Approximation Scheme for the Geometric Connected Facility Location Problem[1]

Flávio K. Miyazawa*, Lehilton L. C. Pedrosa†,
Rafael C. S. Schouery‡ and Renata G. D. de Souza§
Institute of Computing, University of Campinas
Campinas, São Paulo, Brazil, 13083-852
Email: *fkm@ic.unicamp.br, †lehilton@ic.unicamp.br, ‡schouery@ic.unicamp.br, §renataghisloti@gmail.com

***Abstract*—In the well-known *Connected Facility Location Problem* (CFL), we are given sets of clients and locations over a metric space, the cost to install a facility at each location, and a number $M \geq 1$. A solution is a set of locations where to install facilities, a tree spanning these facilities, and an assignment from each client to an open facility. The goal is to obtain a solution that minimizes the cost to connect clients and install facilities plus the cost of the spanning tree multiplied by M. We consider the geometric and prize-collecting version of CFL (GCFL), that is the variant where the metric is the Euclidean plane, one may open a facility at any point at fixed cost f, and may opt to pay a penalty, instead of assigning a given client. Our main result is a polynomial-time approximation scheme for GCFL.**

I. Introduction

The classical metric Facility Location Problem (FLP) is a widely studied problem with many real world applications. For its geometric version, Arora et al. [2] obtained an approximation scheme, by using a dynamic program based on the *shifted dissections*, first used for the Traveling Salesman Problem and the Minimum Steiner Tree Problem [1]. This result was somewhat surprisingly, as the latter depended on the so called *patching lemma*, that has no correspondent in the FLP case. Informally, such a lemma states that a set of edges in the solution that cross a given line may be replaced by a different set of edges that cross this line only a constant number of times, and whose cost is at most a constant times larger. To address the lack of a patching lemma, they noticed that, in the FLP, each client is connected to the closest open facility, and so for the dynamic program it is sufficient to guess this distance up to an additive small error.

For the GCFL, however, the techniques of [2] are not sufficient, as one must also account for the cost of the (Steiner) spanning tree connecting the facilities, and a facility is not necessarily connected to its closest pair. Moreover, contrary to the case of the Minimum Steiner Tree, for which a patching lemma is trivially obtained, the existence of such a lemma for the GCFL is unlikely, since adding a node to the (Steiner) tree incurs in additional cost to install the facility. It was open whether Arora's framework could be used to deal with node-weighted problems [3]. We show that this is indeed the case, by giving a polynomial approximation scheme for the GCFL.

[1] This work has been partially supported by Fapesp (2014/14209-1, 2013/21744-8) and CNPq (311499/2014-7).

II. Approximation Scheme

a) Well-rounded solutions: We reduce the problem to the case in which all points have odd integral coordinates, and the length L of the bounding box containing the points is polynomially bounded. This version is denoted by LCFL.

Lemma 1. *There exist an ε-reduction from GCFL to LCFL.*

b) Shifted dissections: Let $a, b \in \{0, 2, ..., L\}$ be even coordinates drawn at random, and G_0 be the square with vertices $(a - L, b - L)$ and $(a + L, b + L)$. A *quadtree* $QT_{a,b}$ is a recursive partitioning of G_0, such that each square G is divided into four subsquares, until the subquares have side 2.

c) (m, r)-restricted solutions: To obtain a subproblem to the dynamic program, Arora [1] considers only solutions whose edges cross the edges of a square G of $QT_{a,b}$ in m fixed points. A solution with this property is called m-light.

In [1], it is also assumed that a solution crosses the borders of a square only a constant number of times, so that subproblems may be enumerated. To obtain one such solution, one uses a patching lemma. Since there is no patching lemma for GCFL, we proceed to a more involved analysis of the problem. We show that there is a solution whose number of internal components induced in each square is bounded by a constant r. While there might still be much more edges crossing the border of a square, only a constant part of them is considered in each subproblem.

Lemma 2. *Let $\varepsilon > 0$. For sufficiently large $r = O(1)$, and $m = O(\log n)$, there exists an (m, r)-restricted solution whose cost is at most $(1 + \varepsilon)OPT$, where OPT is the optimal cost.*

Theorem 3. *There exists a PTAS for GCFL.*

III. Open problems

An $O(n)$-approximation is used in the reduction to well-rounded case. It is open whether there is a constant factor for the prize-collecting CFL. Also, it would be interesting extending the algorithm to more general node-weighted problems.

References

[1] S. Arora. Polynomial time approximation schemes for Euclidean TSP and other geometric problems. In *FOCS*, pages 2–11, 1996.
[2] S. Arora, P. Raghavan, and S. Rao. Approximation Schemes for Euclidean K-medians and Related Problems. In *STOC*, pages 106–113. ACM, 1998.
[3] J. Remy and A. Steger. Approximation Schemes for Node-Weighted Geometric Steiner Tree Problems. *Algorithmica*, 55(1):240–267, 2007.