained picture of the fitness landscape in form of a tree. We tested our method with instances of the Kauffman NK model and visualized their structure. We also examined the relationship between tree structure and search difficulty for iterated local search. Our results suggest that landscapes with deeper barrier trees have a higher search difficulty. A possible explanation is that a deeper tree means that a traversal to the global optimum has—by average—a longer path. A search algorithm needs to pass more barriers then, and the difficulty is higher.

6.2 Conclusions

The last Section of my thesis summarizes the contributions and reflects on the limitations. I will also refer to aspects that might be interesting for future research.

Deeper insights into LONs and difficulty. The studies presented in my thesis have shed light onto the open question of how properties of LONs are related to search difficulty. This adds a missing piece to the puzzle in which Daolio, Verel, et al. (2012) already achieved significant progress: they found that shorter average path lengths to the global optimum are related to lower search difficulty. This idea is quite close to the concept of centrality, which we used in our studies. Thus, we provide a valuable contribution to this topic in the recent literature.

Highly accurate predictors for search difficulty. We identified a highly accurate, ex-ante method to classify instances of fitness landscapes by their search difficulty. Our studies found that the metrics *PageRank* and *cluster size of the global optimum* have a higher predictive power than the standard metrics which are frequently used in the literature, e.g. *ruggedness* and *deceptiveness*. This might be helpful for future experiments using fitness landscapes, which is relevant for multi-

ple disciplines beyond the field of metaheuristics. For future research, it would be interesting to study if the PageRank values in LONs with basin transition probabilities and LONs with escape edges are correlated. If so, this should indicate a higher validity of PageRank as a candidate for a more universal measure of search difficulty. Such a measure could be useful to determine the difficulty for more popular metaheuristics and principles of heuristic search.

Understanding LONs as stochastic models. Our result on PageRank centrality and search difficulty also facilitates a deeper ontological understanding of LONs. So far, our comprehension of the nature of LONs has been rather vague: this becomes manifest in the exploratory attempts in the recent literature to identify network metrics as predictors of search difficulty. Towards a clarification, we need to consider the results of the first two studies on PageRank centrality. First, edges with basin transition probabilities model the short-distance connections between local optima basins. This reflects the behavior of a metaheuristic that is based on local search. Second, escape edges are defined by a longer distance between basins. The perturbation operator of iterated local search bridges between basins in a way that can be characterized by escape edges. This explains why LONs with basin transitions predicts the success rate of local search with high accuracy, which holds for the combination of escape edges and the performance of iterated local search, too. In a nutshell, LONs are useful to predict the behavior of an algorithm if the edges are modeled according to the rule set of the algorithm. This observation sharpens our understanding of the nature of a LON: it is a model to describe the stochastic process of a particular principle of heuristic search in a fitness landscape.

A complexity perspective on fitness landscapes. The results presented in this thesis shows that a *complexity perspective* on fitness landscapes reveals meaningful and yet unknown aspects. The PageRank in LONs combines the diverse structural properties of fitness landscapes in a single measure. Assuming that a fitness landscape and a search algorithm establish a complex system, it is possible to predict the emerging behavior of the complex system, i.e., the probability to reach a certain state. Reductionist concepts—like ruggedness—reflect only one of numerous structural aspects. This explains why these concepts have poor accuracy. Our results strengthen the right to exist of the complexity approach studies in metaheuristics, in particular in the mode of LONs.

New structural properties of fitness landscapes. The results obtained from our cluster analysis shows that the global structure of fitness landscapes is more nuanced than previously thought. An implication is that we need to rethink the

hypothesis of a single, giant cluster. The existence of multiple clusters also explains why iterated local search may find high-quality solutions, but fails to locate the global optimum. This is important since the principle of this metaheuristic is implemented in many problem-specific heuristics, e.g. the Chained Lin-Kernighan heuristic (Applegate et al., 2003; Lin and Kernighan, 1973).

A new approach to study fitness landscapes. The method introduced to compute the barrier trees exploits our finding that local optima are clustered. It is a new tool to draw a coarse-grained picture of landscapes which retains the global structure. There are two conceivable applications of this approach: (i) for future research on fitness landscapes and metaheuristics, it allows the eventual visualization of larger landscapes, and (ii) it is a general new method to convert—or rather compress—a clustered, complex network into a hierarchical tree structure. This might be useful for other disciplines using network studies, as well.

Limitations and further research. A clear limitation of our studies is the focus on the Kauffman NK model. Future research should study if the results presented here are transferable to other problem types. For instance, there is some evidence towards the traveling salesman problem in the recent literature (Ochoa and Veerapen, 2016b). In this context, it would also be interesting to study LONs for other search operators, e.g. the tunneling crossover networks for genetic algorithms (Ochoa, Chicano, et al., 2015). General structural properties of fitness landscapes are highly relevant for metaheuristics. A direction of further research could be to exploit the methods and structural properties presented here, for the purpose of constructing and improving heuristic search. Finally, the observation that features of local optima networks are correlated to search difficulty does not imply causality. Further research is necessary to strengthen our understanding of the dynamics and trajectories of search algorithms in fitness landscapes.

Closing remarks. In summary, the methods presented in this thesis are supposed to enrich the portfolio available for fitness landscape analysis. The results also contribute to our understanding of the relationship between problem structure, fitness landscapes and LONs. Due to the interdisciplinary applications of fitness landscapes and complex networks, the findings and methods proposed may also be useful beyond the field of evolutionary computation, e.g. in theoretical biology, complex adaptive systems and organizational theory.

Bibliography

Albert, Réka and Albert-László Barabási (2002). Statistical mechanics of complex networks. ***Reviews of Modern Physics,*** 74(1): 47–97.

Applegate, David, William Cook, and Andre Rohe (2003). Chained Lin-Kernighan for Large Traveling Salesman Problems. ***INFORMS Journal on Computing,*** 15(1): 82–92.

Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. In: ***Third International AAAI Conference on Weblogs and Social Media.***: 361–362.

Becker, Oren M. and Martin Karplus (1997). The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. ***The Journal of chemical physics,*** 106(4): 1495–1517.

Billinger, Stephan, Nils Stieglitz, and Terry R. Schumacher (2014). Search on Rugged Landscapes: An Experimental Study. ***Organization Science,*** 25(1): 93–108.

Boccaletti, Stefano, Vito Latora, Yamir Moreno, Martín Chávez Hoffmeister, and Dong-Uk Hwang (2006). Complex networks: Structure and dynamics. ***Physics Reports,*** 424(4-5): 175–308.

Boese, Kenneth D., Andrew B. Kahng, and Sudhakar Muddu (1994). A new adaptive multi-start technique for combinatorial global optimizations. ***Operations Research Letters,*** 16(2): 101–113.

Bonacich, Phillip (2007). Some unique properties of eigenvector centrality. ***Social Networks,*** 29(4): 555–564.

Borgatti, Stephen P. (2005). Centrality and network flow. ***Social Networks,*** 27(1): 55–71.

Borgatti, Stephen P. and Martin G. Everett (2006). A Graph-theoretic perspective on centrality. ***Social Networks,*** 28(4): 466–484.

Borgatti, Stephen P., Ajay Mehra, Daniel J. Brass, and Giuseppe Labianca (2009). Network analysis in the social sciences. ***Science,*** 323(5916): 892–895.

Brin, Sergey and Lawrence Page (1998). The anatomy of a large-scale hypertextual web search engine. ***Computer Networks and ISDN Systems,*** 30(1-7): 107–117.

Chicano, Francisco, Fabio Daolio, Gabriela Ochoa, Sébastien Verel, Marco Tomassini, and Enrique Alba (2012). Local Optima Networks, Landscape Autocorrelation and Heuristic Search Performance. In: ***Parallel Problem Solving from Nature - PPSN XII: 12th International Conference.*** Ed. by Carlos A. Coello Coello, Vincenzo Cutello, Kalyanmoy Deb, Stephanie Forrest, Giuseppe Nicosia, and Mario Pavone. Vol. 7492 LNCS. Berlin and Heidelberg: Springer: 337–347.

Choi, Sung-Soon and Byung-Ro Moon (2008). Normalization for Genetic Algorithms With Nonsynonymously Redundant Encodings. ***IEEE Transactions on Evolutionary Computation,*** 12(5): 604–616.

Daolio, Fabio, Marco Tomassini, Sébastien Verel, and Gabriela Ochoa (2011). Communities of minima in local optima networks of combinatorial spaces. ***Physica A: Statistical Mechanics and its Applications,*** 390(9): 1684–1694.

Daolio, Fabio, Sébastien Verel, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks and the performance of iterated local search. In: ***Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference - GECCO '12.*** Ed. by Terence Soule. Philadelphia, Pennsylvania, USA: ACM Press: 369.

Doye, Jonathan P. K. and Claire P. Massen (2005). Characterizing the network topology of the energy landscapes of atomic clusters. ***Journal of Chemical Physics,*** 122(8): 084105.

Doye, Jonathan P. K., Mark A. Miller, and David J. Wales (1999a). Evolution of the potential energy surface with size for Lennard-Jones clusters. ***The Journal of Chemical Physics,*** 111(18): 8417–8428.

Doye, Jonathan P. K., Mark A. Miller, and David J. Wales (1999b). The double-funnel energy landscape of the 38-atom Lennard-Jones cluster. ***The Journal of Chemical Physics,*** 110(14): 6896–6906.

Easley, David and Jon Kleinberg (2010). ***Networks, crowds, and markets: Reasoning about a highly connected world.*** 1st ed. Cambridge: Cambridge Univ. Press.

Flamm, Christoph, Ivo L. Hofacker, Peter F. Stadler, and Michael T. Wolfinger (2002). Barrier Trees of Degenerate Landscapes. ***Zeitschrift für Physikalische Chemie,*** 216: 155–173.

Fortunato, Santo (2010). Community detection in graphs. ***Physics Reports,*** 486(3-5): 75–174.

Franceschet, Massimo (2011). PageRank: standing on the shoulders of giants. ***Communications of the ACM,*** 54(6): 92–101.

Freeman, Linton C. (1979). Centrality in social networks conceptual clarification. ***Social Networks,*** 1(3): 215–239.

Freeman, Linton C. (2004). ***The Development of Social Network Analysis: A Study in the Sociology of Science.*** Empirical Press.

Frobenius, Ferdinand Georg (1912). Ueber Matrizen aus nicht negativen Elementen. ***Sitzungsberichte Preussische Akademie der Wissenschaft, Berlin,*** 456–477.

Glover, Fred (1986). Future paths for integer programming and links to artificial intelligence. ***Computers & Operations Research,*** 13(5): 533–549.

Hagberg, Aric A., Daniel A. Schult, and Pieter J. Swart (2008). Exploring network structure, dynamics, and function using NetworkX. ***Proceedings of the 7th Python in Science Conference (SciPy 2008),*** 11–15.

Hains, Doug R., Darrel L. Whitley, and Adele E. Howe (2011). Revisiting the big valley search space structure in the TSP. ***Journal of the Operational Research Society,*** 62(2): 305–312.

Hallam, Jonathan and Adam Prügel-Bennett (2005). Large barrier trees for studying search. ***IEEE Transactions on Evolutionary Computation,*** 9(4): 385–397.

He, Jun, Colin Reeves, Carsten Witt, and Xin Yao (2007). A note on problem difficulty measures in black-box optimization: classification, realizations and predictability. ***IEEE Transactions on Evolutionary Computation,*** 15(4): 435–43.

Holland, John H. (1975). ***Adaptation in Natural and Artificial Systems.*** MIT Press Cambridge.

Holland, John H. (1992). Complex Adaptive Systems. ***Daedalus,*** 121(1): 17–30.

Iclanzan, David, Fabio Daolio, and Marco Tomassini (2014). Data-driven local optima network characterization of QAPLIB instances. In: ***Proceedings of the 2014 conference on Genetic and evolutionary computation - GECCO '14.*** Ed. by Christian Igel. Vancouver, BC, Canada: ACM Press: 453–460.

Jacomy, Mathieu, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network

Visualization Designed for the Gephi Software. ***PLoS ONE,*** 9(6). Ed. by Mark R. Muldoon: e98679.

Jones, Terry and Stephanie Forrest (1995). Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms. In: ***Proceedings of the Sixth International Conference on Genetic Algorithms.*** San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.: 184–192.

Kallel, L., B. Naudts, and Colin R. Reeves (2001). Properties of Fitness Functions and Search Landscapes. In: ***Theoretical Aspects of Evolutionary Computing.*** Ed. by Leila Kallel, Bart Naudts, and Alex Rogers. Natural Computing Series. Berlin and Heidelberg: Springer: 175–206.

Kauffman, Stuart A. (1993). ***The origins of order.*** Oxford University Press.

Kauffman, Stuart A. and Simon Levin (1987). Towards a General Theory of Adaptive Walks on Rugged Landscapes. ***Journal of Theoretical Biology,*** 128(1): 11–45.

Kauffman, Stuart A. and Edward D. Weinberger (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. ***Journal of Theoretical Biology,*** 141(2): 211–245.

Kerschke, Pascal, Mike Preuss, Simon Wessing, and Heike Trautmann (2015). Detecting Funnel Structures by Means of Exploratory Landscape Analysis. In: ***Proceedings of the 2015 Genetic and Evolutionary Computation Conference - GECCO '15.*** Ed. by Sara Silva. Madrid, Spain: ACM Press: 265–272.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi (1983). Optimization by simulated annealing. ***Science,*** 220(4598): 671–80.

Laarhoven, P. J. M. and E. H. L. Aarts (1988). ***Simulated Annealing: Theory and Applications.*** Norwell, MA, USA: Kluwer Academic Publishers.

Lazer, David and Allan Friedman (2007). The Network Structure of Exploration and Exploitation. ***Administrative Science Quarterly,*** 52: 667–694.

Lehmer, Derrick H. (1960). Teaching combinatorial tricks to a computer. In: ***Proceedings of the Symposium on Applied Mathematics and Combinatorial Analysis.***

Lenstra, Jan Karel and A. H. G. Rinnooy Kan (1981). Complexity of vehicle routing and scheduling problems. ***Networks,*** 11(2): 221–227.

Levinthal, Daniel A. (1997). Adaptation on Rugged Landscapes. ***Management Science,*** 43(7): 934–950.

Lin, S. and B. W. Kernighan (1973). An Effective Heuristic Algorithm for the Traveling-Salesman Problem. ***Operations Research,*** 21(2): 498–516.

Locatelli, M. (2005). On the Multilevel Structure of Global optimization problems. ***Computational Optimization and Applications,*** 30(1): 5–22.

Lourenço, Helena R., Olivier C. Martin, and Thomas Stützle (2003). Iterated Local Search. In: ***Handbook of Metaheuristics.*** Boston: Kluwer Academic Publishers: 320–353.

Lu, Guanzhou, Jinlong Li, and Xin Yao (2014). Fitness Landscapes and Problem Difficulty in Evolutionary Algorithms: From Theory to Applications. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Emergence, Complexity and Computation. Berlin and Heidelberg: Springer: 133–152.

Lunacek, Monte and Darrell Whitley (2006). The dispersion metric and the CMA evolution strategy. In: ***Proceedings of the 8th annual conference on Genetic and evolutionary computation - GECCO '06.*** New York, New York, USA: ACM Press: 477.

Malan, Katherine M. and Andries P. Engelbrecht (2014). Fitness Landscape Analysis for Metaheuristic Performance Prediction. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Berlin and Heidelberg: Springer: 103–129.

Mason, Winter and Duncan J. Watts (2012). Collaborative learning in networks. ***Proceedings of the National Academy of Sciences,*** 109(3): 764–769.

Massen, Claire P. and Jonathan P. K. Doye (2005). Identifying communities within energy landscapes. ***Physical Review E - Statistical, Nonlinear, and Soft Matter Physics,*** 71(4): 1–13.

Middendorf, Martin, Frank Reischle, and Hartmut Schmeck (2002). Multi Colony Ant Algorithms. ***Journal of Heuristics,*** 8(3): 305–320.

Miller, John H. and Scott E. Page (2007). ***Complex Adaptive Systems: An Introduction to Computational Models of Social Life.*** Princeton. Princeton University Press.

Naudts, B. and L. Kallel (2000). A comparison of predictive measures of problem difficulty in evolutionary algorithms. ***IEEE Transactions on Evolutionary Computation,*** 4(1): 1–15.

Newman, Mark E. J. (2006). Modularity and community structure in networks. ***Proceedings of the National Academy of Sciences,*** 103(23): 8577–8582.

Newman, Mark E. J. and Michelle Girvan (2004). Finding and evaluating community structure in networks. ***Physical Review E - Statistical, Nonlinear, and Soft Matter Physics,*** 69(2): 026113.

Ochoa, Gabriela, Francisco Chicano, Renato Tinós, and Darrell Whitley (2015). Tunnelling Crossover Networks. In: ***Proceedings of the 2015 Genetic and Evolutionary Computation Conference - GECCO '15.*** Ed. by Sara Silva. Madrid, Spain: ACM Press: 449–456.

Ochoa, Gabriela, Marco Tomassini, Sébastien Verel, and Christian Darabos (2008). A study of NK landscapes' basins and local optima networks. In: ***Proceedings of the 10th annual conference on Genetic and evolutionary computation - GECCO '08.*** Ed. by Maarten Keijzer. Atlanta, GA, USA: ACM Press: 555–562.

Ochoa, Gabriela and Nadarajen Veerapen (2016a). Additional Dimensions to the Study of Funnels in Combinatorial Landscapes. In: ***Proceedings of the 2016 Genetic and Evolutionary Computation Conference - GECCO '16.***

Ochoa, Gabriela and Nadarajen Veerapen (2016b). Deconstructing the Big Valley Search Space Hypothesis. In: ***Evolutionary Computation in Combinatorial Optimization: 16th European Conference, EvoCOP 2016, Porto, Portugal, March 30 – April 1, 2016, Proceedings.*** Ed. by Francisco Chicano, Bin Hu, and Pablo Garcia-Sanchez. Porto: Springer: 58–73.

Ochoa, Gabriela, Nadarajen Veerapen, Darrell Whitley, and Edmund K. Burke (2016). The Multi-Funnel Structure of TSP Fitness Landscapes: A Visual Exploration. In: ***Artificial Evolution: 12th International Conference, Evolution Artificielle, EA 2015.*** Lyon: Springer International Publishing: 1–13.

Ochoa, Gabriela, Sébastien Verel, Fabio Daolio, and Marco Tomassini (2014). Local Optima Networks: A New Model of Combinatorial Fitness Landscapes. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Vol. 6. Emergence, Complexity and Computation. Berlin and Heidelberg: Springer: 233–262.

Ochoa, Gabriela, Sébastien Verel, and Marco Tomassini (2010). First-improvement vs. best-improvement local optima networks of nk landscapes. In: ***PPSN'10: Proceedings of the 11th International Conference on Parallel Problem Solving from Nature.*** Ed. by Robert Schaefer, Carlos Cotta, Joanna Kolodziej, and Günter Rudolph. Vol. I. Kraków, Poland: Springer: 104–113.

Oestreicher-Singer, Gal, Barak Libai, Liron Sivan, Eyal Carmi, and Ohad Yassin (2013). The Network Value of Products. ***Journal of Marketing,*** 77(3): 1–14.

Oestreicher-Singer, Gal and Arun Sundararajan (2012). Recommendation Networks and the Long Tail of Electronic Commerce. ***MIS Quarterly,*** 36(1): 65–83.

Oliphant, Travis E. (2007). Python for scientific computing. ***Computing in Science and Engineering,*** 9(3): 10–20.

Perron, Oskar (1907). Zur Theorie der Matrices. ***Mathematische Annalen 1,*** 64(1): 248–263.

Pitzer, Erik and Michael Affenzeller (2012). A Comprehensive Survey on Fitness Landscape Analysis. In: ***Recent Advances in Intelligent Engineering Systems.*** Ed. by János Fodor, Ryszard Klempous, and Carmen Paz Suárez Araujo. Vol. 378. Studies in Computational Intelligence. Berlin and Heidelberg: Springer: 161–191.

Porter, Mason a., Jukka-Pekka Onnela, and Peter J. Mucha (2009). Communities in Networks. ***Notices of the AMS,*** 486(3-5): 1082–1097.

R Development Core Team (2009). ***R: A Language and Environment for Statistical Computing.***

Rechenberg, Ingo (1973). ***Evolutionsstrategie Optimierung technischer Systeme nach Prinzipien der biologischen Evolution.*** Stuttgart: Frommann-Holzboog.

Reeves, Colin R. and Jonathan E. Rowe (2002). ***Genetic Algorithms: Principles and Perspectives.*** Vol. 20. Operations Research/Computer Science Interfaces Series. Boston: Kluwer Academic Publishers.

Reidys, Christian M. and Peter F. Stadler (2002). Combinatorial Landscapes. ***SIAM Review,*** 44(1): 3–54.

Richter, Hendrik (2014). Fitness Landscapes: From Evolutionary Biology to Evolutionary Computation. In: ***Recent Advances in the Theory and Application of Fitness Landscapes.*** Ed. by Hendrik Richter and Andries Engelbrecht. Berlin and Heidelberg: Springer: 3–31.

Rivkin, Jan W (2000). Imitation of Complex Strategies. ***Management Science,*** 46(6): 824–844.

Rothlauf, Franz (2011). ***Design of modern heuristics: Principles and application.*** Berlin and Heidelberg: Springer.

Satuluri, Venu, Srinivasan Parthasarathy, and Duygu Ucar (2010). Markov clustering of protein interaction networks with improved balance and scalability. In: ***Proceedings of the First ACM International Conference on Bioin-***

formatics and Computational Biology - BCB '10. New York, USA: ACM Press: 247.

Schwefel, Hans-Paul (1977). ***Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie.*** Basel: Birkhäuser.

Stadler, Peter F. (1996). Landscapes and their correlation functions. ***Journal of Mathematical Chemistry,*** 20(1): 1–45.

Stillinger, Frank H. (1995). A Topographic View of Supercooled Liquids and Glass Formation. ***Science,*** 267(5206): 1935–1939.

Talbi, E.G. and P. Bessière (1991). A parallel genetic algorithm for the graph partitioning problem. In: ***Proceedings of the 5th international conference on Supercomputing - ICS '91.*** New York, USA: ACM Press: 312–320.

Valente, Thomas W. (1996). Network models of the diffusion of innovations. ***Computational and Mathematical Organization Theory,*** 2(2): 134.

Van Dongen, Stijn (2001). "Graph clustering by flow simulation". PhD thesis. Utrecht University.

Van Stein, Bas, Michael Emmerich, and Zhiwei Yang (2013). Fitness Landscape Analysis of NK Landscapes and Vehicle Routing Problems by Expanded Barrier Trees. In: ***EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation.*** Ed. by Alexandru-Adrian Tantar, Emilia Tantar, Jian-Qiao Sun, Wei Zhang, Qian Ding, Oliver Schütze, Michael Emmerich, Pierrick Legrand, Pierre Del Moral, and Carlos A. Coello Coello. Vol. 227. Advances in Intelligent Systems and Computing. Springer: 75–89.

Verel, Sébastien, Fabio Daolio, Gabriela Ochoa, and Marco Tomassini (2012). Local optima networks with escape edges. In: ***Artificial Evolution.*** Angers, France: Springer: 49–60.

Weinberger, Edward D. (1990). Correlated and uncorrelated fitness landscapes and how to tell the difference. ***Biological cybernetics,*** 336: 325–336.

Wolpert, David H. and William G. Macready (1997). No free lunch theorems for optimization. ***IEEE Transactions on Evolutionary Computation,*** 1(1): 67–82.

Wright, Sewall (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: ***Proceedings of the 6th International Congress of Genetics.*** Ed. by Donald F. Jones. Ithaca, New York: Morgan Kaufmann Publishers Inc.: 356–366.