When the Optimum is also Blind: a New Perspective on Universal Optimization*†

Marek Adamczyk[1], Fabrizio Grandoni[2], Stefano Leonardi[3], and Michał Włodarczyk[4]

1 **University of Bremen, Bremen, Germany**
m.adamczyk@uni-bremen.de
2 **IDSIA, USI-SUPSI, Lugano, Switzerland**
fabrizio@idsia.ch
3 **Sapienza University of Rome, Rome, Italy**
leonardi@dis.uniroma1.it
4 **University of Warsaw, Warsaw, Poland**
m.wlodarczyk@mimuw.edu.pl

Abstract

Consider the following variant of the set cover problem. We are given a universe $U = \{1, ..., n\}$ and a collection of subsets $\mathcal{C} = \{S_1, ..., S_m\}$ where $S_i \subseteq U$. For every element $u \in U$ we need to find a set $\phi(u) \in \mathcal{C}$ such that $u \in \phi(u)$. Once we construct and fix the mapping $\phi : U \mapsto \mathcal{C}$ a subset $X \subseteq U$ of the universe is revealed, and we need to cover all elements from X with exactly $\phi(X) := \bigcup_{u \in X} \phi(u)$. The goal is to find a mapping such that the cover $\phi(X)$ is as cheap as possible.

This is an example of a universal problem where the solution has to be created before the actual instance to deal with is revealed. Such problems appear naturally in some settings when we need to optimize under uncertainty and it may be actually too expensive to begin finding a good solution once the input starts being revealed. A rich body of work was devoted to investigate such problems under the regime of worst case analysis, i.e., when we measure how good the solution is by looking at the worst-case ratio: universal solution for a given instance vs optimum solution for the same instance.

As the universal solution is significantly more constrained, it is typical that such a worst-case ratio is actually quite big. One way to give a viewpoint on the problem that would be less vulnerable to such extreme worst-cases is to assume that the instance, for which we will have to create a solution, will be drawn randomly from some probability distribution. In this case one wants to minimize the expected value of the ratio: universal solution vs optimum solution. Here the bounds obtained are indeed smaller than when we compare to the worst-case ratio.

But even in this case we still compare apples to oranges as no universal solution is able to construct the optimum solution for every possible instance. What if we would compare our approximate universal solution against an optimal universal solution that obeys the same rules as we do? We show that under this viewpoint, but still in the stochastic variant, we can indeed obtain better bounds than in the expected ratio model. For example, for the set cover problem we obtain H_n approximation which matches the approximation ratio from the classic deterministic offline setup. Moreover, we show this for all possible probability distributions over U that have a polynomially large carrier, while all previous results pertained to a model in which elements were sampled independently. Our result is based on rounding a proper configuration IP that captures the optimal universal solution, and using tools from submodular optimization.

The same basic approach leads to improved approximation algorithms for other related problems, including Vertex Cover, Edge Cover, Directed Steiner Tree, Multicut, and Facility Location.

* A full version of the paper is available at http://arxiv.org/abs/1707.01702.

† The second author was partially supported by the ERC Starting Grant NEWNET 279352 and the SNSF Grant APPROXNET 200021_159697/1. The fourth author was partially supported by the National Science Centre of Poland Grant UMO-2016/21/N/ST6/00968.

44th International Colloquium on Automata, Languages, and Programming (ICALP 2017).
Editors: Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl;
Article No. 35; pp. 35:1–35:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2012 ACM Subject Classification F. Theory of Computation, G.1.6 Optimization

Keywords and phrases approximation algorithms, stochastic optimization, submodularity

Digital Object Identifier 10.4230/LIPIcs.ICALP.2017.35

1 Introduction

In a typical online problem part of the input is revealed gradually to an algorithm, which has to react to each new piece of the input by making irrevocable choices. In an online covering problem the online input consists of a sequence of *requests*, which have to be satisfied by the algorithm by buying *items* at minimum total cost.

Some online applications have severe resource constraints, typically in terms of time and/or computational power. Hence even making an online (non-trivial) choice might be too costly. In these settings it makes sense to consider universal algorithms. Roughly speaking, the goal of these algorithms is to pre-compute a reaction to each possible input, so that the online choice can then be made very quickly (say, looking at some pre-computed table). Since the adversary has a lot of power in the universal setting, typically one assumes a stochastic input. In particular, the input is sampled according to some probability distribution π, which is either given in input or that can be sampled multiple times at polynomial cost per sample (*oracle model*).

The most relevant prior work for this paper is arguably due to Grandoni et al. [21] (conference version in [20]). The authors consider the universal stochastic version of some classical NP-hard covering problems such as set cover, non-metric facility location, multicut etc. They provide polynomial-time approximation algorithms for those problems in the *independent activation model*, where each request u is independently sampled with some known probability p_u. Crucially, in their work the approximation ratio is obtained by comparing the expected cost of the approximate solution with the expected cost of the optimal offline solution (that knows the future sampled input). For example, in the set cover case they present a polynomial-time algorithm that computes a mapping of expected cost at most $O(\log(nm))\,\mathbb{E}[OPT_{\text{off}}(X)]$, where the expectation is taken over the sampling of X according to π and $OPT_{\text{off}}(X)$ is the minimum (offline) cost of a set cover of X. Here n is size of the universe and m the number of subsets. For $m \gg n$ this ratio becomes $O\left(\frac{\log m}{\log\log m}\right)$ and is tight. They also consider the universal (non-metric) facility location problem in the independent activation model, and provide a $O(\log n)$ approximation (in the above sense), where n is the total number of clients and facilities. We remark that their method seems not to lead to any improved approximation factor in the metric version of the problem. We finally mention their $O(\log^2 n)$ approximation for universal multicut in the independent activation model, where n is the number of nodes in the graph.

1.1 Our Results and Techniques

Comparing with the offline optimum as in [21] might be too pessimistic. And often when we need to optimize under uncertainty we cannot really find a better benchmark, flagship example of it would be online problems. However, stochastic two-stage [29, 42] and stochastic adaptive [38, 13, 12, 23] problems have proven that one can actually compare an approximate solution with an optimum algorithm that is not omnipotent but obeys the same rules of the model as the approximate one. This inspired us to ask the following question:

Is it also possible in a stochastic universal problem to compare our algorithm with an optimum solution that is restricted by the model in the same way as we are?

In this paper we show that we can do this indeed. In this way we manage to obtain tighter approximation ratios – which of course are compared to a weaker benchmark, but this benchmark itself can be interpreted as more fair and meaningful – and it also allows us to approach more general problems.

1.1.1 Universal Stochastic Set Cover

We shall describe carefully the Universal Stochastic Set Cover problem in this section so that we will fully present the model. For the remaining problems their full statements will appear in appropriate sections.

In the Universal Stochastic Set Cover problem we are given in input a universe $U = \{1, 2, ..., n\}$, and a collection $\mathcal{C} \subseteq 2^U$ of m subsets $S \subseteq U$, each one with an associated cost $c(S)$. We need to a priori map each element $u \in U$ into some set $\phi(u) \in \mathcal{C}$. Then a subset $X \subseteq U$ is sampled according to some probability distribution π (either given in input as a set of scenarios or only by accessing an oracle), and we have to buy the sets $\phi(X) = \bigcup_{u \in X} \phi(u)$ as the cover of X. Our goal is to minimize the expected value of the total cost, i.e., $\mathbb{E}_{X \sim \pi} [c(\phi(X))] = \mathbb{E}_{X \sim \pi} \left[\sum_{S \in \phi(X)} c(S)\right]$. One of the most important aspects in our model is that we do not compare ourselves against the expected value of an optimum offline solution for a given scenario, that is, not against $\mathbb{E}_{X \sim \pi} [OPT_{\text{off}}(X)]$, but what we compare ourselves with is

$$\min_{\phi : \forall_{u \in U} u \in \phi(u)} \mathbb{E}_{X \sim \pi} \left[\sum_{S \in \phi(X)} c(S) \right]$$

which is the expected outcome of an *optimal universal mapping*.

The results depend on the probability distribution with which we deal in an instance of the problem. Possible probability distributions are:

- **Scenario model:** Here we are given in input all the sets $X_1, ..., X_N \subseteq U$ such that $\mathbb{P}_{X \sim \pi} [X = X_i] > 0$ with the associated probability. This model allows for explicit use of all the scenarios in the computations. For the Universal Stochastic Set Cover we obtain $O(\log n)$-approximation in this case even in the weighted case.
- **Oracle model:** This is the most general model. We have a black-box access to an oracle Π from which we can sample a scenario from distribution π. We assume that taking a sample requires polynomial time. In this model we can find an $O(\log n)$-approximation for Universal Stochastic Set Cover in polynomial time only for the unweighted case; in the weighted case we achieve the same approximation factor in pseudo-polynomial time depending on $\max_{S \in \mathcal{C}} c(S)$.
- **Independent activation model:** In this model we assume that every element $u \in U$ is independently sampled with some given probability p_u. This model does not capture correlations of elements, and therefore sometimes it is not fully realistic. Though it cannot be represented by a polynomial number of scenarios, its nice properties allow one to develop good approximation algorithms for several problems. In this setting we are able to approximate Universal Stochastic Set Cover within a factor $O(\log n)$ in polynomial time even in the weighted case.

Our results are obtained by defining a proper configuration LP (with an exponential number of variables) that captures the optimal mapping. We are able to solve this LP via the ellipsoid method using a separation oracle. Somehow interestingly, our separation oracle has to solve a submodular minimization problem. Then we can round the fractional solution in a standard way.

1.1.2 Overview of the results

The robustness of our framework allows us to address universal extensions of several covering problems. After expressing the goal as a true approximation task, we can adapt tools from the rich theory of approximation algorithms.

Here we give an overview of our results. Detailed statements of the theorems appear in appropriate sections.

Scenario model

In this setting, we are able to construct an LP-based polynomial-time $O(\log n)$-approximation to the universal stochastic version of Set Cover (Theorem 6), which generalizes to Non-Metric Facility Location and Constrained Set Multicover. In fact, the latter algorithm achieves an approximation guarantee of exactly H_n. Different rounding procedure leads to a 2-approximation for Universal Stochastic Vertex Cover. All these approximation ratios match the best guarantees obtained in the deterministic world. What is more, the exact polynomial time algorithm for Edge Cover extends to the scenario model. We also present an $O(\sqrt{k})$-approximation for Universal Stochastic Directed Steiner Tree on acyclic graphs, where k is the number of terminals. Except the Set Cover problem, proof of all of these Theorems will appear in the full version of the paper.

Independent activation model

In this setting, we are able to obtain several results in flavour of the $O(1)$-approximation for the Maybecast problem by Karger and Minkoff [33]. We present a 6.33-approximation for Universal Stochastic Metric Facility Location (Theorem 8) and an $O(\log n)$-approximation for Universal Stochastic Multicut (Theorem 12). As an intermediate result, we obtain a 4.75-approximation for Universal Stochastic Multicut on trees.

Oracle model

We can generalize most of our results for the scenario model to this setting, with the restriction that in the weighted case we get a pseudo-polynomial running time. This will be discussed in the full version.

1.2 Related work

Other universal-like problems have been addressed in the literature. For instance, in the universal TSP problem one computes a permutation of the nodes that is then used to visit a given subset of nodes. This problem has been studied both in the worst-case scenario for the Euclidean plane [39, 6] and general metrics [32, 22, 27], as well as in the average-case [31, 7, 43, 19, 45]. (For the related problem of universal Steiner tree, see [33, 32, 22, 19].) Jia et al. [32] introduced the universal set cover and universal facility location problems, and studied them in the worst-case: they show that the adversary is very powerful in such models, and give nearly-matching $\Omega(\sqrt{n})$ and $O(\sqrt{n \log n})$ bounds on the competitive factor. These problems have been later studied by Grandoni et al. in the independent activation model [21], as already mentioned before.

A somewhat related topic is *oblivious routing* [40, 28, 8] (see, e.g., [47, 49] for special cases). A tight logarithmic competitive result as well as a polynomial-time algorithm to compute the best routing is known in the worst case for undirected graphs [5, 41]. For *oblivious routing on directed graphs* in the worst case the lower bound of $\Omega(\sqrt{n})$ [5] nearly matches upper bounds in [24] but for the average case. The authors of [25] give an $O(\log^2 n)$-competitive oblivious routing algorithm when demands are chosen randomly from a known demand-distribution; they also use "demand-dependent" routings and show that these are necessary.

Another closely related notion is the one of *online problems*. These problems have a long history (see, e.g., [9, 16]), and there have been many attempts to relax the strict worst-case notion of competitive analysis: see, e.g., [14, 1, 19] and the references therein. Online problems with stochastic inputs (either i.i.d. draws from some distribution, or inputs arriving in random order) have been studied, e.g., in the context of optimization problems [36, 37, 19, 4], secretary problems [18], mechanism design [26], and matching problems in Ad-auctions [35].

Alon et al. [2] gave the first online algorithm for set cover with a competitive ratio of $O(\log m \log n)$; they used an elegant primal-dual-style approach that has subsequently found many applications (e.g., [3, 10]). This ratio is the best possible under complexity-theoretic assumptions [34]; even unconditionally, no deterministic online algorithm can do much better than this [2]. Online versions of *metric* facility location are studied in both the worst case [36, 17], the average case [19], as well as in the stronger *random permutation model* [36], where the adversary chooses a set of clients unknown to the algorithm, and the clients are presented to us in a random order. It is easy to show that for our problems, the random permutation model (and hence any model where elements are drawn from an *unknown* distribution) are as hard as the worst case.

One can of course consider the (offline) stochastic version of optimization problems. For example, k-stage stochastic set cover is studied in [29, 44], with an improved approximation factor (independent from k) later given in [46].

The result for the Oracle model are based on the Sample Average Approximation approach, see [11] for the application most relevant to our work.

As mentioned before, in two-stage stochastic problems [42, 29] and stochastic adaptive problems [13, 38, 12, 23] it is possible to compare our algorithms with an optimum algorithm which is equally constrained in the model as our problem, and this is what shed a light on the possibility of obtaining results in the same spirit for the universal stochastic optimization.

In two recent papers [48, 15] authors looked at universal optimization over scenarios, but compared against the average offline optimum, and not the optimum universal solution as we do. All the properties of submodular functions used in our work can be found here [50].

2 Preliminaries

Here we give some basic definitions and properties. Given a universe U, we call a function $f : 2^U \to \mathbb{R}$ submodular if $f(A)+f(B) \geq f(A\cap B)+f(A\cup B)$ for each pair of sets $A, B \subseteq U$. Function f is monotone if $f(B) \geq f(A)$ for $A \subseteq B$. When considering a submodular function, we assume that it is implicitly given in the form of an oracle that can be queried on a specific $A \subseteq U$ and returns the value $f(A)$ in constant time.

▶ **Theorem 1** (Iwata et al. [30]). *There is an algorithm to minimize a given submodular function $f : 2^U \to \mathbb{N}$ in polynomial time in $|U|$ and in the number of bits needed to encode the largest value of f.*

Let us introduce a function $g_\pi : 2^U \to \mathbb{R}^+$, $g_\pi(A) = \mathbb{P}[A \cap X \neq \emptyset]$ where X is drawn from the distribution π. Our framework exploits crucially the fact that g_π is a submodular function.

▶ **Lemma 2.** *Function g_π is submodular and monotone.*

Proof. Observe that $g_\pi(A) = \sum_{X \subseteq U} \pi(X) \cdot \mathbf{1}[A \cap X \neq \emptyset]$. Function $A \to \mathbf{1}[A \cap X \neq \emptyset]$ is submodular and a combination of such functions with positive coefficients is submodular. The monotonicity holds trivially by definition. ◀

3 Universal Set Cover

In this section we present our approximation algorithm for universal set cover. We start by considering the case that π is given in input, in the form of a set of N scenarios $X_1, \dots, X_N$, where scenario $X_i \subseteq U$ is sampled with some given probability $p_i = \pi(X_i)$. Then we (partially) extend our approach to the oracle model. Recall that $n = |U|$ is the number of elements in the universe.

Consider the scenario case. Recall that, for $B \subseteq U$, $g_\pi(B) = \mathbb{P}[B \cap X \neq \emptyset]$, where the probability is taken over $X \subseteq U$ sampled from π (namely, one of the scenarios A_i in this case). As mentioned before, g_π is a submodular function over the universe U.

We start by expressing our problem as the following integer program (IP).

$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{C}} c(S) \sum_{B \subseteq S} y_B^S \cdot g_\pi(B) & & \text{(IP-SC)} \\ \text{s.t.} \quad & \sum_{B \ni u} \sum_{S \supseteq B} y_B^S \geq 1 & & \forall_{u \in U} \\ & y_B^S \in \{0,1\} & & \forall_{S \in \mathcal{C}} \forall_{B \subseteq S}. \qquad (1) \end{aligned}$$

We obtain a linear relaxation (LP-SC) of (IP-SC) by replacing the integrality constraints (1) with $y_B^S \geq 0$.

▶ **Lemma 3.** *Integer program (IP-SC) is equivalent to Universal Stochastic Set Cover.*

Proof. It is easy to translate a mapping $\phi : U \to \mathcal{C}$ into some feasible solution to (LP-SC). All variables are zeros by default and for each $S \in \mathcal{C}$ such that $\phi^{-1}(S)$ is non-empty, we set $y^S_{\phi^{-1}(S)} = 1$. Note that always $\phi^{-1}(S) \subseteq S$. In that setting the objective value equals the expected cost of the covering.

Let us fix some feasible solution $\{y_B^S\}_{S \in \mathcal{C}, B \subseteq S}$. We know that for each $u \in U$ there is some pair (B, S) so that $u \in B$ and $y_B^S = 1$ (we will call it a *covering pair*). As long as there are many covering pairs for some u, we replace one of them with $(B \setminus \{u\}, S)$. The new solution is still feasible and the objective value is no greater as function g_π is monotone. Therefore there exists an optimal solution so that each $u \in U$ admits exactly one covering pair (B_u, S_u). We can define $\phi(u) = S_u$ to obtain a covering with expected cost equal to the value of the objective function. ◀

(LP-SC) has an exponential number of variables: in order to solve it we consider its dual, and provide a separation oracle to solve it. Interestingly, our separation oracle uses submodular minimization.

▶ **Lemma 4.** *(LP-SC) can be solved in polynomial time.*

Proof. We show how to solve the dual of (LP-SC), which is as follows:

$$\begin{aligned} \max \quad & \sum_{u \in U} \alpha_u & & \text{(DP-SC)} \\ \text{s.t.} \quad & \sum_{u \in B} \alpha_u \le c(S) \cdot g_\pi(B) & & \forall_{S \in \mathcal{C}} \forall_{B \subseteq S} \\ & \alpha_u \ge 0 & & \forall_{u \in U}. \end{aligned}$$

Observe that (DP-SC) has a polynomial number of variables and an exponential number of constraints. In order to solve (DP-SC), it is sufficient to provide a (polynomial-time) separation oracle, i.e. a procedure that, given a tentative solution $\{\alpha_u\}_{u \in U}$, either determines that it is feasible or provides a violated constraint.

This reduces to check, for each given $S \in \mathcal{C}$, whether there exists $B \subseteq S$ such that $\sum_{u \in B} \alpha_u > c(S) \cdot g_\pi(B)$. In other terms, we wish to determine whether the minimum of function $h_S(B) := c(S) \cdot g_\pi(B) - \sum_{u \in B} \alpha_u$ is negative. Observe that the value of $h_S(B)$ can be computed in polynomial time for a given B, since π is given in input as a set of scenarios. Note also that h_S is submodular: indeed g_π is submodular, costs $c(S)$ are non-negative by assumption, and $-\sum_{u \in B} \alpha_u$ is linear (hence submodular). Hence we can minimize h_S over $B \subseteq S$ in polynomial time via Theorem 1.[1] ◀

Given the optimal solution to (LP-SC), it is sufficient to round it with the usual randomized rounding algorithm for set cover.

▶ **Lemma 5.** *The optimal solution to (LP-SC) can be rounded to an integer feasible solution while increasing the cost by a factor $O(\log n)$ in expected polynomial time.*

Proof. In the optimal solution, for each $S \in \mathcal{C}$, variables $\{y_B^S\}_{B \subseteq S}$ define a probability distribution. We sample from this distribution (independently for each S) for $q = 2 \ln n$ many times. Let $B_1^S, \ldots, B_q^S$ be the sets sampled for S: we let $B^S := \cup_i B_i^S$ and tentatively map elements of B^S into S. In case the same element u belongs to $B^{S'}$ and $B^{S''}$ for $S' \neq S''$, we replace $B^{S'}$ with $B^{S'} \setminus \{u\}$ and iterate: this way each element is mapped into exactly one set. The final sets B^S induce our approximate mapping.

We can upper bound the expected cost of the solution by $\sum_{S \in \mathcal{C}} c(S) \sum_{i=1}^q g_\pi(B_i^S)$. Indeed, by the subadditivity of g_π (which is implied by submodularity and non-negativity), one has $g_\pi(B^S) \le \sum_{i=1}^q g_\pi(B_i^S)$. Furthermore, $g_\pi(B^{S'} \setminus \{u\}) \le g_\pi(B^{S'})$ since g_π is monotone. Trivially

$$\begin{aligned} \mathbb{E}\left[\sum_{S \in \mathcal{C}} c(S) \sum_{i=1}^q g_\pi(B_i^S)\right] &= 2 \ln n \cdot \mathbb{E}\left[\sum_{S \in \mathcal{C}} c(S) g_\pi(B_1^S)\right] \\ &= 2 \ln n \cdot \sum_{S \in \mathcal{C}} c(S) \sum_{B \subseteq S} y_B^S \cdot g_\pi(B^S) \le 2 \ln n \cdot OPT. \end{aligned}$$

And from Markov's inequality

$$\mathbb{P}\left[\sum_{S \in \mathcal{C}} c(S) \sum_{i=1}^q g_\pi(B_i^S) > 4 \ln n \cdot OPT\right] < \frac{1}{2}.$$

---

[1] In order to solve the configuration LP, there is an alternative to finding a separation oracle for the dual. We can transform the configuration LP into a optimization program where we need to minimize a sum of Lovasz's extensions [50], which are convex functions, over a convex region. This approach would be possibly more efficient, but we have chosen the one above for a simpler presentation.

The probability that an element $u \in U$ is not covered with a single sampling over y_B^S is

$$\prod_{B \ni u} \prod_{S \supseteq B} (1 - y_B^S) \leq \prod_{B \ni u} \prod_{S \supseteq B} e^{-y_B^S} = e^{-\sum_{B \ni u} \sum_{S \supseteq B} y_B^S} \leq \frac{1}{e}.$$

Therefore, by the independence of the sampling and the union bound, the probability that at least one element is not covered is at most $n \cdot \frac{1}{e^{2 \ln n}} = \frac{1}{n}$.

Altogether this gives a Monte-Carlo algorithm. As usual, this can be turned into a Las-Vegas algorithm with expected polynomial running time by repeating the procedure when some element is not covered or the cost of the solution is greater than $4 \ln n \cdot OPT$. ◀

The following theorem is a straight-forward consequence of Lemmas 4 and 5.

▶ **Theorem 6.** Universal Stochastic Set Cover *in the scenario model admits a polynomial-time* $O(\log n)$ *approximation algorithm w.r.t. the optimal universal mapping.*

It turns out that not only the randomized rounding technique can be adapted to the universal stochastic model but also the dual fitting technique. This allows us to improve the above result to a purely deterministic algorithm with the approximation ratio exactly equal to H_n. What is more, we are able to consider more general problems where each element u is required to be covered at least $r(u)$ times. Details will appear in the full version.

Our results can be also partially generalized to the oracle model, where the probability distribution is given only by a blackbox oracle. In the unweighted case we are able to give a polynomial time algorithm, while in the weighted case a pseudo-polynomial time one where the complexity depends also on $\max_S c(S)$. Again, details will appear in the full version.

Finally, since in the independent activation model we can compute function g_π in polynomial time, our approach works also in that case.

▶ **Theorem 7.** Universal Stochastic Set Cover *in the independent activation model admits a polynomial-time* $O(\log n)$ *approximation algorithm w.r.t. the optimal universal mapping.*

4 Metric facility location in the independent activation model

In the stochastic universal variant of the metric uncapacitated facility location (*UniStoch-FL*) problem, we have a set of clients C and a set of facilities F. For each client $c \in C$ and facility $f \in F$, there is a cost $d(c, f)$ paid if c is connected to f; furthermore, there is a cost o_f associated with opening facility $f \in F$. In the universal solution we need to assign every client $c \in C$ to a facility $\phi(c) \in F$. Once a set $X \subseteq U$ is realized we need to open all facilities $f \in \bigcup_{c \in X} \phi(c)$ and connect each $c \in X$ to $\phi(c)$. The goal is to minimize the expected total cost of opening the facilities and connecting clients to facilities:

$$\mathbb{E}_{X \sim \pi} \left[\sum_{f \in \bigcup_{c \in X} \phi(c)} o_f + \sum_{c \in X} d(c, \phi(c)) \right].$$

Just by direct modeling of the above formula with the configuration LP we can see that the following is a relaxation of an integer program that solves the problem:

$$\min \sum_{f \in F} o_f \sum_{B \subseteq C} y_B^f \cdot g_\pi(B) + \sum_{c \in C} \sum_{f \in F} g_\pi(c) \cdot d(c, f) \cdot \sum_{B \subseteq C : c \in B} y_B^f \qquad \text{(CONF-LP-FL)}$$

$$\text{s.t. } \forall_{c \in C} : \sum_{f \in F} \sum_{B \subseteq C : c \in B} y_B^f \geq 1 \qquad \text{and} \qquad \forall_{f \in F} \forall_{B \subseteq C} : y_B^f \geq 0.$$

Let $OPT_{CONF-LP-FL}$ be the optimal solution to the above LP. It holds that $OPT_{CONF-LP-FL}$ is a lower-bound on the expected outcome of an optimal universal solution.

What we are able to present in this section is a constant approximation in the case of independent activation where every client c is independently chosen into the scenario with probability p_c. This restriction is due to the fact that we do not know of a good rounding procedure in the general case, while we can solve the configuration LP even with the arbitrary scenario distribution.

▶ **Theorem 8.** *For* Universal Stochastic Facility Location *in the independent activation model there exists an algorithm with expected outcome cost at most* $6.33 \cdot OPT_{CONF\text{-}LP\text{-}FL}$.

4.1 Transforming the LP

We start off with the following two inequalities for any $S \subseteq C$ (see [33] for proof):

$$\min\left(1, \sum_{t\in S} p_t\right) \geq g_\pi(S) = 1 - \prod_{t\in S}(1-p_t) \geq \left(1-\frac{1}{e}\right)\min\left(1, \sum_{t\in S} p_t\right). \tag{2}$$

Now based on this inequality we shall transform and relax the (CONF-LP-FL) to get one that is a bit easier to tackle. The inequality implies that the solution of the following LP is at most $\frac{e}{e-1}$ bigger than the solution of (CONF-LP-FL):

$$\begin{aligned}
&\min \sum_{f\in F} o_f \sum_{B\subseteq C} y_B^f \cdot \min\left(1, \sum_{c\in B} p_c\right) + \sum_{c\in C}\sum_{f\in F} g_\pi(c)\cdot d(c,f)\cdot \sum_{B\subseteq C: c\in B} y_B^f && (3)\\
&\text{s.t.} \sum_{f\in F}\sum_{B\subseteq C: c\in B} y_B^f \geq 1 && \forall_{c\in C}\\
&\quad y_B^f \geq 0 && \forall_{f\in F}\forall_{B\subseteq C}.
\end{aligned}$$

Let Big be a collection of sets $B \subseteq C$ such that $\sum_{t\in B} p_t > 1$, and let Sml be a collection of sets $B \subseteq C$ such that $\sum_{t\in B} p_t \leq 1$. Define x_c^f to be the extent to which c was assigned to f via sets from Big, and $\bar{x}_c^f$ the extent to which c was assigned to f via sets from Sml, i.e., $x_c^f = \sum_{B\in Big: c\in B} y_B^f$ and $\bar{x}_c^f = \sum_{B\in Sml: c\in B} y_B^f$. Also, for every client c there is an obvious inequality $\sum_{B\in Big} y_B^f \geq \sum_{B\in Big: c\in B} y_B^f = x_c^f$. Now we can lower bound (3):

$$\begin{aligned}
&\sum_{f\in F} o_f \left(\sum_{B\in Big} y_B^f + \sum_{B\in Sml} y_B^f \cdot \left(\sum_{j\in B} p_j\right)\right) + \sum_{c\in C}\sum_{f\in F} p_c\cdot d(c,f)\cdot \sum_{B\subseteq C: c\in B} y_B^f\\
&= \sum_{f\in F} o_f \left(\sum_{B\in Big} y_B^f\right) + \sum_{f\in F}\sum_{c\in C} o_f\cdot p_c \left(\sum_{B\in Sml: c\in B} y_B^f\right) + \sum_{c\in C}\sum_{f\in F} p_c\cdot d(c,f)\cdot \sum_{B\subseteq C: c\in B} y_B^f\\
&\geq \sum_{f\in F} o_f\cdot \max_{c\in C}\left(x_c^f\right) + \sum_{f\in F}\sum_{c\in C} o_f\cdot p_c\cdot \bar{x}_c^f + \sum_{c\in C}\sum_{f\in F} p_c\cdot d(c,f)\cdot\left(x_c^f + \bar{x}_c^f\right).
\end{aligned}$$

▶ **Lemma 9.** *Let OPT_{LP-FL} be the optimum solution of the following LP:*

$$\begin{aligned}
&\min \sum_{f\in F} o_f\cdot \max_{c\in C}\left(x_c^f\right) + \sum_{f\in F} o_f\cdot \sum_{c\in C} p_c\cdot \bar{x}_c^f + \sum_{c\in C}\sum_{f\in F} p_c\cdot d(c,f)\cdot\left(x_c^f + \bar{x}_c^f\right) \quad \text{(LP-FL)}\\
&\text{s.t. } \forall_{c\in C}: \sum_{f\in F} x_c^f + \bar{x}_c^f \geq 1 \qquad \text{and} \qquad \forall_{f\in F}\forall_{c\in C}: x_c^f, \bar{x}_c^f \geq 0.
\end{aligned}$$

It holds that $OPT_{LP-FL} \leq \frac{e}{e-1}\cdot OPT_{CONF-LP-FL}$.

4.2 The rounding procedure

LP-FL has a polynomial number of variables and constraints, and so it can be solved in polynomial time; denote by $\left(x_c^f, \bar{x}_c^f\right)_{c\in C, f\in F}$ the optimal solution. Once we solve it we split the clients into two groups:

$$C_{\text{big}} := \left\{ c \in C \,\middle|\, \sum_{f\in F} x_c^f \geq \frac{3}{4} \right\} \text{ and } C_{\text{sml}} = \left\{ c \in C \,\middle|\, \sum_{f\in F} \bar{x}_c^f > \frac{1}{4} \right\}.$$

Dealing with clients from C_{big}

From the definition we get that $\left(\frac{4}{3}x_c^f\right)_{c\in C_{\text{big}}, f\in F}$ is a feasible solution to the following problem, which is a variant of the deterministic facility location problem where the price for the distance for client $c \in C$ is proportional to the factor p_c:

$$\begin{aligned} \min \quad & \sum_{f\in F} o_f \cdot \max_{c\in C_{\text{big}}} \left(z_c^f\right) + \sum_{c\in C}\sum_{f\in F} p_c \cdot d(c,f) \cdot z_c^f \\ \text{s.t.} \quad & \forall_{c\in C_{\text{big}}} : \sum_{f\in F} z_c^f \geq 1 \qquad \text{and} \qquad \forall_{f\in F}\forall_{c\in C_{\text{big}}} : z_c^f \geq 0. \end{aligned}$$

For this problem there exists a 3-approximation algorithm based on primal-dual method (the proof is a simple exercise and will appear in the full version of the paper); let $\left(z_c^f\right)_{c\in C, f\in F}$ be such an integral solution. Thus we have

$$\begin{aligned} & \sum_{f\in F} o_f \cdot \max_{c\in C_{\text{big}}} \left(z_c^f\right) + \sum_{c\in C_{\text{big}}}\sum_{f\in F} p_c \cdot d(c,f) \cdot z_c^f \\ \leq\ & 3 \cdot \sum_{f\in F} o_f \cdot \max_{c\in C_{\text{big}}} \left(\frac{4}{3}x_c^f\right) + 3 \cdot \sum_{c\in C_{\text{big}}}\sum_{f\in F} p_c \cdot d(c,f) \cdot \frac{4}{3}x_c^f \\ \leq\ & 4 \cdot \sum_{f\in F} o_f \cdot \max_{c\in C} \left(x_c^f\right) + 4 \cdot \sum_{c\in C}\sum_{f\in F} p_c \cdot d(c,f) \cdot x_c^f. \end{aligned} \tag{4}$$

Dealing with clients from C_{sml}

Now the $\left(4 \cdot \bar{x}_c^f\right)_{c\in C_{\text{sml}}, f\in F}$ vector is a feasible solution to the following LP:

$$\begin{aligned} \min \quad & \sum_{c\in C_{\text{sml}}}\sum_{f\in F} p_c \left(o_f + d(c,f)\right) \cdot z_c^f \\ \text{s.t.} \quad & \forall_{c\in C_{\text{sml}}} : \sum_{f\in F} z_c^f \geq 1 \qquad \text{and} \qquad \forall_{f\in F}\forall_{c\in C_{\text{sml}}} : z_c^f \geq 0. \end{aligned}$$

We can find an integral solution to this LP easily: just assign every client $c \in C_{\text{sml}}$ to the facility f that minimizes $o_f + d(c,f)$. If $\left(z_c^f\right)_{c\in C_{\text{sml}}, f\in F}$ is such an integral solution, then we have that

$$\sum_{c\in C_{\text{sml}}}\sum_{f\in F} p_c \left(o_f + d(c,f)\right) \cdot z_c^f \leq \sum_{c\in C_{\text{sml}}}\sum_{f\in F} p_c \left(o_f + d(c,f)\right) \cdot 4\bar{x}_c^f. \tag{5}$$

Combining the two solutions

In this way the solution composed from two integer vectors $\left(z_c^f\right)_{c\in C_{\text{big}}, f\in F}$ and $\left(z_c^f\right)_{c\in C_{\text{sml}}, f\in F}$ is a feasible solution to the (LP-FL), and by combining (4) and (5) its value is

$$\begin{aligned}
&\sum_{f\in F} o_f \cdot \max_{c\in C_{\text{big}}} \left(z_c^f\right) + \sum_{f\in F}\sum_{c\in C_{\text{sml}}} o_f \cdot p_c \cdot z_c^f + \sum_{c\in C}\sum_{f\in F} p_c \cdot d\left(c,f\right) \cdot z_c^f \\
\leq\ & 4\cdot \sum_{f\in F} o_f \cdot \max_{c\in C} \left(x_c^f\right) + \sum_{f\in F}\sum_{c\in C} o_f \cdot p_c \cdot 4\bar{x}_c^f + \sum_{c\in C}\sum_{f\in F} p_c \cdot d\left(c,f\right) \cdot \left(4x_c^f + 4\bar{x}_c^f\right) \\
\leq\ & 4\cdot \frac{e}{e-1} OPT_{CONF-LP-FL} < 6.33 \cdot OPT_{CONF-LP-FL}.
\end{aligned}$$

Thus we found a solution $\left(z_c^f\right)_{c\in C, f\in F}$ for the (LP-FL) whose value is at most $6.33 \cdot OPT_{CONF-LP-FL}$. However using the left-side inequality from (2) we can transform it into a solution of the (CONF-LP-FL) without losing its value. Hence we found a solution to the Universal Stochastic Facility Location which is a 6.33-approximation.

5 Multicut in the independent activation model

In the (classical version of the) Multicut problem we are given an undirected graph G with non-negative costs c_e for all $e \in E(G)$ and a set of pairs of vertices $(s_1, t_1), \ldots, (s_k, t_k)$. The goal is to erase a subset of edges F of small cost so that there is no path connecting s_c with t_c for any c. Multicut may be considered a covering problem in which the client c is covered if F contains some (s_c, t_c)-cut.

In the universal stochastic setting we are also given a probability distribution π over the subsets of $C = [1, k]$ and the solution is a mapping $\phi : C \to 2^E$ such that $\phi(c)$ forms a (s_c, t_c)-cut. The expected cost of the solution induced by a mapping ϕ equals $\mathbb{E}_{X\sim\pi}\left[\sum_{\substack{e\in\phi(c)\\ c\in X}} c_e\right]$.

We express the problem with a configuration integer program. Let $\mathcal{P}_c$ denote the family of all paths connecting s_c with t_c and let $y_B^e = 1$ mean that $B = \{c \in C : e \in \phi(c)\}$.

$$\begin{aligned}
\min & \sum_{e\in E} c_e \sum_{B\subseteq C} y_B^e \cdot g_\pi\left(B\right) && \text{(CONF-IP-MC)} \qquad (6) \\
\text{s.t.} & \sum_{e\in P}\sum_{B\ni c} y_B^e \geq 1 && \forall_{c\in C}\,\forall_{P\in\mathcal{P}_c}. \\
& y_B^e \in \{0,1\} && \forall_{e\in E}\,\forall_{B\subseteq C}.
\end{aligned}$$

We next use CONF-LP-MC to denote the linear relaxation of CONF-IP-MC. Likewise for Facility location, we will show that in the independent activation model we can reduce the universal stochastic setting to the buy-at-bulk setting, where each edge e can be either bought for price c_e to serve all the clients or be rented by client c for price $c_e \cdot p_c$. We define variables x^e to indicate that e has been bought and variables $\bar{x}_c^e$ to express the event of e being rented by c. This transition simplifies the linear program CONF-LP-MC to the following one by sacrificing an approximation factor of $\frac{e}{e-1}$.

$$\begin{aligned}
\min & \sum_{e\in E} c_e \cdot x^e + \sum_{e\in E} c_e \sum_{c\in C} p_c \cdot \bar{x}_c^e && \text{(LP-MC)} \\
\text{s.t.} & \sum_{e\in P} (x^e + \bar{x}_c^e) \geq 1 && \forall_{c\in C}\forall_{P\in\mathcal{P}_c} \\
& x^e, \bar{x}_c^e \geq 0 && \forall_{e\in E}\forall_{c\in C}.
\end{aligned}$$

The proofs of the two following lemmas are similar to derivations in Sections 4.1 and 4.2 and so we omit them in this extended abstract.

▶ **Lemma 10.** *If π is an independent activation distribution, then the optimal value of LP-MP is at most $\frac{e}{e-1}$ times larger than the optimal value of CONF-LP-MP.*

▶ **Lemma 11.** *If G is a tree, then one can round a fractional solution to LP-MC to an integral one of cost at most 3 times larger. The procedure runs in polynomial time.*

In order to solve the problem on general graphs, we will embed the graph into a tree that approximately preserves the structure of cuts. The following construction has been introduced by Räcke [41]. We call a tree T decomposition tree of G if

1. the leaves of T correspond to vertices of G,
2. each edge e_t in T has weight $c^T_{e_t}$ equal to the weight of the cut it induces on $V(G)$ (we call this cut $m_T(e_t)$).

For an edge $e \in E(G)$ and a decomposition tree T we define the *relative load* of e as

$$rload_T(e) = \Bigg(\sum_{\substack{e_t \in E(T) \\ e \in m_T(e_t)}} c^T_{e_t} \Bigg) \,/\, c_e.$$

The main result in [41] concerns the relation between multicommodity flows in G and T_i. As our LP formulation is slightly more sophisticated we need to exploit this result in more detail. Section 2.1 in [41] describes how to find (in polynomial time) a convex combination of decomposition trees $\{\lambda_i T_i\}_{i=1}^q$ for a graph G, such that $\max_{e\in E(G)} [\sum_{i=1}^q \lambda_i rload_{T_i}(e)] = O(\log n)$.

▶ **Theorem 12.** Universal Stochastic Multicut *in the independent activation model admits a polynomial-time $O(\log n)$ approximation algorithm.*

Proof. Lemma 11 implies that a fractional solution to LP-MC over T_i can be rounded to an integer solution with ratio 3. Observe that each tree edge on a path between terminals corresponds to some cut between these terminals so any solution to Universal Stochastic Multicut on a decomposition tree induces a solution on the original graph of at most the same cost.

Let L_i denote the optimal value of LP-MC over T_i and L denote the same for G. We are going to show that $\sum_{i=1}^q \lambda_i L_i = O(L \log n)$. In particular this means that $\min_{i=1}^q L_i = O(L \log n)$. After rounding the fractional solution on T_i of the smallest value, we will obtain an integral solution for G of cost $O(L \log n)$, what entails an $O(\log n)$ approximation.

Let us consider the dual linear program of LP-MC.

$$\begin{aligned}
\max \sum_{c\in C} \sum_{P\in\mathcal{P}_c} \alpha_P & & \text{(DP-MC)} \\
\text{s.t.} \sum_{c\in C} \sum_{\substack{P\in\mathcal{P}_c \\ e\in P}} \alpha_P \le c_e & \qquad \forall_{e\in E} & (7) \\
\sum_{\substack{P\in\mathcal{P}_c \\ e\in P}} \alpha_P \le c_e \cdot p_c & \qquad \forall_{c\in C}\, \forall_{e\in E} & (8) \\
\alpha_P \ge 0 & \qquad \forall_{c\in C}\, \forall_{P\in\mathcal{P}_c}. &
\end{aligned}$$

Feasible solutions to (DP-MC) are just multicommodity flows satisfying conditions (7)-(8). Let (β^{T_i}) be an optimal solution to (DP-MC) on a decomposition tree T_i (of value L_i) and let

(α^{T_i}) be a flow on G where a unit flow over e_t in (β^{T_i}) translates into a unit flow over $m_T(e_t)$. Note that the value of shipped commodities remains the same. Consider $(\alpha) = \sum_{i=1}^{q} \lambda_i(\alpha^{T_i})$. It is a (not necessarily feasible) solution of value $\sum_{i=1}^{q} \lambda_i L_i$.

As (β^{T_i}) routes at most $c_{e_t}^{T_i}$ flow through an edge e_t, then (α^{T_i}) routes at most $\sum_{\substack{e_t \in E(T_i) \\ e \in m_{T_i}(e_t)}} c_{e_t}^{T_i}$ flow through an edge e. Therefore constraint (7) is exceeded at most $rload_{T_i}(e)$ times in (α^{T_i}) and $\sum_{i=1}^{q} \lambda_i rload_{T_i}(e)$ times in (α). If we consider vectors $(\beta_c^{T_i})$ given by flow routed between terminals of client c, then we conclude the same for constraint (8).

After scaling (α) down times $M = \max_{e \in E(G)} [\sum_{i=1}^{q} \lambda_i rload_{T_i}(e)]$ we obtain a feasible solution to (DP-MC) for G. This means that L is no less than $\sum_{i=1}^{q} \lambda_i L_i$ divided by M. As we know that $M = O(\log n)$, the claim follows. ◀

References

1 Susanne Albers and Stefano Leonardi. On-line algorithms. *ACM Comput. Surv.*, 31(3es):4, 1999.

2 N. Alon, B. Awerbuch, Y. Azar, N. Buchbinder, and J. Naor. The Online Set Cover Problem. In *STOC'03*, pages 100–105, 2003.

3 Noga Alon, Baruch Awerbuch, Yossi Azar, Niv Buchbinder, and Joseph (Seffi) Naor. A general approach to online network optimization problems. In *SODA'04*, pages 577–586, 2004.

4 Aris Anagnostopoulos, Fabrizio Grandoni, Stefano Leonardi, and Piotr Sankowski. Online network design with outliers. *Algorithmica*, 76(1):88–109, 2016.

5 Yossi Azar, Edith Cohen, Amos Fiat, Haim Kaplan, and Harald Racke. Optimal oblivious routing in polynomial time. In *STOC'03*, pages 383–388, 2003.

6 Dimitris Bertsimas and Michelangelo Grigni. Worst-case examples for the spacefilling curve heuristic for the Euclidean traveling salesman problem. *Oper. Res. Lett.*, 8(5):241–244, 1989.

7 Dimitris J. Bertsimas, Patrick Jaillet, and Amedeo R. Odoni. A priori optimization. *Oper. Res.*, 38(6):1019–1033, 1990.

8 Marcin Bienkowski, Miroslaw Korzeniowski, and Harald Räcke. A practical algorithm for constructing oblivious routing schemes. In *SPAA'03*, pages 24–33, 2003.

9 Allan Borodin and Ran El-Yaniv. *Online computation and competitive analysis*. Cambridge University Press, New York, 1998.

10 Niv Buchbinder and Joseph Naor. Online primal-dual algorithms for covering and packing problems. In *ESA'05*, pages 689–701, 2005.

11 Moses Charikar, Chandra Chekuri, and Martin Pál. Sampling bounds for stochastic optimization. In *Approximation, Randomization and Combinatorial Optimization, Algorithms and Techniques, 8th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2005 and 9th InternationalWorkshop on Randomization and Computation, RANDOM 2005, Berkeley, CA, USA, August 22-24, 2005, Proceedings*, pages 257–269, 2005. doi:10.1007/11538462_22.

12 Ning Chen, Nicole Immorlica, Anna R. Karlin, Mohammad Mahdian, and Atri Rudra. Approximating matches made in heaven. In *Automata, Languages and Programming, 36th International Colloquium, ICALP 2009, Rhodes, Greece, July 5-12, 2009, Proceedings, Part I*, pages 266–278, 2009. doi:10.1007/978-3-642-02927-1_23.

13 Brian C. Dean, Michel X. Goemans, and Jan Vondrák. Approximating the stochastic knapsack problem: The benefit of adaptivity. *Math. Oper. Res.*, 33(4):945–964, 2008. doi:10.1287/moor.1080.0330.

14 Reza Dorrigiv and Alejandro Lopez-Ortiz. A survey of performance measures for on-line algorithms. *SIGACT News*, 36(3):67–81, 2005.

15 Esteban Feuerstein, Alberto Marchetti-Spaccamela, Frans Schalekamp, René Sitters, Suzanne van der Ster, Leen Stougie, and Anke van Zuylen. Scheduling over scenarios on two machines. In *Computing and Combinatorics – 20th International Conference, COCOON 2014, Atlanta, GA, USA, August 4-6, 2014. Proceedings*, pages 559–571, 2014. doi:10.1007/978-3-319-08783-2_48.

16 Amos Fiat and Gerhard J. Woeginger, editors. *Online algorithms*, volume 1442 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1998.

17 Dimitris Fotakis. On the competitive ratio for online facility location. In *ICALP'03*, pages 637–652. Springer, 2003.

18 P. R. Freeman. The secretary problem and its extensions: a review. *Internat. Statist. Rev.*, 51(2):189–206, 1983.

19 Naveen Garg, Anupam Gupta, Stefano Leonardi, and Piotr Sankowski. Stochastic analyses for online combinatorial optimization problems. In *SODA'08*, pages 942–951, 2008.

20 Fabrizio Grandoni, Anupam Gupta, Stefano Leonardi, Pauli Miettinen, Piotr Sankowski, and Mohit Singh. Set covering with our eyes closed. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 347–356, 2008. doi:10.1109/FOCS.2008.31.

21 Fabrizio Grandoni, Anupam Gupta, Stefano Leonardi, Pauli Miettinen, Piotr Sankowski, and Mohit Singh. Set covering with our eyes closed. *SIAM J. Comput.*, 42(3):808–830, 2013.

22 Anupam Gupta, Mohammad Taghi Hajiaghayi, and Harald Räcke. Oblivious network design. In *SODA'06*, pages 970–979, 2006.

23 Anupam Gupta and Viswanath Nagarajan. A stochastic probing problem with applications. In *Integer Programming and Combinatorial Optimization – 16th International Conference, IPCO 2013, Valparaíso, Chile, March 18-20, 2013. Proceedings*, pages 205–216, 2013. doi:10.1007/978-3-642-36694-9_18.

24 Mohammad T. Hajiaghayi, Robert D. Kleinberg, Tom Leighton, and Harald Räcke. Oblivious routing on node-capacitated and directed graphs. In *SODA'05*, pages 782–790, 2005.

25 Mohammad Taghi Hajiaghayi, Jeong Han Kim, Tom Leighton, and Harald Räcke. Oblivious routing in directed graphs with random demands. In *STOC'05*, pages 193–201, 2005.

26 Mohammad Taghi Hajiaghayi, Robert Kleinberg, and David C. Parkes. Adaptive limited-supply online auctions. In *EC'04*, pages 71–80, 2004.

27 Mohammad Taghi Hajiaghayi, Robert D. Kleinberg, and Frank Thomson Leighton. Improved lower and upper bounds for universal tsp in planar metrics. In *SODA'06*, pages 649–658, 2006.

28 Chris Harrelson, Kirsten Hildrum, and Satish Rao. A polynomial-time tree decomposition to minimize congestion. In *SPAA'03*, pages 34–43, 2003.

29 Nicole Immorlica, David Karger, Maria Minkoff, and Vahab Mirrokni. On the costs and benefits of procrastination: Approximation algorithms for stochastic combinatorial optimization problems. In *SODA'04*, pages 684–693, 2004.

30 Satoru Iwata, Lisa Fleischer, and Satoru Fujishige. A combinatorial strongly polynomial algorithm for minimizing submodular functions. *Journal of the ACM (JACM)*, 48(4):761–777, 2001.

31 Patrick Jaillet. A priori solution of a travelling salesman problem in which a random subset of the customers are visited. *Oper. Res.*, 36(6):929–936, 1988.

32 Lujun Jia, Guolong Lin, Guevara Noubir, Rajmohan Rajaraman, and Ravi Sundaram. Universal approximations for tsp, steiner tree, and set cover. In *STOC'05*, pages 386–395, 2005.

33 David R. Karger and Maria Minkoff. Building Steiner trees with incomplete global knowledge. In *FOCS'00*, pages 613–623, 2000.

34 Simon Korman. On the use of randomization in the online set cover problem, 2004. MSc Thesis.

35 Aranyak Mehta, Amin Saberi, Umesh Vazirani, and Vijay Vazirani. AdWords and generalized online matching. *J. ACM*, 54(5):Art. 22, 19 pp., 2007.

36 Adam Meyerson. Online facility location. In *FOCS'01*, pages 426–431. IEEE Computer Society, 2001.

37 Adam Meyerson, Kamesh Munagala, and Serge Plotkin. Designing networks incrementally. In *FOCS'01*, pages 406–415, 2001.

38 Rolf H. Möhring, Andreas S. Schulz, and Marc Uetz. Approximation in stochastic scheduling: the power of lp-based priority policies. *J. ACM*, 46(6):924–942, 1999. doi:10.1145/331524.331530.

39 Loren K. Platzman and John J. Bartholdi, III. Spacefilling curves and the planar travelling salesman problem. *J. ACM*, 36(4):719–737, 1989.

40 Harald Räcke. Minimizing congestion in general networks. In *FOCS'02*, pages 43–52, 2002.

41 Harald Räcke. Optimal Hierarchical Decompositions for Congestion Minimization in Networks. In *STOC'08*, 2008.

42 R. Ravi and Amitabh Sinha. Hedging uncertainty: Approximation algorithms for stochastic optimization problems. In *IPCO'04*, pages 101–115, 2004.

43 Frans Schalekamp and David B. Shmoys. Algorithms for the universal and a priori tsp. *Oper. Res. Lett.*, 36(1):1–3, Jan 2008.

44 David B. Shmoys and Chaitanya Swamy. An approximation scheme for stochastic linear programming and its application to stochastic integer programs. *J. ACM*, 53(6):978–1012, 2006.

45 David B. Shmoys and Kunal Talwar. A constant approximation algorithm for the a priori traveling salesman problem. In *IPCO'08*, 2008.

46 Aravind Srinivasan. Approximation algorithms for stochastic and risk-averse optimization. In *SODA'07*, pages 1305–1313, 2007.

47 L. G. Valiant and G. J. Brebner. Universal schemes for parallel communication. In *STOC'81*, pages 263–277, 1981.

48 Martijn van Ee, Leo van Iersel, Teun Janssen, and René Sitters. A priori TSP in the scenario model. In *Approximation and Online Algorithms – 14th International Workshop, WAOA 2016, Aarhus, Denmark, August 25-26, 2016, Revised Selected Papers*, pages 183–196, 2016. doi:10.1007/978-3-319-51741-4_15.

49 Berthold Vöcking. Almost optimal permutation routing on hypercubes. In *STOC'01*, pages 530–539, 2001.

50 Jan Vondrák. *Submodularity in combinatorial optimization*. PhD thesis, Charles University, 2007.