Approximation Algorithms for the Minimum-Length Corridor and Related Problems

Arturo Gonzalez-Gutierrez*   Teofilo F. Gonzalez†

Abstract

Given a rectangular boundary partitioned into rectangles, the MLC-R problem is to find a Minimum Edge-Length *Corridor* that visits at least one point from the boundary of every rectangle and from the rectangular boundary. We present the first polynomial time constant ratio approximation algorithm for the MLC-R problem, which has been shown to be NP-hard [5].

1 Introduction

An instance I of the MLC-R problem consists of a rectangle F partitioned into a set R of rectangles[1] (or rooms) $R_1, R_2, \ldots, R_r$. A *corridor* $T(I)$ for instance I is a set of connected line segments, each of which lies along the line segments that form F and/or the boundary of the rooms, and must include at least one point from every room and F. The objective of the MLC-R problem is to find a corridor of least total length. A generalization of the MLC-R problem where the rooms are rectilinear polygons is called the MLC problem.

The question as to whether the decision version of each of these problems is NP-complete is raised in [2, 3, 6]. Recently Gonzalez-Gutierrez and Gonzalez [5], and independently Bodlaender, et al. [1], proved that the MLC problem is strongly NP-complete. The reductions in these two papers are different. Gonzalez-Gutierrez and Gonzalez [5] also showed that the decision version of the MLC-R problem, as well as some of its variants, are strongly NP-complete.

When all the possible corridors must include the access point p located on the edges of F, the MLC-R problem is called the *p-access* point or simply the p-MLC-R problem. The decision version of this problem is shown to be NP-complete in [5]. Optimal corridors for the p-MLC-R, MLC-R and MLC problem instances are represented by thick line segments in Figure 1. The arrows represent problem restriction.

*On leave under the PROMEP program from the Universidad Autonoma de Queretaro. Supported in part by a Grant from UC MEXUS-CONACyT, Department of Computer Science, UC Santa Barbara, aglez@cs.ucsb.edu

†Department of Computer Science, UC Santa Barbara, teo@cs.ucsb.edu

[1]We assume that all the rectangles (and rectilinear polygons) are formed by horizontal and vertical line line segments.

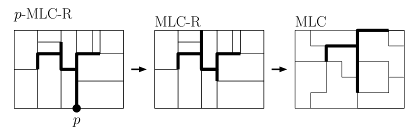

Figure 1: Restriction and generalization of the MLC-R problem.

In this paper we present the first provably polynomial time approximation algorithm for the MLC-R problem with approximation ratio 30. This ratio is twice the one for the case when the rooms are squares [1]. Our approximation algorithm restricts the solution space and the resulting problem, which remains NP-complete, is solved by relaxing the set of feasible solutions. Of course, the resulting solution space is different than the original one. The solution found to the restricted problem instance is rounded to obtain a solution to the original problem instance.

In Section 2 we discuss preliminary results and related problems. In Section 3 we present our parameterized algorithm for the p-MLC-R problem to approximate the MLC-R problem. In Section 4 we present our polynomial time constant ratio approximation algorithm for the p-MLC-R problem. In Section 5 we discuss results for related problems and open problems. For brevity we have omitted the proofs of all lemmas and theorems.

An application for the MLC problem is when laying optical fiber in metropolitan areas and every block (or set of blocks) is connected through its own gateway which may be located on the perimeter of each set of blocks. Our problems have also applications in VLSI and floorplanning when laying wires for clock signals or power, and wires for an electrical network or optical fiber for data communications. There has been recent research activity for related problems arising in spatial databases for trip planning queries.

2 Preliminaries and Related Problems

The point where a vertical and a horizontal edge of one or more rooms and/or F intersect is called a *vertex*. The set $V(R_i)$ of vertices includes all the vertices on the boundary of rectangle R_i. The set $V(F)$ is defined similarly. Using the vertices we may view any instance of the MLC-R problem as a graph. The solution to any instance of the MLC-R problem can be obtained by finding a solution to the p-MLC-R problem at each point $p \in V(F)$ and then selecting the best solution. Based on this claim in what follows we concentrate on developing an approximation algorithm for the p-MLC-R problem. Such an algorithm will also approximate the MLC-R problem within the same constant ratio.

The *Group Steiner Tree* (GST) problem may be viewed as a generalization of the MLC problem [7]. The input to the GST problem is a connected undirected edge-weighted graph $G = (V, E, w)$, where $w : E \rightarrow R^+$ is an edge-weight function; a non-empty set $C \subseteq V$, of *terminals*; and a partition $\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$ of C. The objective is to construct a tree $T(\mathcal{C}) = (V' \subseteq V, E' \subseteq E)$, such that at least one terminal from each set C_i is in $T(\mathcal{C})$ and the total edge-length $\sum_{e \in E'} w(e)$ is minimized. There is no known approximation algorithm for the GST problem that is a constant ratio approximation algorithm for the MLC problem.

Slavik [8, 9] defined the *Tree (Errand) Cover* (TEC) problem which can be seen as a generalization of the GST problem[2]. The TEC problem is more general than the GST problem in that the set $\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$ is not necessarily a partition of C, i.e., each $C_i \subseteq C$ denotes the set of vertices where errand e_i can be performed and the tree includes at least one vertex from each C_i. Slavik [8, 9] developed an approximation algorithm for the TEC problem with approximation ratio 2ρ, when each errand can be performed in at most ρ locations, i.e., $\rho = max\{|C_i|\}$.

The instance of the TEC problem corresponding to the instance (F, R, p) of the p-MLC-R problem is defined for the metric graph $G(V, E, w)$ constructed from (F, R, p) with the errand e_i for each rectangle R_i which can be performed at all the vertices $V(R_i)$, plus the errand e_0 located at vertex p. Every feasible solution for the p-MLC-R problem instance is also a feasible solution for the TEC problem instance and vice versa. Furthermore, the objective function value of every feasible solution to both problem instances is identical.

Let $T(I)$ be any corridor for instance I of the p-MLC-R problem and let $t(I)$ be its edge-length. Let $OPT(I)$ be an optimal corridor for instance I and let $opt(I)$ be its edge-length. An approach to generate suboptimal solution for the p-MLC-R problem is to construct an instance of the TEC problem and then invoke an approximation algorithm to solve it. The solution generated is the solution for the p-MLC-R problem. Currently one uses Slavik's [8, 9] approximation algorithm for the TEC problem. A direct application of this approach to the p-MLC-R problem generates a corridor whose total edge-length is at most $2\rho \cdot opt(I)$, where $\rho = max\{|V(R_i)|\}$ is not bounded above by a constant. In the next section we discuss our parameterized approximation algorithm based on the above approach.

[2]We should point out that in some papers the TEC problem is called the GST problem

3 Parameterized Algorithm

To establish a constant ratio approximation for the p-MLC-R problem, we restrict the solution space by limiting in each rectangle R_i the possible vertices, from which at least one must be part of the corridor. Consider the p-MLC-R$_S$ problem where S is a selector function that defines uniformly in each rectangle R_i a set of points, which are called the *critical points of* R_i. The objective of the p-MLC-R$_S$ problem is to find a minimum edge-length corridor that includes for each R_i at least one of its critical points. Let $k_S \geq 1$ be the maximum number of critical points selected by S from each R_i.

Given S and an instance I of the p-MCL-R problem, we use I_S to denote the instance of the corresponding p-MLC-R$_S$ problem. The instance of the TEC problem, denoted by J_S, is constructed from the instance I_S of the p-MLC-R$_S$ problem using the same approach as the one used for the p-MLC-R problem, but limiting the errands from each rectangle to the critical points of the rectangle. Every feasible solution for the p-MLC-R$_S$ problem instance I_S is also a feasible solution for the instance J_S of the TEC problem, and vice versa. Slavik's algorithm applied to J_S generates a solution $T(J_S)$ from which we construct a corridor $T(I)$ with edge-length $t(I)$ for the p-MLC-R problem. We call our approach the parameterized algorithm $Alg(S)$, where S is the parameter. Let $OPT(I_S)$ be an optimal corridor for I_S and let $opt(I_S)$ be its edge-length. Theorem 1 establishes the approximation ratio for $Alg(S)$. To use the following theorem one needs to prove that $opt(I_S) \leq r_S \cdot opt(I)$ for every instance I of the p-MLC-R problem, where $r_S \geq 1$.

Theorem 1 *Parameterized algorithm $Alg(S)$ generates for every instance I of the p-MLC-R problem a corridor $T(I)$ with length $t(I)$ at most $2k_S \cdot r_S$ times $opt(I)$, provided that $opt(I_S) \leq r_S \cdot opt(I)$.*

Instead of proving that $opt(I_S) \leq r_S \cdot opt(I)$, for every instance I of the p-MLC-R problem, one establishes a stronger result. That is, prove that for every corridor $T(I)$ with edge-length $t(I)$ there is a corridor $T(I_S)$ for I_S with edge-length $t(I_S) \leq r_S \cdot t(I)$. Applying this to $T(I) = OPT(I)$ we know that there is a corridor $T(I_S)$ such that $t(I_S) \leq r_S \cdot opt(I)$. Since $opt(I_S) \leq t(I_S)$, we know that $opt(I_S) \leq r_S \cdot opt(I)$. For the above approach to yield a constant ratio approximation algorithm we need both k_S and r_S to be bounded above by constants.

4 Selector Functions

The selector functions $S(4C)$, $S(F4C)$, and $S(Rk)$ corresponding to the selection of four corners, fewer than four corners, and k points selected randomly, respectively, do not result in constant ratio approximations for $Alg(S)$. The class of problems instances used to show this result have the property that every rectangle can be reached by using a function that selects another vertex called the *special point*.

Depending on the selector function S, a subset of the corner points $C(R_i)$ are called the *fixed points* $F(R_i)$ of rectangle R_i. Now, the set of critical points consists of the union of the pairwise-disjoint sets of fixed points and special points. Special points, in very general terms, are defined as follows. A *special point* is a vertex of R_i, that is not in $F(R_i)$, and the maximum edge-length needed to connect it to each one of the partial corridors in the set $\mathcal{A}$ of all partial corridors that do not include a vertex from rectangle R_i, but include points from all other rectangles (including rectangle R_0) is least possible. Finding special points in this way is in general time consuming. Also, this definition is not valid for all problem instances as the set of partial corridors that included vertices from all the rectangles except from R_i may be empty. We use however an upper bound for the maximum edge-length needed to connect a vertex of R_i to each one of the partial corridors to define special points. This can be defined by computing the tree of shortest paths originating at each vertex $V(R_i)$ [4].

The selector functions $S(+)$, $S(k+)$, and $S(2AC+)$ corresponding to the selection of one special point, k special points, two adjacent corners and a special point, respectively, do not result in constant ratio approximations for $Alg(S)$. The instances for which one can show the negative results suggest that if one includes two opposite corners results in near-optimal solutions to those problem instances. Therefore, we consider two opposite corners and one special point, and show in the next subsection that $Alg(S)$ is a constant ratio algorithm for the p-MLC-R, and therefore for the MLC-R problem.

4.1 Selecting two opposite corners and one special point

We analyze $Alg(S)$ when each set $V(R_i)$ is restricted to two opposite corners and one special point by $S(2OC+)$. Without loss of generality, we select the top-right and bottom-left corners as the two opposite corners. Clearly $k_{S(2OC+)} = 3$, and we prove that $r_{S(2OC+)} = 5$. By Theorem 1 this results in a parameterized algorithm having the approximation ratio 30. This result is established in Theorem 4. We take a top-down approach to prove this theorem. In Theorem 3, we show that given any corridor $T(I)$ there is a corridor $T(I_{S(2OC+)})$ with length $t(I_{S(2OC+)}) < 5 \cdot t(I)$. Applying the theorem to $OPT(I)$ we know there is a corridor $T(I_{S(2OC+)})$ with length $t(I_{S(2OC+)}) < 5 \cdot opt(I)$. Since $opt(I_{S(2OC+)}) \leq t(I_{S(2OC+)})$ it follows that $opt(I_{S(2OC+)}) < 5 \cdot opt(I)$. Applying Theorem 1, the result follows.

Consider the instance I of the p-MLC-R problem and its solution $T(I)$, given in Figure 2(a). Each rectangle results connected through at least one of its critical points except for the two (gray colored) rectangles indicated in the Figure 2(b). In order for the two (gray colored) rectangles to be connected, it is needed to extend $T(I)$. The instance $I_{S(2OC+)}$ of the p-MLC-R$_{S(2OC+)}$ problem and its solution $T(I_{S(2OC+)})$ are given in Figure 2(c). Clearly for this example, $t(I_{S(2OC+)})$ is less than $r_{S(2OC+)} = 5$ times $t(I)$.

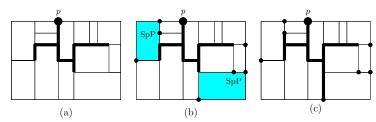

Figure 2: Constructing $T(I_{S(2OC+)})$ from $T(I)$.

Theorem 2 *For every instance I of the p-MLC-R problem, the parameterized algorithm $Alg(S(2OC+))$ generates, in polynomial time with respect to the number of rectangles r, a corridor with length at most $30 \cdot opt(I)$.*

We need to prove that given any corridor $T(I)$, it is possible to construct a corridor $T_S(I)$ with length $t_S(I) \leq 5t(I)$. We establish first some preliminaries and definitions. All the rectangles that do not have a critical point along the corridor $T(I)$ are called *no-critical-point-exposed* (*ncpe*) rectangles. The remaining rectangles are called *critical-point exposed* (*cpe*) rectangles. Notice that the definition of the ncpe rectangles depends on the points selected uniformly by S.

Let τ be the counter-clockwise tour of corridor $T(I)$. This tour starts at the access point p moving in the counter-clockwise direction with respect to p on the exterior of F. Lets mark τ at the first point where each ncpe rectangle is visited for the first time. The order in which the ncpe rectangles are first visited is called the *canonical order* of the ncpe rectangles. Rename the ncpe rectangles according to this canonical order as $VR = \{n_1, n_2, \ldots, n_q\}$. For $1 \leq i \leq q$, we define the *reaching point* ψ_i of the ncpe rectangle n_i as the first point in the ncpe rectangle n_i visited by τ; and the *leaving point* χ_i of the ncpe rectangle n_i as the last point in the ncpe rectangle n_i visited by τ before it visits the reaching point ψ_{i+1} of rectangle n_{i+1}. For $0 \leq i \leq q$, let l_i be the length of the path $\tau(\chi_i, \psi_{i+1})$. Let h_j be the length of the path $\tau(\psi_j, \chi_j)$, for $1 \leq j \leq q$.

By using the shortest path along the corridor $T(I)$ from n_i to n_{i+1}, we define the *exit point* X_i of ncpe n_i and the *entry point* Y_{i+1} of n_{i+1} as the intersection of the shortest path with the edges of n_i and n_{i+1}, respectively. For $i < j$, $T(X_i, Y_j)$ is the shortest path along the corridor $T(I)$ from the exit point X_i to the entry point Y_j, and $t(X_i, Y_j)$ denotes its total edge-length.

The idea is to show that for each ncpe rectangle n_i one can connect one of its critical points to the corridor by adding line segments of length at most $l_{i-1}+h_i+l_i$. We establish this claim in Lemma 4. In Theorem 3 we show this is enough in order to establish our approximation bound.

Theorem 3 *Given any corridor $T(I)$ for any instance I of the p-MLC-R problem, there is a corridor for instance $I_{S(2OC+)}$ of the p-MLC-$R_{S(2OC+)}$ problem with edge-length $t(I_{S(2OC+)}) < 5 \cdot t(I)$.*

To establish our main result we need Lemma 4.

Lemma 4 *For $1 \leq i \leq q$, it is possible to connect at least one of the critical points of every ncpe rectangle $n_i \in VR$ to the corridor, by adding line segments of length at most $l_{i-1} + h_i + l_i$.*

The proof of lemma 4 is based on the following general approach. We define a special type of turns in paths, which is called an *inversion of direction* (or simply iod-subpaths) of a path $T(X_i, Y_j)$. Iod-subpaths are vertical or horizontal. We then define the beam of an iod-subpath as the set of points that are "visible" from an iod-subpath. A path is said to be of type-1 if it has vertical and horizontal iod-subpaths, otherwise it is said to be type-2. We show that if a path has a vertical and a horizontal iod-subpath, then it has two *adjacent* ones whose beams intersect. When this condition holds, we can show that there is at least one rectangle in R that is completely inside the "region" of the adjacent iod-subpaths. This latter property is used to establish that from either end of the path that includes the adjacent iod-subpaths, or from a point which we call the bifurcation point, one can reach any point in one of the rectangles located completely inside the region of the iod-subpath by using line segments of length at most $t(X_i, Y_j)$. Then we use this fact to show that the length of the path from a vertex of n_i to a rectangle in R is at most $l_i+h_i+l_{i+1}$, and by using the definition of a special point, we show the special point of n_i can be connected to the corridor by using line segments of length at most $l_i + h_i + l_{i+1}$.

So if path $T(X_{i-1}, Y_i)$ or $T(X_{i-1}, Y_{i+1})$ is type-1, we know it is possible to connect the critical point of n_i to the corridor by adding line segments of length at most $l_i + h_i + l_{i+1}$. However, it may be that both paths ($T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$) are type-2. In this case we need to characterize the structure of paths $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$. Then we use this characterization to show that the path $T(Y_i, Y_{i+1})$ is type-1. Then by using arguments similar to the ones above, one can show that the special point of n_i can be connected to the corridor by using line segments of length at most $l_i + h_i + l_{i+1}$. Because of page limitations we cannot include additional details.

5 Additional Results and Discussion

Our approximation algorithm can be adapted to the MLC_k problem, i.e., when F is partitioned into c-gons, for $c \leq k$, resulting in a constant ratio approximation algorithm, when k is bounded above by a constant.

We have adapted our algorithm (with the same constant ratio) to restricted versions of the MLC problem, but so far we have not been able to adapt it to all cases. The existence of a constant ratio approximation algorithm for the MLC problem remains a challenging open problem. An equally challenging problem is to develop approximation algorithm for the MLC-R with a significant smaller approximation ratio, for example two.

Another interesting problem is the group-TSP problem when restricted to rectangles as in the MLC-R problem, which we call the *rectangular group-TSP*. We claim that a parameterized algorithm similar to the one presented in this paper that uses the selector $S(2OC+)$ generates a solution to the group-TSP and the algorithm has approximation ratio 60. However a better algorithm for this problem with approximation ratio 22.5 [4] exists based on Slavik's approximation algorithm for the *Errand Scheduling* problem [8, 9], rather than on the approximation algorithm for the TEC problem.

References

[1] H. L. Bodlaender, C. Feremans, A. Grigoriev, E. Penninkx, R. Sitters, and T. Wolle. On the minimum corridor connection and other generalized geometric problems. In *Proc. 4th. WAOA*, vol. 4368 of *LNCS*, pages 69–82, 2006.

[2] E. D. Demaine and J. O'Rourke. Open problems from CCCG 2000. In *Proc. 13th CCCG*, pages 185–187, 2001.

[3] D. Eppstein. Some open problems in graph theory and computational geometry. ICS Dept., UCI, Nov., 2001.

[4] A. Gonzalez-Gutierrez and T. F. Gonzalez. Approximation algorithms for the minimum-length corridor and related problems. Technical Report TR UCSB 2007-03, Department of Computer Science, UCSB, 2007.

[5] A. Gonzalez-Gutierrez and T. F. Gonzalez. Complexity of the minimum-length corridor problem. *Comput. Geom. Theory Appl.*, 37(2):72–103, 2007.

[6] L. Y. Jin and O. W. Chong. The minimum touching tree problem. School of Computing, National University of Singapore, 2003.

[7] G. Reich and P. Widmayer. Beyond Steiner's problem: a VLSI oriented generalization. In *Proc. of WG '89: LNCS*, pages 196–210.

[8] P. Slavik. The errand scheduling problem. Technical Report 97-02, CSE Deparment, University of New York at Buffalo, 1997.

[9] P. Slavik. Approximation Algorithms for Set Cover and Related Problems. Ph. D. Thesis 98-06, CSE Deparment, University of New York at Buffalo, 1998.