Publications of Ralf Rothenberger

This document lists all peer-reviewed publications of Ralf Rothenberger, Chair for Algorithm Engineering, Hasso Plattner Institute, Potsdam, Germany. This listing was automatically generated on February 28, 2021. An up-to-date version is available online at hpi.de/friedrich/docs/publist/rothenberger.pdf.

Journal articles

[1] Chauhan, A., Friedrich, T., Rothenberger, R., Greed is Good for Deterministic Scale-Free Networks. In: *Algorithmica* 82, pp. 3338–3389, 2020.

Large real-world networks typically follow a power-law degree distribution. To study such networks, numerous random graph models have been proposed. However, real-world networks are not drawn at random. Therefore, Brach, Cygan, Lacki, and Sankowski [SODA 2016] introduced two natural deterministic conditions: (1) a power-law upper bound on the degree distribution (PLB-U) and (2) power-law neighborhoods, that is, the degree distribution of neighbors of each vertex is also upper bounded by a power law (PLB-N). They showed that many real-world networks satisfy both deterministic properties and exploit them to design faster algorithms for a number of classical graph problems. We complement the work of Brach et al. by showing that some well-studied random graph models exhibit both the mentioned PLB properties and additionally also a power-law lower bound on the degree distribution (PLB-L). All three properties hold with high probability for Chung-Lu Random Graphs and Geometric Inhomogeneous Random Graphs and almost surely for Hyperbolic Random Graphs. As a consequence, all results of Brach et al. also hold with high probability or almost surely for those random graph classes. In the second part of this work we study three classical NP-hard combinatorial optimization problems on PLB networks. It is known that on general graphs with maximum degree Δ, a greedy algorithm, which chooses nodes in the order of their degree, only achieves a $\Omega(\ln \Delta)$-approximation for Minimum Vertex Cover and Minimum Dominating Set, and a $\Omega(\Delta)$-approximation for Maximum Independent Set. We prove that the PLB-U property suffices for the greedy approach to achieve a constant-factor approximation for all three problems. We also show that all three combinatorial optimization problems are APX-complete even if all PLB-properties holds hence, PTAS cannot be expected unless P=NP.

[2] Friedrich, T., Krejca, M. S., Rothenberger, R., Arndt, T., Hafner, D., Kellermeier, T., Krogmann, S., Razmjou, A., Routing for On-Street Parking Search using Probabilistic Data. In: *AI Communications* 32, pp. 113–124, 2019.

A significant percentage of urban traffic is caused by the search for parking spots. One possible approach to improve this situation is to guide drivers along routes which are likely to have free parking spots. The task of finding such a route can be modeled as a probabilistic graph problem which is NP-complete. Thus, we propose heuristic approaches for solving this problem and evaluate them experimentally. For this, we use probabilities of finding a parking spot, which are based on publicly available empirical data from TomTom International B.V. Additionally, we propose a heuristic that relies exclusively on conventional road attributes. Our experiments show that this algorithm comes close to the baseline by a factor of 1.3 in our cost measure. Last, we complement our experiments with results from a field study, comparing the success rates of our algorithms against real human drivers.

Conference papers

[3] Bläsius, T., Friedrich, T., Göbel, A., Levy, J., Rothenberger, R., The Impact of Heterogeneity and Geometry on the Proof Complexity of Random Satisfiability. In: *Symposium on Discrete Algorithms (SODA)*, 2021.

Satisfiability is considered the canonical NP-complete problem and is used as a starting point for hardness reductions in theory, while in practice heuristic SAT solving algorithms can solve large-scale industrial SAT instances very efficiently. This disparity between theory and practice is believed to be a result of inherent properties of industrial SAT instances that make them tractable. Two characteristic properties seem to be prevalent in the majority of real-world SAT instances, heterogeneous degree distribution and locality. To understand the impact of these two properties on SAT, we study the proof complexity of random k-SAT models that allow to control heterogeneity and locality. Our findings show that heterogeneity alone does not make SAT easy as heterogeneous random k-SAT instances have superpolynomial resolution size. This implies intractability of these instances for modern SAT-solvers. On the other hand, modeling locality with an underlying geometry leads to small unsatisfiable subformulas, which can be found within polynomial time. A key ingredient for the result on geometric random k-SAT can be found in the complexity of higher-order Voronoi diagrams. As an additional technical contribution, we show an upper bound on the number of non-empty Voronoi regions, that holds for points with random positions in a very general setting. In particular, it covers arbitrary p-norms, higher dimensions, and weights affecting the area of influence of each point multiplicatively. Our bound is linear in the total weight. This is in stark contrast to quadratic lower bounds for the worst case.

[4] Friedrich, T., Göbel, A., Neumann, F., Quinzan, F., Rothenberger, R., Greedy Maximization of Functions with Bounded Curvature Under Partition Matroid Constraints. In: *Conference on Artificial Intelligence (AAAI)*, pp. 2272–2279, 2019.

We investigate the performance of a deterministic GREEDY algorithm for the problem of maximizing functions under a partition matroid constraint. We consider non-monotone submodular functions and monotone subadditive functions. Even though constrained maximization problems of monotone submodular functions have been extensively studied, little is known about greedy maximization of non-monotone submodular functions or monotone subadditive functions. We give approximation guarantees for GREEDY on these problems, in terms of the curvature. We find that this simple heuristic yields a strong approximation guarantee on a broad class of functions. We discuss the applicability of our results to three real-world problems: Maximizing the determinant function of a positive semidefinite matrix, and related problems such as the maximum entropy sampling problem, the constrained maximum cut problem on directed graphs, and combinatorial auction games. We conclude that GREEDY is well-suited to approach these problems. Overall, we present evidence to support the idea that, when dealing with constrained maximization problems with bounded curvature, one needs not search for (approximate) monotonicity to get good approximate solutions.

[5] Friedrich, T., Rothenberger, R., The Satisfiability Threshold for Non-Uniform Random 2-SAT. In: *International Colloquium on Automata, Languages and Programming (ICALP)*, pp. 61:1–61:14, 2019.

Propositional satisfiability (SAT) is one of the most fundamental problems in computer science. Its worst-case hardness lies at the core of computational complexity theory, for example in the form of NP-hardness and the (Strong) Exponential Time Hypothesis. In practice however, SAT instances can often be solved efficiently. This contradicting behavior has spawned interest in the average-case analysis of SAT and has triggered the development of sophisticated rigorous and non-rigorous techniques for analyzing random structures. Despite a long line of research and substantial progress, most theoretical work on random SAT assumes a uniform distribution on the variables. In contrast, real-world instances often exhibit large fluctuations in variable occurrence. This can be modeled by a non-uniform distribution of the variables, which can result in distributions closer to industrial SAT instances. We study satisfiability thresholds of non-uniform random 2-SAT with n variables and m clauses and with an arbitrary probability distribution $(p_i)_{i\in[n]}$ with $p_1 \geq p_2 \geq \cdots \geq p_n > 0$ over the n variables. We show for $p_1^2 = \Theta(\sum_{i=1}^n p_i^2)$ that the asymptotic satisfiability threshold is at $m = \Theta((1-\sum_{i=1}^n p_i^2)/(p_1(\sum_{i=2}^n p_i^2)^{1/2}))$ and that it is coarse. For $p_1^2 = o(\sum_{i=1}^n p_i^2)$ we show that there is a sharp satisfiability threshold at $m = (\sum_{i=1}^n p_i^2)^{-1}$. This result generalizes the seminal works by Chvatal and Reed [FOCS 1992] and by Goerdt [JCSS 1996].

[6] Peters, J., Stephan, D., Amon, I., Gawendowicz, H., Lischeid, J., Salabarria, J., Umland, J., Werner, F., Krejca, M. S., Rothenberger, R., Kötzing, T., Friedrich, T., Mixed Integer Programming versus Evolutionary Computation for Optimizing a Hard Real-World Staff Assignment Problem. In: *International Conference on Automated Planning and Scheduling (ICAPS)*, pp. 541–554, 2019.

Assigning staff to engagements according to hard constraints while optimizing several objectives is a task encountered by many companies on a regular basis. Simplified versions of such assignment problems are NP-hard. Despite this, a typical approach to solving them consists of formulating them as mixed integer programming (MIP) problems and using a state-of-the-art solver to get solutions that closely approximate the optimum. In this paper, we consider a complex real-world staff assignment problem encountered by the professional service company KPMG, with the goal of finding an algorithm that solves it faster and with a better solution than a commercial MIP solver. We follow the evolutionary algorithm (EA) metaheuristic and design a search heuristic which iteratively improves a solution using domain-specific mutation operators. Furthermore, we use a flow algorithm to optimally solve a subproblem, which tremendously reduces the search space for the EA. For our real-world instance of the assignment problem, given the same total time budget of 100 hours, a parallel EA approach finds a solution that is only 1.7% away from an upper bound for the (unknown) optimum within under five hours, while the MIP solver Gurobi still has a gap of 10.5 %.

[7] Friedrich, T., Rothenberger, R., Sharpness of the Satisfiability Threshold for Non-Uniform Random k-SAT. In: *International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 6151–6155, 2019.

We study non-uniform random k-SAT on n variables with an arbitrary probability distribution p on the variable occurrences. The number $t = t(n)$ of randomly drawn clauses at which random formulas go from asymptotically almost surely (a.a.s.) satisfiable to a.a.s. unsatisfiable is called the satisfiability threshold. Such a threshold is called sharp if it approaches a step function as n increases. We show that a threshold t(n) for random k-SAT with an ensemble $(p_n)_{n\in\mathbb{N}}$ of arbitrary probability distributions on the variable occurrences is sharp if $\|p\|_2^2 = O_n(t^{-2/k})$ and $\|p\|_\infty = o_n(t^{-k/(2k-1)} \log^{-(k-1)/(2k-1)}(t))$. This result generalizes Friedgut's sharpness result from uniform to non-uniform random k-SAT and implies sharpness for thresholds of a wide range of random k-SAT models with heterogeneous probability distributions, for example such models where the variable probabilities follow a power-law distribution.

[8] Bläsius, T., Eube, J., Feldtkeller, T., Friedrich, T., Krejca, M. S., Lagodzinski, J. A. G., Rothenberger, R., Severin, J., Sommer, F., Trautmann, J., Memory-restricted Routing With Tiled Map Data. In: *Systems, Man, and Cybernetics (SMC)*, pp. 3347–3354, 2018.

Modern routing algorithms reduce query time by depending heavily on preprocessed data. The recently developed Navigation Data Standard (NDS) enforces a separation between algorithms and map data, rendering preprocessing inapplicable. Furthermore, map data is partitioned into tiles with respect to their geographic coordinates. With the limited memory found in portable devices, the number of tiles loaded becomes the major factor for run time. We study routing under these restrictions and present new algorithms as well as empirical evaluations. Our results show that, on average, the most efficient algorithm presented uses more than 20 times fewer tile loads than a normal A*.

[9] Friedrich, T., Rothenberger, R., Sharpness of the Satisfiability Threshold for Non-Uniform Random k-SAT. In: *Theory and Applications of Satisfiability Testing (SAT)*, pp. 273–291, 2018. Best Paper Award.

We study non-uniform random k-SAT on n variables with an arbitrary probability distribution p on the variable occurrences. The number $t = t(n)$ of randomly drawn clauses at which random formulas go from asymptotically almost surely (a.a.s.) satisfiable to a.a.s. unsatisfiable is called the satisfiability threshold. Such a threshold is called sharp if it approaches a step function as n increases. We show that a threshold t(n) for random k-SAT with an ensemble $(p_n)_{n\in\mathbb{N}}$ of arbitrary probability distributions on the variable occurrences is sharp if $\|p\|_2^2 = O_n(t^{-2/k})$ and $\|p\|_\infty = o_n(t^{-k/(2k-1)} \log^{-(k-1)/(2k-1)}(t))$. This result generalizes Friedgut's sharpness result from uniform to non-uniform random k-SAT and implies sharpness for thresholds of a wide range of random k-SAT models with heterogeneous probability distributions, for example such models where the variable probabilities follow a power-law distribution.

[10] Friedrich, T., Krohmer, A., Rothenberger, R., Sutton, A. M., Phase Transitions for Scale-Free SAT Formulas. In: *Conference on Artificial Intelligence (AAAI)*, pp. 3893–3899, 2017.

Recently, a number of non-uniform random satisfiability models have been proposed that are closer to practical satisfiability problems in some characteristics. In contrast to uniform random Boolean formulas, scale-free formulas have a variable occurrence distribution that follows a power law. It has been conjectured that such a distribution is a more accurate model for some industrial instances than the uniform random model. Though it seems that there is already an awareness of a threshold phenomenon in such models, there is still a complete picture lacking. In contrast to the uniform model, the critical density threshold does not lie at a single point, but instead exhibits a functional dependency on the power-law exponent. For scale-free formulas with clauses of length $k = 2$, we give a lower bound on the phase transition threshold as a function of the scaling parameter. We also perform computational studies that suggest our bound is tight and investigate the critical density for formulas with higher clause lengths. Similar to the uniform model, on formulas with $k \geq 3$, we find that the phase transition regime corresponds to a set of formulas that are difficult to solve by backtracking search.

[11] Friedrich, T., Krohmer, A., Rothenberger, R., Sauerwald, T., Sutton, A. M., Bounds on the Satisfiability Threshold for Power Law Distributed Random SAT. In: *European Symposium on Algorithms (ESA)*, pp. 37:1–37:15, 2017.

Propositional satisfiability (SAT) is one of the most fundamental problems in computer science. The worst-case hardness of SAT lies at the core of computational complexity theory. The average-case analysis of SAT has triggered the development of sophisticated rigorous and non-rigorous techniques for analyzing random structures. Despite a long line of research and substantial progress, nearly all theoretical work on random SAT assumes a uniform distribution on the variables. In contrast, real-world instances often exhibit large fluctuations in variable occurrence. This can be modeled by a scale-free distribution of the variables, which results in distributions closer to industrial SAT instances. We study random k-SAT on n variables, $m = \Theta(n)$ clauses, and a power law distribution on the variable occurrences with exponent β. We observe a satisfiability threshold at $\beta = (2k-1)/(k-1)$. This threshold is tight in the sense that instances with $\beta < (2k-1)/(k-1) - \varepsilon$ for any constant $\varepsilon > 0$ are unsatisfiable with high probability (w.h.p.). For $\beta \geq (2k-1)/(k-1) + \varepsilon$, the picture is reminiscent of the uniform case: instances are satisfiable w.h.p. for sufficiently small constant clause-variable ratios m/n; they are unsatisfiable above a ratio m/n that depends on β.

[12] Arndt, T., Hafner, D., Kellermeier, T., Krogmann, S., Razmjou, A., Krejca, M. S., Rothenberger, R., Friedrich, T., Probabilistic Routing for On-Street Parking Search. In: *European Symposium on Algorithms (ESA)*, pp. 6:1–6:13, 2016.

An estimated 30% of urban traffic is caused by search for parking spots. Traffic could be reduced by suggesting effective routes leading along potential parking spots. In this paper, we formalize parking search as a probabilistic problem on a road graph and show that it is NP-complete. We explore heuristics that optimize for the driving duration and the walking distance to the destination. Routes are constrained to reach a certain probability threshold of finding a spot. Empirically estimated probabilities of successful parking attempts are provided by TomTom on a per-street basis. We release these probabilities as a dataset of about 80,000 roads covering the Berlin area. This allows to evaluate parking search algorithms on a real road network with realistic probabilities for the first time. However, for many other areas, parking probabilities are not openly available. Because they are effortful to collect, we propose an algorithm that relies on conventional road attributes only. Our experiments show that this algorithm comes close to the baseline by a factor of 1.3 in our cost measure. This leads to the conclusion that conventional road attributes may be sufficient to compute reasonably good parking search routes.

[13] Chauhan, A., Friedrich, T., Rothenberger, R., Greed is Good for Deterministic Scale-Free Networks. In: *Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pp. 33:1–33:15, 2016.

Large real-world networks typically follow a power-law degree distribution. To study such networks, numerous random graph models have been proposed. However, real-world networks are not drawn at random. Therefore, Brach, Cygan, Lacki, and Sankowski [SODA 2016] introduced two natural deterministic conditions: (1) a power-law upper bound on the degree distribution (PLB-U) and (2) power-law neighborhoods, that is, the degree distribution of neighbors of each vertex is also upper bounded by a power law (PLB-N). They showed that many real-world networks satisfy both deterministic properties and exploit them to design faster algorithms for a number of classical graph problems. We complement the work of Brach et al. by showing that some well-studied random graph models exhibit both the mentioned PLB properties and additionally also a power-law lower bound on the degree distribution (PLB-L). All three properties hold with high probability for Chung-Lu Random Graphs and Geometric Inhomogeneous Random Graphs and almost surely for Hyperbolic Random Graphs. As a consequence, all results of Brach et al. also hold with high probability or almost surely for those random graph classes. In the second part of this work we study three classical NP-hard combinatorial optimization problems on PLB networks. It is known that on general graphs with maximum degree Δ, a greedy algorithm, which chooses nodes in the order of their degree, only achieves a $\Omega(\ln \Delta)$-approximation for Minimum Vertex Cover and Minimum Dominating Set, and a $\Omega(\Delta)$-approximation for Maximum Independent Set. We prove that the PLB-U property suffices for the greedy approach to achieve a constant-factor approximation for all three problems. We also show that all three combinatorial optimization problems are APX-complete even if all PLB-properties holds hence, PTAS cannot be expected unless P=NP.

[14] Rothenberger, R., Grau, S., Rossberg, M., Dominating an s-t-Cut in a Network. In: *Current Trends in Theory and Practice of Computer Science (SOFSEM)*, pp. 401–411, 2015.

We study an optimization problem with applications in design and analysis of resilient communication networks: given two vertices s, t in a graph $G = (V, E)$, find a vertex set $X \subset V$ of minimum cardinality, such that X and its neighborhood constitute an $s - t$ vertex separator. Although the problem naturally combines notions of graph connectivity and domination, its computational properties significantly differ from these relatives. In particular, we show that on general graphs the problem cannot be approximated to within a factor of $2^{\log^{1-\delta} n}$, with $\delta = 1/\log\log^c n$ and arbitrary $c < 1/2$ (if P $\neq$ NP). This inapproximability result even applies if the subgraph induced by a solution set has the additional constraint of being connected. Furthermore, we give a $2\sqrt{n}$-approximation algorithm and study the problem on graphs with bounded node degree. With Δ being the maximum degree of nodes $V \setminus \{s, t\}$, we identify a $(\Delta + 1)$ approximation algorithm.

[15] Bringmann, K., Friedrich, T., Hoefer, M., Rothenberger, R., Sauerwald, T., Ultra-Fast Load Balancing on Scale-Free Networks. In: *International Colloquium on Automata, Languages and Programming (ICALP)*, pp. 516–527, 2015.

The performance of large distributed systems crucially depends on efficiently balancing their load. This has motivated a large amount of theoretical research how an imbalanced load vector can be smoothed with local algorithms. For technical reasons, the vast majority of previous work focuses on regular (or almost regular) graphs including symmetric topologies such as grids and hypercubes, and ignores the fact that large networks are often highly heterogenous. We model large scale-free networks by Chung-Lu random graphs and analyze a simple local algorithm for iterative load balancing. On n-node graphs our distributed algorithm balances the load within $O((log\ log\ n)^2)$ steps. It does not need to know the exponent $2 < \beta < 3$ of the power-law degree distribution or the weights w_i of the graph model. To the best of our knowledge, this is the first result which shows that load-balancing can be done in double-logarithmic time on realistic graph classes.