PASS Approximation:
A Framework for Analyzing and Designing Heuristics

Uriel Feige[1⋆], Nicole Immorlica[2⋆], Vahab S. Mirrokni[3⋆], and Hamid Nazerzadeh[4⋆]

[1] Weizmann Institute, Rehovot, Israel. uriel.feige@weizmann.ac.il
[2] Northwestern University, Chicago, IL. nickle@eecs.northwestern.edu
[3] Google Research, New York, NY. mirrokni@theory.csail.mit.edu
[4] Stanford University, Stanford, CA. hamidnz@stanford.edu

Abstract. We introduce a new framework for designing and analyzing algorithms. Our framework applies best to problems that are inapproximable according to the standard worst-case analysis. We circumvent such negative results by designing guarantees for classes of instances, parameterized according to properties of the optimal solution. We also make sure that our parameterized approximation, called *PArametrized by the Signature of the Solution (PASS)* approximation, is the best possible. We show how to apply our framework to problems with additive and submodular objective functions such as the capacitated maximum facility location problems. We consider two types of algorithms for these problems. For greedy algorithms, our framework provides a justification for preferring a certain natural greedy rule over some alternative greedy rules that have been used in similar contexts. For LP-based algorithms, we show that the natural LP relaxation for these problems is not optimal in our framework. We design a new LP relaxation and show that this LP relaxation coupled with a new randomized rounding technique is optimal in our framework.

In passing, we note that our results strictly improve over previous results of Kleinberg, Papadimitriou and Raghavan [JACM 2004] concerning the approximation ratio of the greedy algorithm.

1 Introduction

Many important optimization problems in practice are inapproximable in theory. Practitioners deal with inapproximability issues by designing heuristics that, while provably bad on some instances, appear to perform well in practice. But for theoreticians, designing a formal framework to help guide algorithmic development for inapproximable problems has proved largely elusive.

In this paper, we present a new framework, called *PArametrized by the Signature of the Solution (PASS) approximations*. Our framework attempts to categorize instances according to how "easy" or "hard" they are, and design guarantees for all instances simultaneously with a single algorithm (the offered guarantee

⋆ Work performed in part at Microsoft Research.

depends on the class of the instance and will degrade to arbitrarily bad factors for inapproximable problems, but in a controlled way). We show how this framework can be applied to a general class of optimization problem, including *capacitated maximum facility location*, that can be described as maximizing a non-decreasing submodular revenue function minus a linear cost function. We then show how the new framework affects the choice of algorithms. Two standard approaches for handling such problems are via greedy and LP-based algorithms. We study a natural greedy algorithm and prove that it is an optimal PASS approximation whereas other greedy algorithms that give optimal worst-case approximations are not. For LP-based algorithms, we show that a natural LP relaxation cannot be used to design an optimal PASS approximation. Instead, we provide a different LP relaxation and an associated rounding technique that is optimal. Our new LP relaxation is unconventional in the sense that instead of providing an upper bound on the optimal solution (this is a maximization problem), it provides a lower bound.

The current paper outlines the theory of PASS approximations. We describe the general technique and how to apply this technique to a wide range of theoretical problems using both greedy and LP-based algorithms. In a companion paper [3], we apply the notion of approximation developed here to a specific problem (banner advertising) of practical significance. This problem is a special case of the broad class of problems studied in this paper.

The rest of the paper is organized as follows. After defining the problems, in Section 2, we describe the theory of PASS approximation, and compare it with previous approaches proposed to deal with the hardness of approximation. The summary of results is given in Section 3. The greedy and an LP-based algorithms for these problems are presented respectively in Sections 4 and 5.

Problem Studied

In this paper, we mainly focus on the maximum facility location problem [1, 2]:

Maximum Facility Location (MFL). *A set $\mathcal{F}$ of m facilities is given. For every facility i, there is an opening cost of c_i. There is also a set $\mathcal{J}$ of n clients. The revenue of connecting client j to facility i is $u_{ij} \geq 0$ (this may be interpreted as a client revenue minus a connection cost). Every client can connect to at most one open facility (or none). The goal in MFL is to open some facilities and connect clients to them so as to maximize the total revenue from the connected clients minus the total cost of the opened facilities.*

For comparison with some previous work [6], we shall discuss also the following problem that [6] call the variable catalog segmentation problem.

Catalog Segmentation Problem. A company has a collection of products and a collection of potential clients. Clients have various levels of interest associated with each type of product. The company wishes to produce several types of catalogs, each type containing a subset of the products (the number of products in a catalog may be limited by considerations such as weight), and mail to every potential client at most one catalog (presumably, of a type that would be of interest to the client). Assuming that producing a catalog-type has unit cost, and that for each type i and client j there is a expected revenue of u_{ij} from mailing a catalog of type i to client j, which catalogs should the company produce in order to maximize its expected profit (expected benefit minus production cost)? If all potential types of catalogs can be listed beforehand and all values u_{ij} are known, then this is a special case of MFL, with the catalogs serving as facilities. (In [6] it is assumed that all types of catalogs cost the same to produce, and we follow this assumption in our presentation. More generally, we may associate a cost c_i for producing the catalog of type i, and then the problem becomes equivalent to MFL.)

Most of our results apply to a general class of maximizing submodular set functions, called *submodular maximum facility location*, that can be described as maximizing a non-decreasing submodular revenue function minus a linear cost function.

Submodular Maximum Facility Location (SMFL). Consider a set N of n facilities and a set function $f : 2^N \to R^+$. For any subset $S \subset N$, $f(S) = R(S) - c(S)$, where R is a non-negative non-decreasing submodular set function corresponding to the revenue, and $c(S) = \sum_{i \in S} c_i$ is a linear cost function. As a result, set function f is a non-monotone submodular function and the goal is to find a subset S that maximizes $f(S)$. [5] We assume a value oracle for the revenue function R and a description for the cost c (this is of polynomial size) are given.

MFL is a special case of SMFL. Moreover, one can show (details omitted) that capacitated maximum facility location (CMFL), in which every facility has a capacity that limits the number of clients that it can serve, is also a special case of SMFL. Other examples include a variety of optimization problems such as set buying, catalog segmentation [6], banner ad allocation problem with guaranteed delivery [3], maximizing influence in social networks [5, 8], and optimal sensor installation for outbreak detection [7].

2 The theory of PASS approximation

In this section we describe the notion of PASS approximation for the maximum facility location problem; in Section 4.2 we show how this notion extends to submodular maximum facility location. First note that for these problems the value of the objective function may be negative for some feasible solutions. As is often the case with objective functions that may be negative, the MFL problems are NP-hard to approximate within any constant factor (see for example Theorem 1). Therefore, researchers have attempted to present other types of performance guarantees. As we discuss in the next section, most of previous attempts suffer in that they do not prove guarantees with regard to the real

[5] Note that function f can be possibly negative and therefore the result of Feige et al. [4] does not apply.

optimum value on every instance. Nonetheless, there are large classes of interesting instances in which the approximation ratio can be much better than the worst-case guarantees, for example if the cost of opening facilities is far from the revenue one can get from the open facilities. Our goal is to get a better understanding of the approximation ratio, exposing classes of input instances for which a constant approximation ratio is possible.

Let us first describe an attempt that fails to resolve our concerns.

Relatively small costs. Based on the intuition of the previous paragraph, for $0 < \alpha < 1$, let us call an instance α-bounded if for every facility, the cost of opening the facility is at most α times the revenue one gets by connecting all clients to the facility. Is it the case that when α is sufficiently small there is a constant approximation for MFL for α-bounded instances? The answer is negative. The proof involves starting from a hard to approximate instance of MFL, and adding an additional client that provides revenue $\max(c_i/\alpha)$ regardless of which facility services it. This forces the instance to be α-bounded, while increasing the value of an optimal solution by only $\max(c_i/\alpha)$. An appropriate choice of parameters leads to the desired hardness result.

We now discuss a performance measure introduced by Kleinberg et al [6]. Unlike the notion of an α-bounded instance discussed previously, the idea is to use the notion of α-boundedness not with respect to the input instance, but rather with respect to its optimal solution. Namely, call a solution α-bounded if the total cost of opening the facilities in this solution is at most an α-fraction of the total revenue derived from all clients in the solution. [6] In [6], it is shown that for the catalog segmentation problem, whenever α is bounded away from 1, the approximation ratio of a natural greedy algorithm is a constant (that tends to 1 as α tends to 0). An exact statement of this result of [6] appears in Theorem 4. Our notion of performance guarantee can be viewed as a generalization of the notion used in [6]. A more detailed comparison between our work and that of [6] will appear in Section 4.

Consider an arbitrary MFL instance I and an arbitrary feasible solution S. For each facility i open in S, let c_i be its opening cost, and let $r_i = \sum u_{ij}$ (where the sum is taken over clients j connected to facility i in S) be the total revenue derived from clients connected in S to facility i. [7] Let α_i denote the ratio c_i/r_i. Intuitively, a value of α_i close to 0 indicates that opening the facility i was a favorable decision, because the revenue r_i that resulted from this opening came at relatively little cost. A value of α_i close to 1 indicates that the opening of facility i may have been questionable, as most of the revenue r_i is offset by the cost c_i. On a global scale, the total revenue of S is $R(S) = \sum_{i \in S} r_i$, the total cost is $C(S) = \sum_{i \in S} c_i$, and the value of solution S is $V(S) = R(S) - C(S)$. Similar to the local values α_i, we shall use α to denote an aggregate value $\alpha = C(S)/R(S)$.

---

[6] Technically, in [6] a different parameter μ is considered, which in our terminology is $\mu = \frac{1}{\alpha} - 1$. It is straightforward to translate results expressed in terms of μ to results expressed in terms of α and vice versa.

[7] A better notation might be to write $r_i(S)$ instead of r_i, but we use r_i for brevity.

Definition 1. *Given an instance I of MFL and a feasible solution S, and using notation as above.*

- *The* expanded signature *of S is the collection $\{(q_i, \alpha_i)\}$, where i ranges over all facilities open in S, $q_i = r_i/R(S)$, and $\alpha_i = c_i/r_i$.*
- *The* signature $sig(S)$ *of S is the collection $\{(q_i, \alpha_i)\}$ obtained from the expanded signature by unifying components that share the same value of α_i. Namely, in the signature i no longer refers to a specific facility, all α_i are distinct, r_i denotes the total revenue that comes from open facilities which share the same α_i value (namely $r_i = \sum_{\text{facilities } i':\alpha'_i=\alpha_i} r_{i'}$), and q_i denotes the fraction of revenue that comes from open facilities which share the same α_i value (namely, $q_i = r_i/R(S)$).*

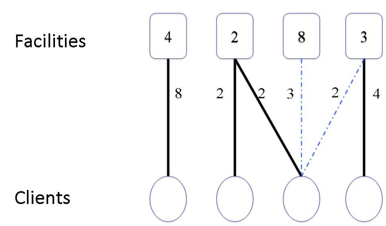

Fig. 1. An instance of the facility location problem: the costs of the facilities and the revenue from the clients (on the edges). An optimal solution is depicted by solid lines. The expanded signature is $\{(\frac{1}{2}, \frac{1}{2}), (\frac{1}{4}, \frac{1}{2}), (\frac{1}{4}, \frac{3}{4})\}$, the signature is $\{(\frac{3}{4}, \frac{1}{2}), (\frac{1}{4}, \frac{3}{4})\}$, and the summary signature is $\frac{9}{16}$.

Note that for every signature $\sum q_i = 1$, and that if all open facilities in S have the same value α_i then the signature is $(1, \alpha)$, in which case we abbreviate it to α. When open facilities have different values of α_i we may view $\alpha = C(S)/R(S)$ as a parameter that to some extent summarizes the signature, even though it does not have the same distinguishing power among solutions as the signature does. Using α as a *summary signature* will be convenient when we compare our results against previous results of [6]. Also, it is important to distinguish between the *expanded signature* and the *signature* in order to be able to talk about asymptotics in the hardness results as for any fixed expanded signature there are a fixed number of facilities. For the positive results, the notions of expanded signature and signature are interchangeable by changing the index of summation, and the reader may find it easier to interpret the positive results using the expanded signature.

In our framework of *PASS approximation*, we express the approximation ratios of algorithms as a function of the signature. Observe that an instance may have multiple different signatures (one for each feasible solution). Our approximation ratios will apply to all of them (and hence to the best of them). Nevertheless, the reader may find it convenient to think of the signature of an instance as that of (one of) its optimal solution(s). Given any feasible solution (e.g., an optimal one) with signature S, for every index i solution S generates a value of $r_i(1-\alpha_i)$ from facilities with α value equal to α_i. Our algorithms may open facilities different than those opened by S, but our accounting method will show that our algorithms recover value at least $\hat{v}_i = r_i(1-\alpha_i-\alpha_i \ln \frac{1}{\alpha_i})$ in exchange to the value generated by S from index i. This parameter $\hat{v}_i$ is therefore called the *recoverable value*, and, as we will prove, it is the optimal recoverable value (i.e., it is NP-hard to recover more). Note that $0 \leq \hat{v}_i \leq r_i$, with $\hat{v}_i = 0$ when $\alpha_i = 1$ (i.e., we can't recover any value from facilities whose cost equals their revenue) and $\hat{v}_i = r_i$ when $\alpha_i = 0$ (i.e., we can recover all the revenue from facilities with zero cost). To simplify the presentation in this paper, and with no significant effect on the results, we pretend that quantities such as $\ln x$ can be computed exactly in polynomial time for every x.

3 Our Results

We first present a hardness result and then give tight greedy and LP-based algorithms for MFL. The proof is omitted in the current version.

Theorem 1. *Let sig* $= \{(q_i, \alpha_i)\}$ *be an arbitrary signature, and consider the class of MFL instances that have an optimal solution with signature sig. For simplicity of notation, for each such instance, normalize the costs and revenues such that the revenue of the optimal solution having signature sig is 1, and hence its value is* $1 - \sum q_i \alpha_i$. *Then on this class of instances, for every* $\epsilon > 0$, *it is NP-hard to find a solution of value* $\sum \hat{v}_i + \epsilon$ *where* $\hat{v}_i = q_i(1-\alpha_i-\alpha_i \ln \frac{1}{\alpha_i})$.

Corollary 1. *For any* $\epsilon > 0$ *and* $\alpha = C(S)/R(S)$, *for any optimal solution* S, *it is NP-hard to approximate MFL within a ratio better than* $\frac{1-\alpha-\alpha \ln \frac{1}{\alpha}}{1-\alpha} + \epsilon$.

We show that there are algorithms with approximation ratios that match the hardness results, and moreover do so for wider classes of problems. The first class of algorithms that we consider is that of greedy algorithms. We shall distinguish between two types of greedy algorithms depending on whether it is greedy with respect to margin or to rate. Only one of these versions is optimal in our framework.

Theorem 2. *Let* I *be an arbitrary instance of MFL, let* S *be an arbitrary feasible solution and let* $\{(q_i, \alpha_i)\}$ *be the signature of* S. *For simplicity of notation, normalize the costs and revenues in* I *such that the revenue of* S *is 1, and hence its value is* $1 - \sum q_i \alpha_i$. *Then the greedy-rate algorithm produces a solution of value at least* $\sum \hat{v}_i$ *where* $\hat{v}_i = q_i(1-\alpha_i-\alpha_i \ln \frac{1}{\alpha_i})$.

Corollary 2. *The greedy-rate algorithm approximates MFL within a ratio of at least* $\frac{1-\alpha-\alpha \ln \frac{1}{\alpha}}{1-\alpha}$, *where* $\alpha = C(S)/R(S)$ *for any optimal solution* S.

Remark: As shown in Section 4.2, the result above holds for SMFL problem, under the appropriate definition of signature. We also remark that Corollaries 1 and 2 are each stronger than previous results proved in [6]. These issues will be discussed in Section 4.

The next class of algorithms that we consider is based on linear programming. It is also possible to show that the natural linear programming relaxation does not result in approximation ratios that match the hardness results of Theorem 1 (the details are omitted in this version). Hence, we introduce a new linear program, called the *recoverable value LP*, whose objective is to maximize the (fractional) recoverable value rather than the (fractional) true value.[8] We then show that the LP can be rounded to give a feasible solution of value not lower than the recoverable value of the LP.

Theorem 3. *The recoverable value LP for MFL can be solved in polynomial time. For every input instance and feasible solution S with signature $\{(q_i, \alpha_i)\}$, the LP has a solution of value at least as high as $\sum \hat{v}_i$, where $\hat{v}_i = r_i(1 - \alpha_i - \alpha_i \ln \frac{1}{\alpha_i})$. Any solution of the LP can be rounded in random polynomial time to give a feasible solution of expected value at least as high as the value of the objective function in the LP solution.*

3.1 Why use our notion of PASS approximation?

In this section we present arguments in favor of our notion of PASS approximation. The point that we will try to make is that performance measures guide the design of algorithms, and our performance measure appears to us to be a very good guide. We assume in the discussion below that the true goal is to maximize revenue minus cost, and compare various approaches that can be used in order to circumvent the inapproximability results for this measure. Recall that our approach of PASS approximation is to express the approximation ratio not as a function of the size of the input instance, but as of its signature.

As mentioned, MFL and SMFL are NP-hard to approximate. Therefore, researchers have attempted to present other types of performance guarantees. One existing theoretical approach for coping with inapproximable problems is to change the objective function in a way that preserves the spirit of the original problem. For example, one might consider the complement of the objective (e.g., vertex cover as opposed to independent set), or bicriteria approximations (e.g., bisection in graphs). In [2, 1] the approach taken was to measure the quality of a solution on a shifted scale which is always nonnegative. This is equivalent to changing the objective function by adding to it a sufficiently large constant that ensures that all solutions have nonnegative value. As an example, consider algorithms for MFL based on linear programming. In [1] a combination of a linear program and rounding technique is designed. They show that the approximation ratio of $2(\sqrt{2}-1)$ that they obtain is best possible (matches the integrality gap), but with respect to a shifted scale of the objective function. As we do not claim the same about our linear programming approach, then clearly there are instances in which the algorithm of [1] is better than ours. Likewise, there are instances on which our LP plus rounding gives better results (because we are

[8] We note this relaxation is not a relaxation in the usual sense, because the value of the objective function of the LP is a lower bound on the value of an optimal solution, rather than an upper bound.

optimal with respect to the structural approximation measure, whereas [1] are not). Hence it appears as if the results are incomparable. Nevertheless, we would like to convince the reader that even though the result of [1] is interesting mathematically, it does not really provide the kind of algorithmic insights that are relevant to the original problem. To obtain an approximation ratio of $2(\sqrt{2}-1)$ with respect to the shifted scale, it is safe to open every facility with probability at least $1-2(\sqrt{2}-1)$ (and at most 1), regardless of the cost of the facility, and regardless of whether any client wants to connect to the facility. This is a simple (and obviously counterproductive) rule of thumb that comes out of the shifted scale performance measure, and in fact the algorithm of [1] follows it. We view this as evidence that the shifted scale performance measure is not a good guide in the design of algorithms (with respect to the original objective function).

One can also show that other approaches designed to analyze heuristics, including optimizing with respect to a budget constraint, average-case and smoothed analysis, do not provide a desirable performance measure for the MFL problem. Details are omitted.

In comparison with [6], note that the approximation ratio in [6] is expressed as a function of one parameter that we refer to as the summary signature α. What is the advantage of presenting the more complicated signature $\{(q_i, \alpha_i)\}$? We see two advantages (beyond the obvious advantage of always providing a performance guarantee that is at least as good as that provided by the summary signature). One is *prescriptive*: the design of our LP is a natural consequence of our signature, valuing each star according to its own recoverable value. It would have been very difficult to design and analyze it without having at least implicitly a notion similar to the detailed signature. The other advantage is *conceptual*: our signature enjoys closure properties that the summary signature does not have. Given two disjoint instances of SMFL, the detailed signature becomes simply the union of the original detailed signatures, and the output guarantee (approximation ratio times value of optimal solution) is simply the sum of output guarantees of the two instances. For the summary signature, this is not true.

LP-based vs. greedy algorithms. The performance guarantees that we prove for the greedy algorithm and the LP-based algorithm are the same, and for the greedy algorithm we prove this performance guarantee for a wider range of instances (SMFL rather than just MFL). So what is the point of having an LP-based algorithm? There are several reasons to do this.

Most importantly, there is a conceptual difference between the use of PASS approximation framework for our greedy versus LP-based algorithms. For our greedy algorithm, the theory of structural approximation is *descriptive*. It describes the approximation ratios of existing algorithms, and may guide us in the choice of the greedy rule to use. For the LP-based approach, however, the theory of PASS approximation is not only descriptive, but also *prescriptive*. It guides us in the design of new algorithms. The definitions of the signature and recoverable value define for us the linear program and the rounding technique. While in our examples, the greedy algorithms happens to be tight, the LP-based approach may still be of value for other problems precisely because of it's prescriptive nature – *it is designed to produce tight algorithms.* For this reason, we consider the LP-based approach to be a significant contribution of our paper.

Another reason is so as to diversify our algorithmic toolbox. Even though the current paper is concerned with a class of problems for which the LP approach does not seem to offer significant advantages over the greedy approach, this need not be the case for other classes of problems. The development of a methodology of how to use linear programming relaxations in the context of structural approximation (which turns out to be different than the way linear programming relaxations are typically used in "classical" approximation) is anticipated to lead to rewards in future work.

4 A greedy approach

One standard approach for the MFL problems are greedy algorithms. In this section, we describe two plausible greedy algorithms for SMFL, and prove that one of them is optimal with regards to the PASS approximation. Given a set S of facilities, let $C(S)$ denote the total cost of facilities in S, let $R(S)$ denote the revenue of the optimum assignment given that the open facilities are those in S, and let $V(S) = R(S) - C(S)$ denote the total value of S. Given a facility i, let $M(i|S)$ denote the marginal revenue of i with respect to S. Namely, $M(i|S) = R(S \cup \{i\}) - R(S)$. If $i \in S$ then $M(i|S) = 0$.

The greedy algorithms construct a solution iteratively by selecting facilities that maximize some function of the marginal revenue $M(i|S)$. Given a partial solution (set of open facilities) S, the *greedy-rate* algorithm opens the facility i which maximizes the *rate* of increase in value, i.e., $\frac{M(i|S)-c_i}{M(i|S)}$, provided that this rate is positive. The *greedy-margin* algorithm simply opens the facility with the largest marginal value, i.e., $M(i|S) - c_i$, provided that this value is positive.

The greedy step can be implemented in polynomial time for the special cases of SMFL mentioned in Section 1. (For example, for CMFL, implementing the greedy step involves computing the optimal assignment of clients to the open facilities subject to the capacity constraints. This can be solved in polynomial time via an algorithm for the so called B-matching problem in bipartite graphs.) However, in general, the greedy step for SMFL might be NP-hard.

4.1 Comparison to KPR

The greedy-margin algorithm was studied by Kleinberg, Papadimitriou, and Raghavan [6] for the catalogue segmentation problem (and generalizations which maintain the property of uniform-cost facilities). [9] They proved the following theorem (Theorems 2.3 and 2.4 in [6]):

[9] The greedy algorithm specified prior to Theorem 2.3 in [6] does not specify a rule of which facility to open next, as long as its marginal revenue is larger than its cost. However, the proof of Theorem 2.4 in [6] is based on the use of a greedy-margin rule, without stating this explicitly.

Theorem 4. **[Kleinberg, Papadimitriou and Raghavan.]** *For the catalogue segmentation problem, the greedy-margin algorithm achieves an approximation ratio of at least* $1 + \alpha - 2\sqrt{\alpha}$*, where* $\alpha = C(S)/R(S)$ *for any optimal solution* S*. There are instances on which the approximation ratio of the algorithm is no better than* $1 - \alpha$*.*

In this section, we will improve upon this result by generalizing the analysis to accommodate non-uniform facility costs and providing improved approximation guarantees (Theorem 2) together with a matching NP-hardness result (Theorem 1). In doing so, we must be careful with our choice of greedy algorithm. We have defined two natural greedy algorithms – the greedy-rate and the greedy-margin algorithm – and in fact in uniform-cost settings such as that of [6] these two algorithms coincide as the rate of a facility is monotone in its marginal revenue. But for non-uniform facility costs, as the following simple example illustrates, the greedy-margin algorithm gives very poor results.

Example. There are n clients and $n+1$ facilities. Facility i, $1 \le i \le n$, has cost 1, revenue 2 for client i, and 0 revenue for all other clients; Facility $n+1$ has cost $n-2$ and revenue 1 per client. The optimal solution will open the first n facilities for a value of n, whereas the greedy-margin rule will open only facility $n+1$ for a value of 2.

By contrast, for the greedy-rate algorithm, the approximation ratio that we prove is strictly better than that proved in [6]. See Figure 2 for a detailed comparison of our bounds with those of [6].

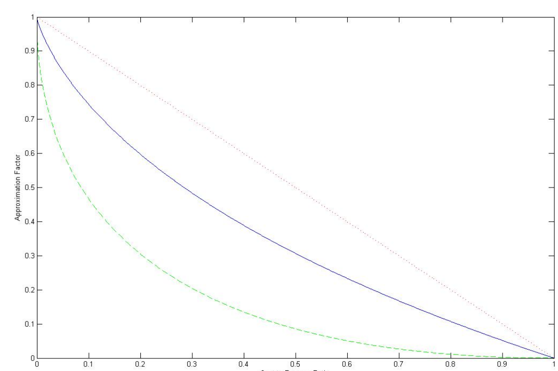

Fig. 2. The solid line is the approximation ratio of the greedy-rate algorithm plotted for $\alpha \in [0,1]$, as proved in Corollary 2. For every value of α improving over this approximation ratio is NP-hard, as proved in Theorem 1. The dashed and dotted lines depict the lower and upper bounds proved in [6].

4.2 Approximation as a function of α

In this section, we prove Corollary 2 which gives an approximation factor for the greedy rate algorithm as a function of the summary signature α. The following simple observation (appearing in [6]) is of key importance, and hence we state it as a lemma.

Lemma 1. *Let i be a facility and let S and T be sets of facilities. Then the marginal revenue of facility i with respect to S is at least as large as the loss in marginal revenue of T when facility i is added to S.*

$$M(i|S) \geq M(T|S) - M(T|S \cup \{i\})$$

Proof. By the fact that the revenue function is nondecreasing, we have $R(S \cup \{i\} \cup T) \geq R(S \cup T)$. Breaking each revenue to a sum of marginal revenues we have $R(S) + M(i|S) + M(T|S \cup \{i\}) \geq R(S) + M(T|S)$. Canceling the $R(S)$ and subtracting $M(T|S \cup \{i\})$ from both sides, the lemma is proved. □

We now proceed to prove our improved bounds. As stated, our analysis applies to any problem with a nondecreasing submodular revenue function and linear cost function.

Lemma 2. *The value of the greedy-rate algorithm is at least $R(O)(1-\alpha-\alpha \ln \frac{1}{\alpha})$ where O is an optimal solution, and $\alpha = C(O)/R(O)$.*

Proof. We analyze the value of greedy-rate up to the first point in time in which its total revenue meets or exceeds $R(O) - C(O)$. Let $\mu(x)$ denote the rate at which the value obtained by the greedy-rate algorithm increases when it has already made a revenue of x. Observe that at a point when greedy has already made a revenue of $x < R(O) - C(O)$, the marginal revenue of O is at least $R(O) - x$ (by Lemma 1). By submodularity of the revenue function, at this point there must be at least one facility of O with rate $\frac{R(O)-x-C(O)}{R(O)-x}$. The rate at which the value increases at each point in time is at least as high as the rate one would get by choosing the highest rate among the facilities of O at the same time. Therefore, $\mu(x) \geq \frac{R(O)-x-C(O)}{R(O)-x}$. As the total value of greedy-rate, $V(G)$, is the integral of the rate of increase of the value, we have:

$$\begin{aligned}
V(G) = \int \mu(x)dx \geq \int_0^{R(O)-C(O)} & \frac{R(O) - x - C(O)}{R(O) - x} dx \\
&= R(O) \int_0^{1-\alpha} \frac{1 - x - \alpha}{1 - x} dx \\
&= R(O) \int_0^{1-\alpha} (1 - \frac{\alpha}{1 - x}) dx \\
&= R(O)(1 - \alpha - \alpha(\ln \frac{1}{1 - (1 - \alpha)} - \ln 1)) \\
&= R(O)(1 - \alpha - \alpha \ln \frac{1}{\alpha})
\end{aligned}$$

□

The approximation ratio of Corollary 2 follows from Lemma 2 together with the fact that $V(O) = R(O)(1-\alpha)$. The NP-hardness results appear in Theorem 1, and they naturally provide examples where the approximation ratio of greedy-rate is no better than claimed even in the special case of linear revenue functions and uniform costs.

4.3 Approximation as a function of the signature

The approximation ratio in Corollary 2 is expressed as a function of the summary signature α, whereas stronger performance guarantees can be given by expressing the approximation ratio as a function of the signature $\{q_i, \alpha_i\}$, as stated in Theorem 2.

In fact, we prove Theorem 2 for the general submodular facility location problems. To do so, we should extend the notion of a signature to a solution for submodular maximum facility location. The difficulty is that even though the cost of every open facility is well defined, its revenue is not. Hence we refine the notion of a solution to be represented not as a set of open facilities, but as an ordered set (a tuple). Namely, the open facilities are given (after renaming) in some order $1, 2, \ldots$ (even though this order is irrelevant to the actual value of the solution). Thereafter, a refined parameter α'_i in the expanded signature is defined relative to the marginal revenue of facility i with respect to this order. That is, $\alpha'_i = c_i/M_i$, where here M_i is shorthand notation for $M(i|\{1, \ldots, i-1\})$. Likewise, we define $q'_i = M_i / \sum M_j$. Using this notation, we can now strengthen Lemma 2. The proofs of the lemmas are omitted.

Lemma 3. *Let S be an arbitrary (ordered) solution for submodular maximum facility location with expanded signature $\{(q'_i, \alpha'_i)\}$ and total revenue normalized to 1. Then the value of the greedy-rate algorithm is at least $\sum_{i \in S} q'_i(1 - \alpha'_i - \alpha'_i \ln \frac{1}{\alpha'_i})$.*

To motivate the following lemma, observe that Lemma 3 by itself does not capture the notion of PASS approximation that we have for the special case of MFL. For a given set of open facilities in MFL, the optimal choice of allocation of clients might not correspond to revenues r_i per facility that are equal to M_i for any ordering of facilities. For example, if there are two facilities and two clients, where client i has revenue 2 if connected to facility i and revenue 1 if connected to the other facility, then in the optimal solution $r_1 = r_2 = 2$, whereas for any ordering $M_1 = 3$ and $M_2 = 1$.

Lemma 4. *Let S be an arbitrary solution for maximum facility location with expanded signature $\{(q_i, \alpha_i)\}$. Then there is an ordering of the facilities of S giving $\sum_{i \in S} q'_i(1 - \alpha'_i - \alpha'_i \ln \frac{1}{\alpha'_i}) \geq \sum_{i \in S} q_i(1 - \alpha_i - \alpha_i \ln \frac{1}{\alpha_i})$.*

The combination of Lemmas 3 and 4 imply Theorem 2 (and also the generalization of Theorem 2 to SMFL).

5 A linear programming approach

In this section, we develop an LP-based approach for MFL. It is based on an interplay between the notions of the true value of a solution and the recoverable value of the solution. Recall that the recoverable value (see definition in Section 2), which in general is lower than the true value, represents our approximation goal in the sense that we wish to find a solution of true value at least equal to that of the recoverable value of the best integral solution. First we introduce a new LP relaxation for the general problem called the *recoverable value relaxation*. This LP captures the natural constraints for the MFL problem, but has an objective function describing the recoverable value of the solution rather than the true value. Hence the LP provides a fractional solution that maximizes the recoverable value, and we denote this value by $\hat{V}_f$. We round this fractional solution to an integral one of (expected) true value at least $\hat{V}_f$, thus meeting our approximation goal. Moreover, we can solve the recoverable value LP in polynomial time.

While our approach is general, we have been unable to analyze it for CMFL or more general variants, and leave this as an open question.

5.1 An LP relaxation

Recall that in the MFL problem, each facility $i \in \mathcal{F}$ has an opening cost of c_i and each client $j \in \mathcal{J}$ has a revenue u_{ij} for being connected to facility i. We call pair (i,T) of a facility i and a subset T of clients connected to it a *star*. Let x_{iT} be an indicator variable of star (i,T), i.e., that facility i is opened and connected to clients $j \in T$. The revenue of connecting the clients in T to facility i is $r_{iT} = \sum_{j \in T} u_{ij}$. For every star (i,T) we associate a *recoverable value* which is $\hat{v}_{iT} = r_{iT}(1 - \alpha_{iT} - \alpha_{iT} \ln \frac{1}{\alpha_{iT}})$, where $\alpha_{iT} = \frac{c_i}{r_{iT}}$. Then the optimal fractional recoverable value is described by the following LP, called the *recoverable value LP relaxation*.

$$\begin{aligned} \text{maximize} \quad & \sum \hat{v}_{iT} x_{iT} && && (1) \\ \text{subject to} \quad & \sum_{i,T: j \in T} x_{iT} \le 1 && j \in \mathcal{J} \\ & \sum_{T \subseteq \mathcal{J}} x_{iT} \le 1 && i \in \mathcal{F} \\ & x_{iT} \ge 0 && i \in \mathcal{F}, T \subseteq \mathcal{J} \end{aligned}$$

The first inequality guarantees that each client contributes revenue to at most one facility and the second inequality guarantees that each facility is opened at most once. Every integral solution satisfies these constraints. Hence the value of the LP is at least as large as the recoverable value of the best integer solution (the one maximizing the recoverable value).

Let $\hat{V}_f$ be the optimal fractional recoverable value, namely, the optimal value to the above LP. Let V_f be the fractional true value associated with this solution, namely $\sum (r_{iT} - c_i) x_{iT}$. Typically, LP-relaxations provide upper bounds for maximization problems. In contrast, it is not in general true that V_f provides an upper bound on the true value of the best integer solution. Instead, as Lemma 5 will show, $\hat{V}_f$ provides a lower bound.

Our LP has exponentially many variables; however, we can solve it using the ellipsoid method. We solve the separation oracle of the dual linear program using a greedy algorithm. This algorithm exploits concavity of the recoverable values and some other structural properties of the dual.

Our randomized rounding procedure is composed of two steps: The first step considers facilities independently. Facilities of 0-cost are always opened. For the remaining facilities, $\alpha_{iT} > 0$. For each such facility i and each star (i, T) let $\beta_{iT} = x_{iT} \ln \frac{1}{\alpha_{iT}}$. Let $\beta_i = \sum_T \beta_{iT}$. We open facility i with probability $\min[\beta_i, 1]$. The first step might open several facilities with overlapping sets of clients. In the second step, we assign any over-demanded client j to the facility to which it contributes the maximum revenue. The following lemma yields Theorem 3. For the lack of space, the proof is omitted here.

Lemma 5. *Consider an optimal fractional solution of LP (1), with fractional recoverable value $\hat{V}_f$. For the MFL problem, our randomized rounding technique achieves an integral solution of expected (true) value at least $\hat{V}_f$.*

Acknowledgements

The work of the first author was supported in part by The Israel Science Foundation (grant No. 873/08).

References

1. Ageev, A., Sviridenko, M.: An 0.828-approximation algorithm for uncapacitated facility location problem. Discrete Applied Mathematics 93 (1999) 289–296
2. Cornuejols, G., Nemhauser, G., Wolsey, L.: Locations of bank accounts to optimize float: an analytic study of exact and approximate algorithms. Management Science 23 (1977) 789–810
3. Feige, U., Immorlica, N., Mirrokni, V., Nazerzadeh, H.: A combinatorial allocation mechanism with penalties for banner advertising. To appear in the Proceedings of the 17th International World Wide Web Conference (WWW) (2008)
4. Feige, U., Mirrokni, V., Vondrak, J.: Maximizing non-monotone submodular functions. In: Proceedings of the forty-eigth annual IEEE symposium on Foundations of Computer Science (FOCS). (2007) 461–471
5. Kempe, D., Kleinberg, J., Éva Tardos: Maximizing the spread of influence through a social network. In: KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, New York, NY, USA, ACM (2003) 137–146
6. Kleinberg, J.M., Papadimitriou, C.H., Raghavan, P.: Segmentation problems. J. ACM 51(2) (2004) 263–280
7. Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., VanBriesen, J., Glance, N.S.: Cost-effective outbreak detection in networks. In: KDD. (2007) 420–429
8. Mossel, E., Roch, S.: On the submodularity of influence in social networks. In: Proceedings of ACM STOC. (2007) 128–134