Approximation Algorithms for Facility Location Problems

David B. Shmoys

Cornell University, Ithaca NY 14853, USA

Abstract. One of the most flourishing areas of research in the design and analysis of approximation algorithms has been for facility location problems. In particular, for the metric case of two simple models, the uncapacitated facility location and the k-median problems, there are now a variety of techniques that yield constant performance guarantees. These methods include LP rounding, primal-dual algorithms, and local search techniques. Furthermore, the salient ideas in these algorithms and their analyses are simple-to-explain and reflect a surprising degree of commonality. This note is intended as companion to our lecture at CONF 2000, mainly to give pointers to the appropriate references.

1 A tale of two problems

In the past several years, there has been a steady series of developments in the design and analysis of approximation algorithms for two facility location problems: the uncapacitated facilily location problem, and the k-median problem. Furthermore, although these two problems were always viewed as closely related, some of this recent work has not only relied on their interrelationship, but also given new insights into the ways in which algorithms for the former problem yield algorithms, and performance guarantees, for the latter.

In the *k-median problem*, the input consists of a parameter k, and n points in a metric space; that is, there is a set $\mathcal{N}$ and for each pair of points $i, j \in \mathcal{N}$, there is a given distance $d(i, j)$ between them that is symmetric (i.e., $d(i, j) = d(j, i)$, for each $i, j \in \mathcal{N}$), satisfies the triangle inequality (i.e., $d(i, j) + d(j, k) \geq d(i, k)$, for each $i, j, k \in \mathcal{N}$), and also has the property that $d(i, i) = 0$ for each $i \in \mathcal{N}$. The aim is to select k of the n points to be *medians*, and then assign each of the n input points to its closest median so as to minimize the average distance that an input point is from its assigned median. Early work on the k-median problem was motivated by applications in facility location: each median corresponds to a facility to be built, and the input set of points corresponds to the set of clients that need to be serviced by these facilities; there are resources sufficient to build only k facilities, and one wishes to minimize the total cost of servicing the clients.

In the *uncapacitated facility location problem*, which is also referred to as the *simple plant location problem*, the strict requirement that there be k facilities is relaxed, by introducing a cost associated with building a facility; these costs are then incorporated into the overall objective function. More precisely, the input consists of two sets of points (which need not be disjoint), the potential facility location points $\mathcal{F}$, and the set of clients $\mathcal{C}$; in this case, we shall let n denote $|\mathcal{F} \cup \mathcal{C}|$. For each point $i \in \mathcal{F}$, there is a given cost f_i that reflects the cost incurred in opening a facility at this location. We wish to decide which facilities to open so as to minimize the total cost for opening them, plus the total cost of servicing each of the clients from its closest open facility. The uncapacitated facility location problem is one of the most well-studied problems in the Operations Research literature, dating back to the work of Balinski [2], Kuehn & Hamburger [11], Manne [14], and Stollsteimer [18, 19] in the early 60's.

Throughout this paper, a *ρ-approximation algorithm* is a polynomial-time algorithm that always finds a feasible solution with objective function value within a factor of ρ of optimal. Hochbaum [8] showed that the greedy algorithm is an $O(\log n)$-approximation algorithm for this problem, and provided instances to verify that this analysis is asymptotically tight. In fact, this result was shown for the more general setting, in which the input points need not belong to a metric space.

Lin & Vitter [12] gave an elegant technique, called *filtering*, for rounding fractional solutions to linear programming relaxations. As one application of this technique for designing approximation algorithms, they gave another $O(\log n)$-approximation algorithm for the uncapacitated facility location problem. Furthermore, Lin & Vitter gave an algorithm for the k-median problem that finds a solution for which the objective is within a factor of $1+\epsilon$ of the optimum, but is infeasible since it opens $(1+1/\epsilon)(\ln n+1)k$ facilities. Both of these results hold for the general setting; that is, the input points need not lie in a metric space. In a companion paper, Lin & Vitter [13] focused attention on the metric case, and showed that for the k-median problem, one can find a solution of cost no more than $2(1+\epsilon)$ times the optimum, while using at most $(1+1/\epsilon)k$ facilities.

The recent spate of results derive algorithms that can be divided, roughly speaking, into three categories. There are *rounding algorithms* that rely on linear programming in the same way as the work of Lin & Vitter, in that they first solve the linear relaxation of a natural integer programming formulation of the problem, and then round the optimial LP solution to an integer solution of objective function value no more than factor of ρ greater, thereby yielding a ρ-approximation algorithm. The second type of algorithm also relies on the linear programming relaxation, but only in an implicit way; in a *primal-dual algorithm*, the aim is to simultaneously derive a feasible integer solution for the original problem, as well as a feasible solution to the dual linear program to its linear relaxation. If one can show that the objective function value of the former always is within a factor of ρ of the latter, then this also yields a ρ-approximation algorithm. Finally, there are *local search algorithms*, where one maintains a feasible solution to the original problem, and then iteratively attempts to make a minor modification (with respect to a prescribed notion of "neighboring" solutions) so as to yield a solution of lower cost. Eventually, one obtains a local optimum; that is, a solution of cost no more than that of each of its neighboring solutions. In this case, one must also derive the appropriate structural properties in order to conclude that any locally optimal solution is within a factor of ρ of the global optimum. These three classes of algorithms will immediately be blurred; for example, some algorithms will start with an LP rounding phase, but end with a local search phase.

One other class of algorithmic techniques should also be briefly mentioned. Arora, Rao, & Raghavan [1] considered these two problems for geometrically-defined metrics. For the 2-dimensional Euclidean case of the k-median problem (either when the medians must be selected from among the input points, or when they are allowed to be selected arbitrarily from the entire space) and the uncapacitated facility location problem, they give a randomized polynomial approximation scheme; that is, they give a randomized $(1+\epsilon)$-approximation algorithm, for any fixed $\epsilon > 0$. No such schemes are likely to exist for the general metric case: Guha & Khuller [7] proved lower bounds, respectively, of 1.463 and $1 + 1/e$ (based on complexity assumptions, of course) for the uncapacitated facility location problem and k-median problems.

2 LP rounding algorithms

The first approximation algorithm with a constant performance guarantee for the (metric) uncapacitated facility problem was given by Shmoys, Tardos, & Aardal [17]. Their algorithm is an LP rounding algorithm; they give a natural extension of the techniques of Lin & Vitter [13] to yield an algorithm with performance guarantee equal to $3/(1-e^{-3}) \approx 3.16$.

Guha & Khuller [7] observed that one can strengthen the LP relaxation by approximately guessing the proportion of the overall cost incurred by the facilities in the optimal solution, and adding this as a constraint. Since there are only a polynomial number of reasonably-spaced guesses, we can try them all (in polynomial time). Guha & Khuller showed that by adding a local search phase, starting with the solution obtained by rounding the optimum to the "right" LP, one obtains a 2.41-approximation algorithm. For their result, the local search is extremely simple; one checks only whether some additional facility can be opened so that the overall cost decreases, and if so, one adds the facility that most decreases the overall cost.

The LP rounding approach was further strengthened by Chudak & Shmoys [5, 6] who used only the simple LP relaxation (identical to the one used by Lin & Vitter and Shmoys, Tardos, & Aardal), but relied on stronger information about the structure of optimal solutions to the linear programming relaxation to yield a performance guarantee of $1 + 2/e \approx 1.74$. Another crucial ingredient in their approach is that of randomized rounding, which is a technique introduced by Raghavan & Thompson [16] in the context of multicommodity flow; in this approach, the fractional values contained in the optimal LP solution are treated as probabilities. Chudak & Shmoys showed how to incorporate the main decomposition results of [17] so as to obtain a variant that might best be called *clustered randomized rounding*. Chudak & Shmoys also showed that the technique of Guha & Khuller could be applied to their algorithm, and by doing so, one improves the performance guarantee by a microscopic amount.

The first constant performance guarantee for the k-median problem also relied an LP rounding approach: Charikar, Guha, Tardos, & Shmoys [4] gave a 20/3-approximation algorithm. A wide class of combinatorially-defined linear programs has the property that there exists an optimal solution with the property that each variable is equal to 0, 1/2, or 1; as will be discussed by Hochbaum in another invited talk at this workshop, this property often yields a 2–approximation algorithm as a natural corollary. Unfortunately, the linear relaxation used for the k-median problem does not have this property. However, the algorithm of [4] is based on the idea that one can reduce the problem to an input for which there is such an 1/2-integral solution, while losing only a constant factor, and then subsequently round the 1/2-integral solution to an integer one.

3 Primal-dual algorithms

In one of the most exciting developments in this area, Jain & Vazirani [9] gave an extremely elegant primal-dual algorithm for the uncapacitated facility location problem; they showed that their approach yields a 3-approximation algorithm, and it can be implemented to run in $O(n^2 \log n)$ time. In fact, Jain & Vazirani proved a somewhat stronger performance guarantee; they showed that if we overcharge each facility used in the resulting solution by a factor of 3, then the total cost incurred is still within a factor of 3 of the optimum cost for the unaltered input. In fact, this overcharging is also within a factor of the 3 of the value of the LP relaxation with the original cost data. Mettu & Plaxton [15] give a variant of this algorithm (which is not explicitly a primal-dual algorithm, but builds substantially on its intuition) that can be implemented to run in $O(n^2)$ time, which is linear in the size of the input; this variant does not have the stronger guarantee (but is still a 3-approximation algorithm).

This stronger performance guarantee for the uncapacitated facility location problem has significant implications for the k-median problem. One natural connection between the two problems works as follows: the facility costs can be viewed as Lagrangean multipliers that enforce the constraint that exactly k facilities are used. Suppose that we start with an instance of the k-median problem, and define an input to the uncapacitated facility location problem by letting $\mathcal{C} = \mathcal{F} = \mathcal{N}$, and setting the cost of each facility to be a common value ϕ. If $\phi = 0$, then clearly all facilities will be opened in the optimal solution, whereas for a sufficiently large value of ϕ, only one facility will be opened in the optimal solution. A similar trade-off curve will be generated by any (reasonable) algorithm for the uncapacitated facility location problem.

Jain & Vazirani observed that if their approximation algorithm is used, and ϕ is set so that the number of facilities opened is exactly k, then the resulting solution for the k-median problem has cost within a factor of 3 of optimal. Unfortunately, the trade-off curve need not be continuous, and hence it is possible that no such value of ϕ exists. However, if this unlucky event occurs, then one can generate two solutions from essentially equal values of ϕ, where one solution has more than k facilities and the other has fewer; Jain & Vazirani show how to combine these two solutions to obtain a new solution with k facilities of cost that is within a factor of 6 of optimal. Charikar & Guha [3] exploited the fact that these two solutions have a great deal of common structure in order to give a refined analysis; in this way, they derive a 4-approximation algorithm for this problem.

Mettu & Plaxton [15] also show how to extend their approach to the uncapacitated facility location problem to obtain an $O(n^2)$-time algorithm for the k-median problem; in fact, their algorithm has the miraculous property that it outputs a single permutation of the input nodes such that, for *any* k, the first k nodes constitute a feasible solution within a constant factor of the optimal k-node solution. Thorup [20] has also given linear-time constant-factor approximation algorithms for the k-median problem; if the metric is defined with respect to a graph with m edges, then he gives a $12+o(1)$-approximation algorithm that runs in $O(m)$ time.

4 Local search algorithms

Local search is one of the most successful approaches to computing, in practice, good solutions to NP-hard optimization problems. Indeed, many of the most visible methods to the general scientific community are variations on this theme; simulated annealing, genetic algorithms, even neural networks can all be viewed as coming from this family of algorithms. In a local search algorithm, more narrowly viewed, one defines a graph on the space of all (feasible) solutions, where two solutions are neighbors if one solution can be obtained from the other by a particular type of modification. One then searches for a node that is locally optimal, that is, whose cost is no more than each of its neighbors, by taking a walk in the graph along progressively improving nodes.

Korupolu, Plaxton, & Rajaraman [10] analyzed a local search algorithm in which two nodes are neighbors if exactly one facility is added (or, symmetrically, deleted) in comparing the two solutions, or both one facility is added *and* one facility is deleted. They show that a local optimum with respect to this neighborhood structure has cost within a factor of 5 of optimal. One further issue needs to be considered in deriving an approximation algorithm: the algorithm needs to run in polynomial time. This can be accomplished in a variety of ways, but typically involves computing a "reasonably good" solution with which to start the search, as well as insisting that a step not merely improve the cost, but improve the cost "significantly". In this way, they show that, for any $\epsilon > 0$, local search can yield a $(5+\epsilon)$-approximation algorithm. Charikar & Guha [3] give a more sophisticated neighborhood structure that yields another simple 3-approximation algorithm. Furthermore, they show that rescaling the relative weight of the facility and the assignment costs leads to a 2.415-approximation algorithm based on local search. Finally, they show that all of the ideas above can be combined: LP rounding (on multiple LPs augmented as in [7] and augmented with a greedy improvement phase), primal-dual algorithms (improved with the rescaling idea), and the improved local search algorithm. In this way, they improve the performance guarantee of 1.736 due to Chudak & Shmoys to 1.728. This might be the best guarantee known as of this writing, but for these two problems, it seems unlikely to be the last word.

References

1. S. Arora, P. Raghavan, and S. Rao. Approximation schemes for Euclidean k-medians and related problems. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 106–113, 1998.
2. M. L. Balinksi. On finding integer solutions to linear programs. In *Proceedings of the IBM Scientific Computing Symposium on Combinatorial Problems*, pages 225–248. IBM, 1966.
3. M. Charikar and S. Guha. Improved combinatorial algorithms for the facility location and k-median problems. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 378–388, 1999.
4. M. Charikar, S. Guha, É. Tardos, and D. B. Shmoys. A constant-factor approximation algorithms for the k-median problem. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 1–10, 1999.
5. F. A. Chudak. Improved approximation algorithms for uncapacitated facility location. In R. E. Bixby, E. A. Boyd, and R. Z. Ríos-Mercado, editors, *Integer Programming and Combinatorial Optimization*, volume 1412 of *Lecture Notes in Computer Science*, pages 180–194, Berlin, 1998. Springer.
6. F. A. Chudak and D. B Shmoys. Improved approximation algorithms for the uncapacitated facility location problem. Submitted for publication.
7. S. Guha and S. Khuller. Greedy strikes back: Improved facility location algorithms. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 649–657, 1998.
8. D. S. Hochbaum. Heuristics for the fixed cost median problem. *Math. Programming*, 22:148–162, 1982.
9. K. Jain and V. V. Vazirani. Primal-dual approximation algorithms for metric facility location and k-median problems. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 2–13, 1999.
10. M. R. Korupolu, C. G. Plaxton, and R. Rajaraman. Analysis of a local search heuristic for facility location problems. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1–10, 1998.
11. A. A. Kuehn and M. J. Hamburger. A heuristic program for locating warehouses. *Management Sci.*, 9:643–666, 1963.
12. J.-H. Lin and J. S. Vitter. ϵ-approximations with minimum packing constraint violation. In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pages 771–782, 1992.
13. J.-H. Lin and J. S. Vitter. Approximation algorithms for geometric median problems. *Inform. Proc. Lett.*, 44:245–249, 1992.
14. A. S. Manne. Plant location under economies-of-scale-decentralization and computation. *Management Sci.*, 11:213–235, 1964.
15. R. R. Mettu and C. G. Plaxton. The online median problem. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, 2000, to appear.

16. P. Raghavan and C. D. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7:365–374, 1987.
17. D. B. Shmoys, É. Tardos, and K. I. Aardal. Approximation algorithms for facility location problems. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing*, pages 265–274, 1997.
18. J. F. Stollsteimer. *The effect of technical change and output expansion on the optimum number, size and location of pear marketing facilities in a California pear producing region.* PhD thesis, University of California at Berkeley, Berkeley, California, 1961.
19. J. F. Stollsteimer. A working model for plant numbers and locations. *J. Farm Econom.*, 45:631–645, 1963.
20. M. Thorup. Quick k-medians. Unpublished manuscript, 2000.