Finding Rectilinear Least Cost Paths in the Presence of Convex Polygonal Congested Regions * †

Avijit Sarkar

School of Business, University of Redlands
1200 East Colton Avenue, Redlands, CA 92373, USA

Rajan Batta and Rakesh Nagi

Department of Industrial and Systems Engineering, 438 Bell Hall
University at Buffalo (SUNY), Buffalo, NY 14260, USA

July 2007

Abstract

This paper considers the problem of finding the least cost rectilinear distance path in the presence of convex polygonal congested regions. We demonstrate that there are a finite, though exponential number of potential staircase least cost paths between a specified pair of origin-destination points. An upper bound for the number of entry/exit points of a rectilinear path between two points specified *a priori* in the presence of a congested region is obtained. Based on this key finding, a "memory-based probing algorithm" is proposed for the problem and computational experience for various problem instances is reported. A special case where polynomial time solutions can be obtained has also been outlined.

Keywords: Least Cost Path, Congested Regions, Rectilinear Distance Metric.

1 Introduction

The problem of finding the rectilinear least cost path between two points on a plane is trivially solved. There is an infinite number of alternate optima except when the points share abscissa or y-coordinate. However, the situation is more complicated in the presence of obstacles or barriers, and originates in many fields, such as wire routing in VLSI, motion/path planning in robotics, plant and facility layout, urban planning and transportation, geographical information systems, etc. It has been extensively studied in the past. Lee, Yang and Wong [19, 11] emphasized its application to VLSI design and presented efficient algorithms for finding rectilinear collision-free paths between two given points in the presence of rectilinear obstacles. Guha and Suzuki [8] is a study of proximity problems for points on a rectilinear plane with rectangular obstacles.

*Corresponding Author: Avijit Sarkar. E-mail: avijit_ sarkar@redlands.edu

†Submitted to *Computers and Operations Research*

Schuierer [18] presented a data structure to preprocess a rectangular polygon such that the shortest distance between two query points can be determined employing the rectilinear distance metric. Larson and Li [9] proposed an efficient algorithm for determining the shortest rectilinear path between two points in $\Re^2$ in the presence of arbitrarily shaped polygonal barriers. The algorithm of [9] runs in $O(m(m^2 + n^2))$ time where n is the total number of vertices of the barriers and m is the number of origin-destination points. Algorithms for the same problem were also developed by Clarkson, Kapoor and Vaidya [6] and Mitchell [13]. A related problem is that using the Euclidean distance. This is considered by Mitchell and Papadimitriou [14].

A barrier which can be penetrated but through which travel is at a higher cost is referred to in the literature as a congested region. Butt and Cavalier [3] considered the rectilinear p-median problem in the presence of convex polygonal ***congested region***s. They claimed that at least one least cost path will always coincide with segments of the grid obtained by drawing horizontal and vertical lines through the existing facilities and the vertices of the congested regions. Sarkar, Batta and Nagi [16] established with the aid of a counter-example that the grid structure proposed by Butt and Cavalier to solve the rectilinear p-median problem is inadequate. They also proved that Butt and Cavalier's grid structure suffices for rectangular congested regions. However a generalized solution methodology to determine the least cost path between two points in the presence of non-rectangular congested regions still does not exist. In [17], Sarkar, Batta and Nagi addressed the problem of placing a new department of given area but unknown dimensions in a shop floor layout in the presence of other departments, where all departments are modeled as rectangular congested regions. This paper proposes a "memory based probing procedure" to determine the least cost path between a pair of origin-destination points in the presence of **convex polygonal** congested regions.

The remainder of this paper is organized as follows. Section 2 describes some preliminaries to the problem. Section 3 defines the problem. Section 4 demonstrates the existence of an exponential number of potential least cost paths. In Section 5 we establish an upper bound for the number of entry/exit points for a single convex polygonal congested region. Based on these results, a "memory-based probing" solution algorithm is proposed in Section 6. Computational results are reported in Section 7. In addition a special case of the problem has been outlined where it is possible to obtain polynomial time solutions. Finally, conclusions and directions for future research are outlined in Section 8.

2 Preliminaries

2.1 Larson and Li's [9] probing approach

Larson and Li developed an efficient algorithm to determine the minimum rectilinear distance path between a given set of origin-destination points in the presence of polygonal barriers. They defined the origin-destination points and the barrier vertices as nodes. From every node, the authors constructed a "vertex seeking tree (VST)" that seeks all other vertices which "communicate" with the start node and contains the minimal length path from the start node to each such vertex. A VST, rooted at an origin-destination point (node) D having coordinates (x, y), is the union of four "probes" originating from D, in the $+x$, $-x$, $+y$ and $-y$ travel directions. They defined the following:

1. A "positive x probe" originating from D is a horizontal line, drawn from D in the $+x$ direction. A "positive x probe" of D is denoted as $r_{x^+}(D)$.

2. A "negative x probe" originating from D is a horizontal line, drawn from D in the $-x$ direction. A "negative x probe" of D is denoted as $r_{x^-}(D)$.

3. A "positive y probe" originating from D is a vertical line, drawn from D in the $+y$ direction. A "positive y probe" of D is denoted as $r_{y^+}(D)$.

4. A "negative y probe" originating from D is a vertical line, drawn from D in the $-y$ direction. A "negative y probe" of D is denoted as $r_{y^-}(D)$.

There are three possibilities for each of these probes. They are illustrated in Figure 2 of [9].

1. A probe intersects no barrier, i.e., proceeds to infinity and becomes a ray.

2. A probe intersects another node, upon which it terminates.

3. A probe intersects an edge of a barrier partitioning the edge into two sub-edges. From the point of intersection, the probe proceeds along the sub-edge that lies at an obtuse angle to the incident probe and terminates at the end-point (a barrier vertex or a node) of the sub-edge. Clearly in this case, the probe itself is the union of a vertical (or a horizontal) line segment and the sub-edge that lies at an obtuse angle. It is pertinent to add that if the probe is incident on an edge at a right angle, then the probe terminates.

2.2 Need for "memory-based" probing

In our problem, the endeavor is to determine a rectilinear path of least cost between origin-destination points in the presence of congested regions. Since congested regions allow travel through, a probe, on intersecting an edge of a congested region, will pass through the region. In doing so, the probe will generate an entry point (point at which the probe enters the region) and an exit point (point at which the probe leaves the region). Each such entry and exit point can potentially generate three probes, one in each travel direction apart from the direction of the probe that generated the point. The same is true for any following entry/exit point, potentially resulting in an exponential number of probes. For ease of understanding, let us consider the example illustrated in Figure 1. The first entry point of $r_{y^+}(O)$ to the congested region is E_1. Probing "without memory" will result in the entry/exit points E_2, E_3, E_4, etc. Note that the probe originating from E_1 is a positive y probe, while the probe originating from E_3 is a negative y probe. Consequently, as depicted in Figure 1, an infinite sequence of turning steps (a turning step is defined in Larson and Li as a probe directly connecting two horizontal or two vertical probes whose directions of travel are reversed) may result. Hence a simple "barrier extension" of Larson and Li's vertex seeking tree approach will not suffice for our problem.

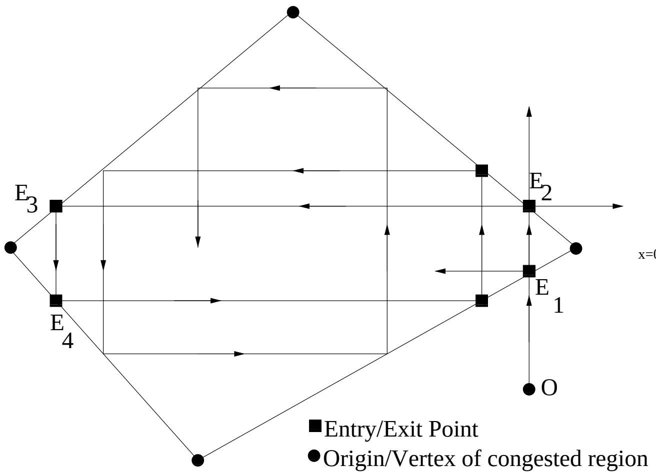

Figure 1: Memory-less probing

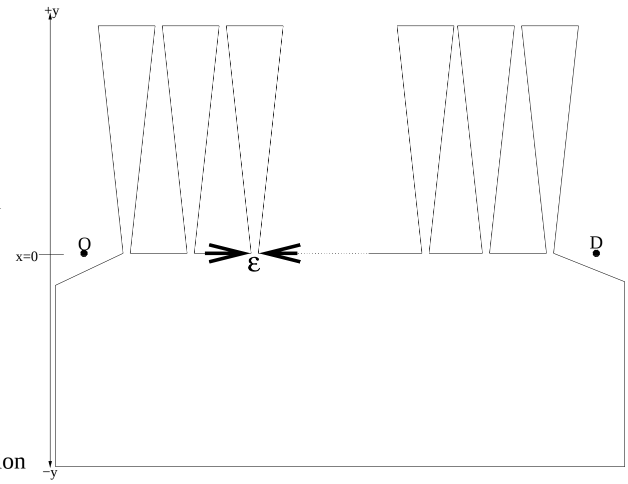

Figure 2: Arbitrarily shaped congested region

2.3 Need for convex polygonal restriction

Figure 2 illustrates a single non-convex congested region with congestion factor $\alpha \to 0$ and a pair of origin-destination points O and D lying on $x = 0$. By increasing the height of the congested region in both the $+y$ and $-y$ directions, we can ensure that the least cost path between O and D will never bypass the congested region. Among paths that penetrate the congested region, the one with least cost is the straight line joining O and D along $x = 0$, at a cost of $d(O, D) + M\epsilon\alpha$, where $d(O, D)$ is the length of the shortest rectilinear path between O and D, ϵ is the width of each wedge as shown in Figure 2 and M, a very large number, denotes the number of entry/exit points. If we let $\alpha \to 0$, $\epsilon \to 0$ and $M \to \infty$ under the condition that $M\epsilon\alpha \to 0$, we get $d(O, D) + M\epsilon\alpha \to d(O, D)$. Thus we have a limiting example where the least cost path between O and D can enter/exit a congested region an arbitrarily large number of times. To overcome this difficulty, we consider **convex polygonal** congested regions in the remainder of this paper (except in Section 7.4 when a special case of our problem is discussed). A convex shape helps limit the number of entry/exits. A polygonal restriction helps us to identify the entry/exit points.

We proceed in our analysis by:

1. Determining an upper bound on the number of entry/exit points for a single congested region, and

2. Associating memory with the probes such that a path never contains a "turning step".

3 Problem Definition

Let G_i (an open set) denote the set of points $(x, y) \in \Re^2$ contained strictly within the i^{th} congested region. We also define $\overline{G}_i = G_i \cup \{\text{boundary of congested region } i\}$, a closed convex polygonal set. The set $\overline{G}_i$ represents the i^{th} congested region. Let $v(m, i)$ denote the m^{th} clockwise ordered vertex of congested region $\overline{G}_i$, with coordinates $(x(m, i), y(m, i))$. Each congested region is characterized by a finite number of such vertices through which vertical and/or horizontal lines of tangency can be passed. Traveling through the interior of a congested region is permitted at an additional cost per unit distance. This additional cost per unit distance for traveling in congested region $\overline{G}_i$ is denoted by α_i, $0 \leq \alpha_i < \infty$, and is referred to as its congestion factor. Thus if w is the cost of travel per unit distance between two points lying outside or on the boundary of $\overline{G}_i$, then the cost of travel (between the same points) along a path entirely contained in G_i would be $(1 + \alpha_i)w$. It is pertinent to note that there is no congestion along the boundary of a congested region: hence the distinction between the inside and the boundary of a congested region. Note that the boundary of a congested region can be written as $\overline{G}_i \setminus G_i$. We assume that there are p non-intersecting congested regions. Thus $\overline{G}_i \cap \overline{G}_{i'} = \emptyset \quad \forall i, i' = 1, \ldots, p, i \neq i'$. We let $\overline{G} = \cup_i \overline{G}_i$.

Consider a pair of origin-destination points O and D in $\Re^2 \setminus \overline{G}$. The objective is to find the least cost path between O and D in the presence of p convex polygonal congested regions. Note that the least cost path between any such pair of points may want to bypass a congested region $\overline{G}_i$ or enter it depending on the value of α_i.

4 Initial Insights

In this section, we state two observations which provide substantial insight into the nature of the problem. However, before doing so, we present a few definitions from Larson and Li, suitably adapted for our purposes and some facts about least cost paths.

4.1 Least Cost Paths

A *staircase* path between $O = (x_o, y_o)$ and $D = (x_d, y_d)$ is a rectilinear path having length $|x_o - x_d| + |y_o - y_d|$. Two points are said to *communicate* if there exists at least one feasible staircase path between them. Here the term feasible implies a path that does not contain a point belonging to the interior of a congested region.

We now present some cases for a path between O and D:

1. If O and D communicate, the least cost path between O and D will never enter the interior of a congested region.

2. The shortest length path between O and D may not necessarily be the least cost path between O and D. This is illustrated later by examples in Figures 4(c) and 4(d).

3. Two paths of equal length between O and D may not necessarily have equal cost. In fact, we later observe (Observation 2) that among two paths of equal length, the path with lesser cumulative weighted travel (product of the rectilinear travel distance inside a congested region and the congestion factor of the congested region) is the path of smaller cost.

4. The least cost path between O and D may enter and exit a congested region more than once. This is illustrated later by an example in Figure 4(b). When α_i becomes prohibitively high, the least cost path may bypass the congested region.

4.2 Observations

Observation 1: *In the presence of p congested regions in $\Re^2$, there are in the worst case 2^{2p+2} uniquely different candidate staircase least cost paths between an origin and a destination.*

The observation is based on the construction of 2^{2p+2} unique staircase paths between an origin O and a destination D in the presence of p congested regions.

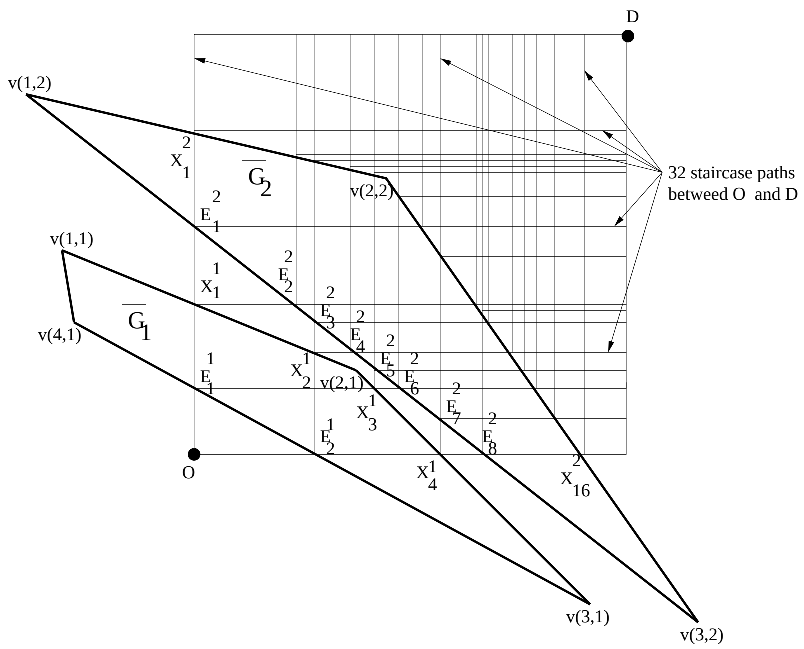

Figure 3: Exponential growth in number of staircase paths between O and D

We start by considering two congested regions $\overline{G}_1$ and $\overline{G}_2$ as depicted in Figure 3. Let $(x(m,i), y(m,i))$ denote the coordinates of the m^{th} clockwise ordered vertex $v(m,i)$ of congested region $\overline{G}_i$. $\overline{G}_1$ and $\overline{G}_2$ are constructed such that $x(1,1) < x(4,1) < x(2,1) < x(3,1)$; $y(3,1) < y(2,1) < y(4,1) < y(1,1)$ for $\overline{G}_1$ and $x(1,2) < x(2,2) < x(3,2)$; $y(3,2) < y(2,2) < y(1,2)$ for $\overline{G}_2$. We denote an entry point of a path passing through the interior of $\overline{G}_i$ as E^i_s (using increasing subscript in decreasing order of y-coordinates of the entry points). From each such entry point, the path splits in $+y$ and $+x$ travel directions and exit points

X_s^i (again using increasing subscript in decreasing order of y-coordinates of the exit points) are generated. For example, in Figure 3, the entry points (on $\overline{G}_1$) are E_1^1 and E_2^1 whereas the exit points are X_1^1, X_2^1, X_3^1 and X_4^1. From each exit point, the paths again split in $+y$ and $+x$ travel directions. Note that the paths (from O) are constructed according to the probing procedure described in Section 6.1. Due to the way $\overline{G}_2$ is constructed, (eight) probes from all (four) exit points of $\overline{G}_1$ are incident on $\overline{G}_2$. Consequently there are sixteen exit points from $\overline{G}_2$.

Clearly, a third congested region $\overline{G}_3$ can be added such that all 32 probes exiting $\overline{G}_2$ are incident on $\overline{G}_3$ (and there are 128 staircase paths from O to D); a fourth congested region $\overline{G}_3$ can be added such that all 128 probes exiting $\overline{G}_3$ are incident on $\overline{G_4}$ (and there are 512 staircase paths from O to D) and so on. Hence, in the worst case, there could possibly be 2^{2p+1} uniquely different staircase paths from O to D that enter and exit each congested region exactly once, where p is the number of congested regions that cause O and D to not communicate.

Analogously in the worst case, 2^{2p+1} uniquely different staircase paths (that enter and exit each congested region exactly once) can be constructed by probing from D back to O. Note that during backward probing (from D to O), some of the probes generated may intersect entry/exit points that were previously generated during forward probing (from O to D). In such cases, the backward probes are terminated at those entry/exit points. It is pertinent to mention here that 2-entry 2-exit staircase paths can be captured by the union of these two sets of 2^{2p+1} uniquely different staircase paths which result due to forward and backward probing. For an example of a 2-entry 2-exit staircase path, we refer the reader to the example illustrated in Figure 4(b). Also note that if the two sets of 2^{2p+1} uniquely different staircase paths do not intersect, the total number of paths is additive; hence the 2^{2p+2} uniquely different candidate staircase least cost paths between O and D.

Finally, a staircase path that enters and exits a convex polygonal congested region more than two times cannot exist (proved later in Theorem 5.1).

Observation 2: *Among all staircase paths P_{OD} of length l_{OD} between origin-destination points O and D that pass through congested regions $\overline{G}_i$, $i = 1, \ldots, p$ with congestion factors α_i, the least cost path is the one that minimizes $\sum_{i=1}^{p} \phi_i \alpha_i$, where ϕ_i is the length of a portion of P_{OD} in the interior of $\overline{G}_i$.*

Without loss of generality, let $w = 1$. The cost incurred to travel along path P_{OD} is $\sum_{i=1}^{p} \phi_i(1+\alpha_i) + \delta$, where δ is the total length of portions of P_{OD} that lie outside the congested regions. Since $\sum_{i=1}^{p} \phi_i(1+\alpha_i) + \delta = \sum_{i=1}^{p} \phi_i \alpha_i + l_{OD}$, the observation is valid.

From Observation 1, at most 2^{2p+2} uniquely different candidate staircase least cost paths can be constructed, i.e., at most $O(4^p)$ entry/exit points may be generated between a given origin j and destination k. However, due to Observation 2, none of these paths can be pruned *a priori*. Note that the *staircase* least cost path P^*_{jk} is the one that minimizes $\sum_{i=1}^{p} \phi_i \alpha_i$, which can be interpreted as a weighted sum of the distances traversed in the interior of congested region $\overline{G}_i$, α_i's being the weights. Let us consider, for the sake of argument, a dynamic programming-like approach, that selects the best exit point E_b (among all exit points) for the first congested region intersected and prunes all other exit points. Now consider such a pruned exit point E_{pr}. The weighted travel of a *staircase* path through E_{pr} may be less than the corresponding travel through E_b; because we do not have *a priori* knowledge of the exact amount of "interior" travel of staircase paths through subsequent entry/exit points, lying on the *staircase* path passing through E_{pr}. Hence an "*a priori* pruning procedure" (of paths) cannot be designed for our problem. Consequently, there could be at most $O(4^p)$ candidate *staircase* least cost paths between a given origin and destination.

The remainder of the paper focuses on the development of an algorithm (entitled "memory-based probing algorithm", described later in Section 6) to determine the least cost path between a pair of origin-destination points in the presence of convex polygonal congested regions. It is pertinent to note here that a least cost path could be non-staircase too (note examples in Section 5). To this end, the following section establishes an upper bound on the number of entry/exit points of a rectilinear (staircase or non-staircase) path between two points for a single convex polygonal congested region.

5 Establishment of Maximum Number of Entry-Exit Points

In this section, we demonstrate, with the aid of some examples (in Figure 4), that a least cost rectilinear path between two points specified *a priori* may enter and exit a convex polygonal congested region up to two times. Then we prove that a least cost rectilinear path cannot enter and exit a convex polygonal congested region three or more times because travel along a 1-entry 1-exit path and a 2-entry 2-exit path cannot simultaneously cost more than a 3-entry 3-exit path. Determining an upper bound on the number of entry/exit points is crucial to the development of our solution methodology because it prohibits the formation of turning steps in the memory-based probing procedure. Let $v(m)$ denote the m^{th} clockwise ordered vertex of congested region $\overline{G}_1$.

1. In Figure 4(a), the least cost path between A and B does not enter congested region $\overline{G}_1$ because A and B communicate simply.

2. In Figure 4(a), the least cost path between O and D is OE_1E_4D. It enters and exits the congested region $\overline{G}_1$ once, at E_1 and E_4 respectively, for $0 < \alpha \leq 1.33$. For $\alpha > 1.33$, the least cost path Ov_4v_1D bypasses $\overline{G}_1$.

3. In Figure 4(b), the least cost path between O and D (now at $(12.5, 10)$) is $OE_1E_2E_3E_4D$. It enters and exits $\overline{G}_1$ two times for $0 < \alpha \leq 0.59$. The first entry and exit points are E_1 and E_2 respectively and the second entry and exit points are E_3 and E_4 respectively.

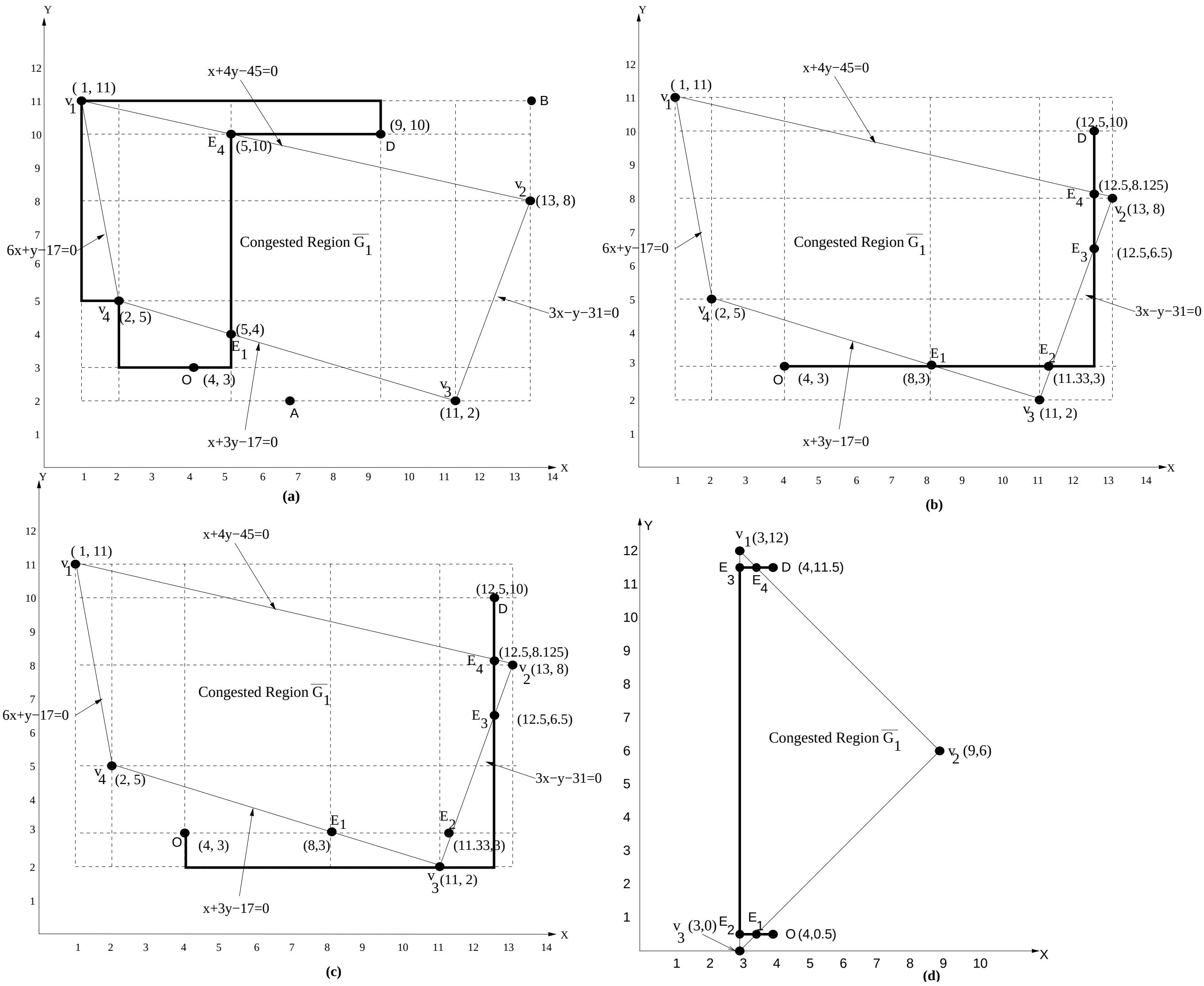

Figure 4: Examples of 1-entry/exit and 2-entry/exits

4. However for $\alpha = 0.61$, the least cost path between O and D is $Ov_3E_3E_4D$. It enters and exits $\overline{G}_1$ once, at E_3 and E_4 respectively. Note that in this case, the least cost path between O and D is no longer the path of shortest length between O and D. This is illustrated in Figure 4(c).

5. In Figure 4(d), the least cost path between O and D is $OE_1E_2E_3E_4D$. It enters and exits $\overline{G}_1$ two times for $\alpha = 0.30$. The first entry and exit points are E_1 and E_2 respectively and the second entry and exit points are E_3 and E_4 respectively. Note that in this case also, the least cost path between O and D is no longer the path of shortest length between O and D.

We now prove that a rectilinear path between a pair of origin-destination points cannot enter and exit a convex polygonal congested region three or more times. Consider a convex polygonal congested region $\overline{G}_1$, as shown in Figure 5. Let Λ be the set of points contained strictly within $\overline{G}_1$. Let $v(m)$, $m = 1, \ldots, q$, denote the m^{th} clockwise ordered vertex of $\overline{G}_1$. Note that since we consider one congested region, we use the notation $v(m)$ instead of $v(m, i)$, as introduced earlier in Section 3. Also, let $v(1)$, $v(r)$, $v(b)$ and $v(l)$ denote the vertices of $\overline{G}_1$ with the highest y-coordinate, highest x-coordinate, lowest y-coordinate and lowest x-coordinate respectively. Clearly $l \geq b \geq r \geq 1$. We denote as $\beta(m)$ the "vertex angle" of $v(m)$. We now define the "vertex angle" and state some observations simultaneously:

1. Definition: The "vertex angles" $\beta_1(1)$ and $\beta_2(1)$ of $v(1)$ are the angles between a horizontal line drawn through $v(1)$ and the edges $\overline{v(1)v(2)}$ and $\overline{v(1)v(q)}$ respectively.

2. Definition: The "vertex angles" $\beta_1(b)$ and $\beta_2(b)$ of $v(b)$ are the angles between a horizontal line drawn through $v(b)$ and the edges $\overline{v(b)v(b-1)}$ and $\overline{v(b)v(b+1)}$ respectively.

3. Observation: Vertices $v(1)$ and $v(b)$ have two edge angles each, denoted as $(\beta_1(1), \beta_2(1))$ and $(\beta_1(b), \beta_2(b))$ respectively, as mentioned earlier in items 1 and 2.

4. Observation: At most one of the angles $\beta_1(1)$, $\beta_2(1)$, $\beta_1(b)$ and $\beta_2(b)$ can be obtuse in order to maintain convexity. However in such cases, it can be trivially established that a feasible rectilinear path that enters and exits $\overline{G}_1$ exactly three times simply cannot be constructed.

5. Definition: For $2 \leq m \leq r-1$, the "vertex angle" $\beta(m)$ of $v(m)$ is the acute angle between a horizontal drawn through $v(m)$ and the edge $\overline{v(m)v(m+1)}$. Observe that $\beta_1(1) < \beta(2) < \beta(m) < \beta(r-1)$.

6. Definition: For $r+1 \leq m \leq b-1$, the "vertex angle" $\beta(m)$ of $v(m)$ is the acute angle between a horizontal drawn through $v(m)$ and the edge $\overline{v(m-1)v(m)}$. Observe that $\beta(r+1) > \beta(m) > \beta(b-1) > \beta_1(b)$.

7. Definition: For $b+1 \leq m \leq l-1$, the "vertex angle" $\beta(m)$ of $v(m)$ is the acute angle between a horizontal drawn through $v(m)$ and the edge $\overline{v(m)v(m+1)}$. Observe that $\beta_2(b) < \beta(b+1) < \beta(m) < \beta(l-1)$.

8. Definition: For $l+1 \leq m \leq q$, the "vertex angle" $\beta(m)$ of $v(m)$ is the acute angle between a horizontal drawn through $v(m)$ and the edge $\overline{v(m-1)v(m)}$. Observe that $\beta(l+1) > \beta(m) > \beta(q) > \beta_2(1)$.

9. Observation: None of the angles $\beta(m)$ defined in points 5, 6, 7, and 8 above can be obtuse due to the convexity assumption.

10. Observation: The edge angles of $v(l)$ and $v(r)$ are undefined.

Let $\gamma(m) = 90° - \beta(m)$. We define $E(\overline{G}_1)$ to be the smallest rectangle that encloses $\overline{G}_1$. Let Δ be the set of points contained strictly within $E(\overline{G}_1)$. Consider an origin O and a destination D, such that $O, D \in \Delta - \Lambda$, and such that O, D do not communicate.

Theorem 5.1. *A least cost rectilinear path between a pair of origin-destination points that enters and exits a convex polygonal congested region three or more times does not exist.*

Proof: Without loss of generality, let us assume that the cost of travel per unit distance, $w = 1$.

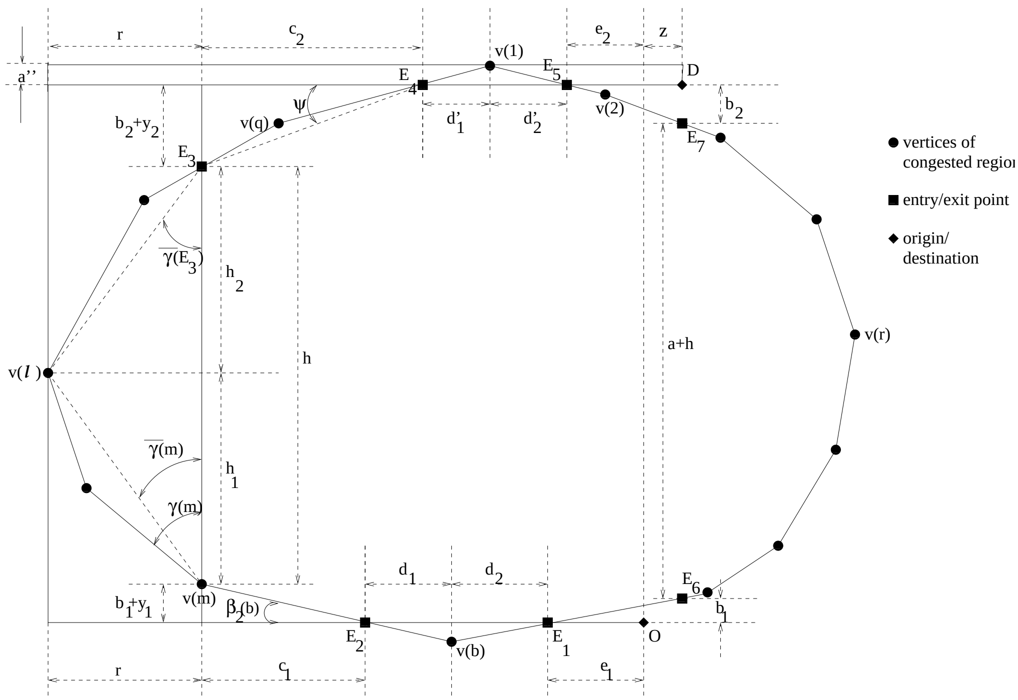

Figure 5: Figure for Proof of Theorem 5.1

Stage 1: Stating four paths under consideration

For the purpose of argument, consider the rectilinear path $OE_1E_2v(m)E_3E_4E_5D$ that enters and exits the congested region exactly three times. We prove later that such a path cannot exist. Without loss of generality, the first, second and third entry points are E_1, $v(m)$ and E_4 respectively, whereas the corresponding exit points are E_2, E_3 and E_5. Note that the path is non-staircase due to the way O and D are located. Cost (of travel) along the path is $e_1 + e_2 + c_1 + c_2 + b_1 + y_1 + b_2 + y_2 + z + (h + d_1 + d_2 + d_1' + d_2')(1 + \alpha)$, where e_1, e_2, c_1, c_2, b_1, y_1, b_2, y_2, z, h, d_1, d_2, d_1', d_2' are as shown in Figure 5 and α is the congestion factor of $\overline{G}_1$.

$OE_1E_2v(l)E_4E_5D$ is a rectilinear path that enters and exits the congested region exactly two times. Without loss of generality, the first and second entry points are E_1 and E_4 whereas the corresponding exit points are E_2 and E_5 respectively. Note that this path too is non-staircase. Cost of the path is $e_1 + e_2 + c_1 + c_2 + 2r + b_1 + y_1 + b_2 + y_2 + h + z + (d_1 + d_2 + d_1' + d_2')(1 + \alpha)$.

OE_6E_7D is a rectilinear staircase path that enters and exits the congested region exactly once. Without loss of generality, the entry and exit points are E_6 and E_7 respectively. Cost of the path is $b_1 + b_2 + z + (a + h)(1 + \alpha)$.

$OE_1E_2v(l)v(1)D$ is a rectilinear non-staircase path that enters and exits the congested region exactly once. Without loss of generality, the entry and exit points are E_1 and E_2 respectively. Cost of the path is $e_1 + e_2 + c_1 + c_2 + 2r + 2a'' + h + b_1 + y_1 + b_2 + y_2 + z + d_1' + d_2' + (d_1 + d_2)(1 + \alpha)$.

Stage 2: Comparing 3 entry-exit with staircase 1 entry-exit (Case A)

For the 3-entry 3-exit path to be cheaper than the staircase 1-entry 1-exit path, the following must hold: $e_1+e_2+c_1+c_2+b_1+y_1+b_2+y_2+z+(h+d_1+d_2+d_1'+d_2')(1+\alpha) < b_1+b_2+z+(a+h)(1+\alpha)$.

On simplification, we get

$$\alpha > \frac{e_1 + e_2 + c_1 + c_2 + d_1 + d_2 + d_1' + d_2'}{a - d_1 - d_2 - d_1' - d_2'}. \tag{1}$$

Stage 3: Comparing 3 entry-exit with 2 entry-exit (Case B)

For the 3-entry 3-exit path to be cheaper than the 2-entry 2-exit path, the following must hold:

$e_1 + e_2 + c_1 + c_2 + b_1 + y_1 + b_2 + y_2 + z + (h + d_1 + d_2 + d_1' + d_2')(1 + \alpha) < e_1 + e_2 + c_1 + c_2 + 2r + b_1 + y_1 + b_2 + y_2 + h + z + (d_1 + d_2 + d_1' + d_2')(1 + \alpha)$.

On simplification, we get

$$\alpha < \frac{2r}{h}. \tag{2}$$

From Figure 5, $h = h_1 + h_2$, $\frac{r}{h_1} = \tan\overline{\gamma}(m)$, $\frac{r}{h_2} = \tan\overline{\gamma}(E_3)$.

Hence, substituting h in (2), we get

$$\alpha < \frac{2}{\cot\overline{\gamma}(m) + \cot\overline{\gamma}(E_3)}. \tag{3}$$

Stage 4: Comparing 3 entry-exit with non-staircase 1 entry-exit (Case C)

For the 3-entry 3-exit path to be cheaper than the non-staircase 1-entry 1-exit path, the following must hold:

$e_1 + e_2 + c_1 + c_2 + b_1 + y_1 + b_2 + y_2 + z + (h + d_1 + d_2 + d_1' + d_2')(1 + \alpha) < e_1 + e_2 + c_1 + c_2 + 2r + 2a'' + h + b_1 + y_1 + b_2 + y_2 + z + (d_1 + d_2)(1 + \alpha)$.

On simplification, we get

$$\alpha < \frac{2(r + a'')}{(h + d_1' + d_2')}. \tag{4}$$

Stage 5: Bounds on α

Combining (1), (2), (4), we get

$$\frac{e_1 + e_2 + c_1 + c_2 + d_1 + d_2 + d_1' + d_2'}{a - d_1 - d_2 - d_1' - d_2'} \;<\; \alpha \;<\; min\{\frac{2r}{h} \;,\; \frac{2(r + a'')}{(h + d_1' + d_2')}\}. \tag{5}$$

Let $min\{\frac{2r}{h} \;,\; \frac{2(r+a'')}{(h+d_1'+d_2')}\} = \frac{2r}{h}$. We first prove that the condition $\frac{e_1+e_2+c_1+c_2+d_1+d_2+d_1'+d_2'}{a-d_1-d_2-d_1'-d_2'} < \alpha < \frac{2r}{h}$ can never exist. In other words, we establish (in stage 7) that both cases A and B cannot simultaneously occur.

Later we consider the case $min\{\frac{2r}{h} \;,\; \frac{2(r+a'')}{(h+d_1'+d_2')}\} = \frac{2(r+a'')}{(h+d_1'+d_2')}$ and establish in stage 8 that both cases A and C cannot simultaneously occur..

Stage 6: Fact from convexity

To ensure convexity of the congested region, $\gamma(m) < 90° - \beta_2(b)$, where $\beta_2(b)$ is the second edge angle of $v(b)$. Since for any convex polygon, $\overline{\gamma}(m) < \gamma(m)$, $\overline{\gamma}(m) < 90° - \beta_2(b)$.

Since $0° \leq \overline{\gamma}(m) \leq 90°$,

$$\tan \overline{\gamma}(m) < \tan(90° - \beta_2(b)) \Rightarrow \tan \overline{\gamma}(m) < \cot \beta_2(b) \Rightarrow \cot \overline{\gamma}(m) > \tan \beta_2(b) \tag{6}$$

Adding $\cot \overline{\gamma}(E_3)$ to both sides of (6), we get

$\cot \overline{\gamma}(m) + \cot \overline{\gamma}(E_3) > \tan \beta_2(b) + \cot \overline{\gamma}(E_3)$, i.e., $\frac{1}{\cot \overline{\gamma}(m)+\cot \overline{\gamma}(E_3)} < \frac{1}{\tan \beta_2(b)+\cot \overline{\gamma}(E_3)}$

$$\Rightarrow \frac{2}{\cot \overline{\gamma}(m) + \cot \overline{\gamma}(E_3)} < \frac{2}{\tan \beta_2(b) + \cot \overline{\gamma}(E_3)}. \tag{7}$$

Stage 7: Establishing both cases A and B cannot simultaneously occur

Combining (1), (3), (7), we get

$\frac{e_1+e_2+c_1+c_2+d_1+d_2+d_1'+d_2'}{a-d_1-d_2-d_1'-d_2'} < \alpha < \frac{2}{\cot \overline{\gamma}(m)+\cot \overline{\gamma}(E_3)} < \frac{2}{\tan \beta_2(b)+\cot \overline{\gamma}(E_3)}$

$\Rightarrow \frac{e_1+e_2+c_1+c_2+d_1+d_2+d_1'+d_2'}{a-d_1-d_2-d_1'-d_2'} < \frac{2}{\tan \beta_2(b)+\cot \overline{\gamma}(E_3)}$

$$\Rightarrow (e_1 + e_2 + c_1 + c_2 + d_1 + d_2 + d_1' + d_2')(\tan \beta_2(b) + \cot \overline{\gamma}(E_3)) < 2(a - d_1 - d_2 - d_1' - d_2'). \tag{8}$$

Note that for any convex polygon, $\overline{\gamma}(E_3) < 90° - \psi$. Since $\psi, \overline{\gamma}(E_3) < 90°$, $\cot \overline{\gamma}(E_3) > \cot(90° - \psi)$, i.e.,

$$\cot \overline{\gamma}(E_3) > \tan \psi. \tag{9}$$

From (8) and (9),

$$(e_1 + e_2 + c_1 + c_2 + d_1 + d_2 + d_1' + d_2')(\tan \beta_2(b) + \tan \psi) < 2(a - d_1 - d_2 - d_1' - d_2'). \tag{10}$$

From Figure 5, $\frac{b_1+y_1}{c_1} = \tan \beta_2(b) \ \Rightarrow y_1 = c_1 \tan \beta_2(b) - b_1$.

Similarly $\frac{b_2+y_2}{c_2} = \tan \psi \ \Rightarrow y_2 = c_2 \tan \psi - b_2$.

Since $h + (b_1 + y_1) + (b_2 + y_2) = a + h + b_1 + b_2$, we get $a = y_1 + y_2$.

Hence $a = c_1 \tan \beta_2(b) - b_1 + c_2 \tan \psi - b_2$.

Substituting a in (10), we get

$$\begin{aligned}(e_1 + e_2 + c_1 + c_2 + d_1 + d_2 + d_1' + d_2')(\tan \beta_2(b) + \tan \psi) < \\ 2(c_1 \tan \beta_2(b) - b_1 + c_2 \tan \psi - b_2 - d_1 - d_2 - d_1' - d_2').\end{aligned} \tag{11}$$

Note that

$$(c_1 + d_1 + d_2 + e_1) = (c_2 + d_1' + d_2' + e_2). \tag{12}$$

Simplifying (11) by using (12), we get

$$(e_1 + d_1 + d_2) \tan \beta_2(b) + (e_2 + d_1' + d_2') \tan \psi < -(d_1 + d_2 + d_1' + d_2' + b_1 + b_2). \tag{13}$$

Due to the convexity of the polygon, $0^\circ \leq \psi \leq 90^\circ$ and $0^\circ \leq \beta_2(b) \leq 90^\circ$. Hence the LHS of (13) is always positive and inequality (13) presents a contradiction. Thus the 3-entry 3-exit path cannot be simultaneously cheaper than the staircase 1-entry 1-exit path and 2-entry 2-exit path.

Stage 8: Establishing both cases A and C cannot simultaneously occur

We now consider the case $min\{\frac{2r}{h} , \frac{2(r+a'')}{(h+d_1'+d_2')}\} = \frac{2(r+a'')}{(h+d_1'+d_2')}$. This implies that the upper bound imposed on α by the non-staircase 1-entry 1-exit path is tighter than the upper bound imposed by the 2-entry 2-exit path; in other words, the non-staircase 1-entry 1-exit path does not enter the congested region $\overline{G}_1$ a second time (bypasses the interior travel between entry point E_4 and exit point E_5 through the vertex $v(1)$) and is hence cheaper than the 2-entry 2-exit path. Since $\overline{G}_1$ is convex, an imaginary path with 3 entries and three exits (such as $OE_1E_2v(m)E_3E_4E_5D$) would also bypass the interior travel between E_4 and E_5, thus becoming a 2-entry 2-exit path, which in this case is costlier than the non-staircase 1-entry 1-exit path under consideration. Hence a 3-entry 3-exit path that is simultaneously cheaper than a non-staircase 1-entry 1-exit path and a 2-entry 2-exit path does not exist.

Stage 9: More than 3 entry/exits

If a path enters and exits the congested region more than 3 times, we simply consider a sub-path that enters and exits exactly three times and observe that this sub-path cannot be of minimal cost. Thus we can keep reducing the number of entry/exits till we hit the 3 entry/exit case, which as shown previously cannot hold as well.

The theorem follows. ■

As an immediate implication of Theorem 5.1, we associate memory with probes (described later in step 7(b) of Section 6.1) in the "memory-based probing" procedure.

6 Memory-Based Probing Algorithm

In this section, the "memory based probing procedure" between any pair of origin-destination points in the presence of convex polygonal congested regions is discussed. Our solution methodology to determine the least cost path P^*_{OD} between origin-destination points O and D contains two major steps:

(i) the probing procedure whose steps are described in Section 6.1, and

(ii) generation of the network, described in Section 6.3. Note that the network generated as a result of the probing procedure contains staircase as well as non-staircase paths.

6.1 Probing Algorithm

The steps of the "memory based probing procedure" are outlined as follows:

1. Choose any pair of origin-destination points O, D.
2. Check if the origin and destination simply communicate following criteria listed in [9]. If so, draw one probe each from the origin and destination that intersect at a common point. Terminate the probes

at the point of intersection.

3. Check if the origin and destination communicate but do not communicate simply. If so, draw a feasible staircase path between the points, applying the Path-Push and Amalgamation procedure of [9].
4. If the origin and destination do not communicate, draw $E(\overline{G})$, the smallest rectangle that encloses O, D, and the set of all congested regions $\overline{G}$. The edges of $E(\overline{G})$ are parallel to the horizontal and vertical axes.
5. Probe in all four rectilinear directions from the origin.
6. For each probe, the following possibilities exist:
   (a) A probe intersects no congested region. Such a probe is terminated when it intersects $E(\overline{G})$.
   (b) A probe intersects an edge $\overline{v(m,i)v(m+1,i)}$ of a congested region $\overline{G}_i$, generating an entry/exit point E and partitioning the edge into two sub-edges $\overline{v(m,i)E}$ and $\overline{Ev(m+1,i)}$. From E, the probe proceeds along the sub-edge that lies at an obtuse angle to the incident probe and terminates at the end point of the sub-edge (a vertex of the congested region). We denote such a vertex as $b(m,i)$ and refer to it as a "bypass vertex". However, if the probe is incident at a right angle, then it terminates at the point of incidence E.
7. For entry/exit point E:
   (a) If E communicates (simply or otherwise) with the destination, probe as in steps 2 and 3.
   (b) Else, probe in all possible rectilinear directions apart from:
      i. the direction that would split the (obtuse) angle of incidence into two smaller angles (an acute angle and a right angle). For the ease of understanding, consider the example illustrated in Figure 6 in which the dotted direction should not be probed. The same probing rule applies if the angle of incidence is inside a congested region.
      ii. the directions that were probed by two previous entry/exit points, so that the formation of a turning step within a congested region (as illustrated in Figure 1) is prohibited. This can be achieved by associating a memory with each probe, which stores its direction of probing.

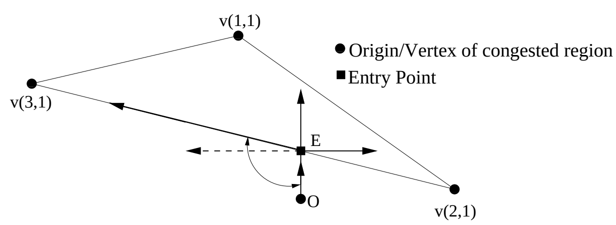

Figure 6: Probing procedure: step $7b(i)$

8. For each "bypass vertex" $b(m,i)$:

(a) If $b(m,i)$ communicates (simply or otherwise) with the destination, probe as in steps 2 and 3.

(b) Else, consider $b(m,i)$ as a pseudo-origin and repeat steps 5 to 7. The necessity of step $8b$ is explained in the following sub-section.

9. Let $B = \cup_{i,m}\{v(m,i)\} \setminus \cup_{i,m}\{b(m,i)\}$ represent the set of congested region vertices where a probe never terminated. For any $v_t \in B$, $t = 1,\ldots,Kp$, where p is the number of congested regions and K, a constant, is an upper bound on the number of edges of each congested region, repeat steps 2 to 8 considering $v_t \in B$ as a pseudo-origin.
10. Interchange origin and destination and repeat steps 5 to 8.
11. Sequentially combine probes to form paths and eliminate all 3-entry 3-exit paths.

6.2 Correctness of the Probing Procedure

Lemma 6.2.1. *The memory-based probing procedure of Section 6.1 is correct.*

Proof: The correctness of the probing procedure follows from the correctness of its various steps. The correctness of Steps 2 and 3 follows from the definition of simply communicating vertices and Theorem 1 of [9] respectively. Step 4 is necessary because the least cost path (union of probes) between O and D must be contained within $E(\overline{G})$. Step 5 ensures that all paths are rectilinear. Step $6b$ is identical to step 3 of the vertex seeking tree construction procedure of Larson and Li [9]. It only differs from [9] in the generation of entry/exit points (since our paper considers congested regions and not barriers). Step $7b(ii)$ eliminates turning steps within congested regions (as illustrated in Figure 1). Step $8b$ follows from the probing procedure of [9], when the least cost path bypasses a congested region via its vertices (i.e., the congested region acts as a barrier to travel) and thus recognizes the possibility that a least cost path may be non-staircase. Consider the example illustrated in Figure 4(a), in which the least cost path Ov_4v_1D between O and D bypasses $\overline{G}_1$ through the vertices v_4 and v_1 for $\alpha > 1.33$. Step 9 is necessitated by the exact same reason as step $8b$. Step 10 is essential because our probing procedure yields unidirectional arcs. Step 11 follows from Theorem 5.1. ■

6.3 Generation of the Network and Shortest Path Application

The network for our problem is $H(W,L)$, where $W = (\cup_{i,m}\{v(m,i)\} \cup (\cup_i E_i))$ is the set of nodes (E_i denotes the set of entry/exit points of congested region $\overline{G}_i$) and L is the set of arcs (i.e., probes). We partition the arc set L into subsets L_1 and L_2 where $L = L_1 \cup L_2$. L_1 denotes the set of arcs lying outside congested regions. The length of all such probes is equal to the length of the shortest rectilinear path between the start and end points of the probe. L_2 denotes the set of arcs lying inside congested regions. The length of such an arc is equal to $(1+\alpha_i)$ times the length of the shortest rectilinear path between the start and end points of the arc, where α_i denotes the congestion factor of $\overline{G}_i$ in which the arc is fully contained. P^*_{OD} can be obtained by solving the shortest path problem on this constructed network $H(W,L)$. Depending on the size of the network generated, different shortest path algorithms can be employed. We discuss these issues in greater detail later in Section 7.2.

7 Computational Results

7.1 Generation of Problem Instances

The numerical tests were performed on different problem instances. User input parameters of each problem were the number of congested regions, their congestion factors, and the coordinates of the origin and destination. The number of vertices of each congested region was generated using standard random number generators but never exceeded 10. The coordinates of the vertices were obtained by geocoding the vertices of convex and non-intersecting land polygons (in an ArcGIS 9.1 framework) scattered across major metropolitan areas in a five county region (San Bernardino, Riverside, Los Angeles, Orange County, and San Diego) in southern California. The origin and destination were placed such that they were rectilinearly invisible *vis-à-vis* most congested regions. Low congestion factors (less than 1) were assigned to each region to maximize the chances of a potential least cost path to pass through most congested regions. This was necessary to ensure that the computation results reveal the true extent of the number of entry/exit points generated when most of the congested regions are intersected. The memory-based probing algorithm was implemented in C programming language.

7.2 Numerical Test Results

The results in Table 1 indicate the variation in the number of entry/exit points generated (ρ) and the runtime (τ) with the number of congested regions intersected (ϕ) due to probing. Note that (i) ρ and τ are functions of ϕ, and not p, as all congested regions are not necessarily intersected during probing, (ii) τ reflects the time taken by the probing procedure to generate all entry/exit points, and not to determine the least cost path between the origin and destination, and (iii) the results in Table 1 are dependent on the locations of the congested regions in $\Re^2$, their relative proximity and the locations of the origin-destination points relative to the congested regions.

The plots of ϕ versus ρ and ϕ versus τ are illustrated in Figures 7 and 8 respectively. The inherent exponential nature of the problem is clearly evident from these plots.

A plot of ϕ versus $\log_2 \rho$ is illustrated in Figure 9. The relationship between ϕ and $\log_2 \rho$ can be approximated by a straight line (using standard regression approaches), which has a slope of approximately 0.47; implying that, for the numerical tests conducted, our probing algorithm generated $O(2^{0.47\phi})$, i.e., $O(1.385^{\phi})$ entry/exit points. Clearly the algorithm performs much better in terms of the total number of entry and exit points generated, the upper bound of which is $O(4^p)$.

As evident from the computational test results, the number of entry/exit points generated for a reasonably large-sized problem instance (with 35 congested regions) is approximately $160,000$. The size of the

Table 1: Test Results

Problem #	p	ϕ	ρ	τ (CPU secs.)
1	1	1	15	7
2	2	2	11	22
3	4	3	77	37
4	6	5	156	31
5	10	7	1586	61
6	15	9	3194	89
7	20	12	11407	174
8	25	17	50383	555
9	30	26	100823	736
10	35	31	160919	1505

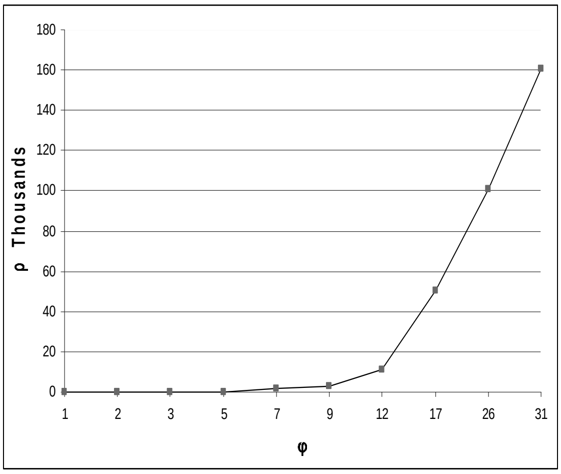

Figure 7: ϕ versus ρ

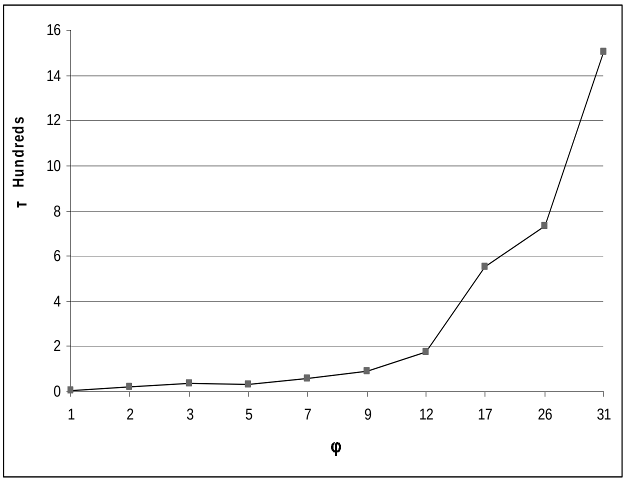

Figure 8: ϕ versus τ

constructed network $H(W, L)$ is the main determinant of the shortest path algorithm employed to obtain P^*_{OD}. The standard shortest path algorithm given by Dijkstra [7] and its various efficient implementations, well studied in Ahuja, Magnanti and Orlin [1] would suffice for most problem instances. However, if computation times become prohibitively high, either (i) large scale shortest path algorithms ([5]), or (ii) other heuristic approaches have to be employed to determine P^*_{OD}. Zhan and Noon [20] provides a comprehensive and objective performance evaluation of 15 shortest path algorithms using a variety of real road networks and recommends Dijkstra's algorithm with "approximate buckets" implementation for the one-to-one (i.e., between one pair of nodes) shortest path problem.

7.3 Discussion on Computational Complexity

Larson and Li's [9] Polypath algorithm to determine the minimum rectilinear distance path(s) in the presence of barriers is polynomially bounded in the number of nodes (comprised of barrier vertices and origin-destination points) of the constructed network. In our case, the size of the generated network is also a func-

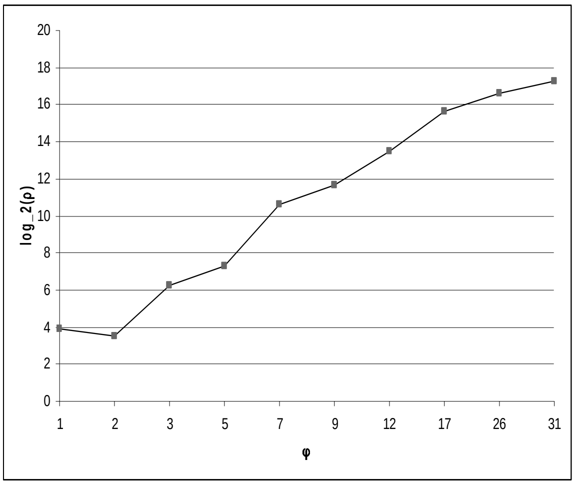

Figure 9: ϕ versus $\log_2 \rho$

tion of the number of entry/exit points generated by the memory based proving algorithm. Clearly the number of entry/exit points is a function of the number of congested regions intersected during probing (ϕ) and the number of times each congested region is intersected during probing. Our computational experience suggests that ϕ itself is a function of (i) size, shape, and number of vertices of the congested regions, (ii) their relative placement in $\Re^2$, and (iii) location of the origin-destination points relative to the congested regions. It is pertinent to mention here that these three factors compound the worst case performance analysis of the algorithm. Additionally, recall that as a result of Observation 2 (discussed earlier in Section 4.2), *a priori* pruning of probes is not possible, and consequently the state space (of entry/exit points) seems to grow exponentially. Our empirical analysis (reported in the literature [1] as one of three widely adopted approaches for measuring the performance of an algorithm) suggests that the memory-based probing algorithm is an exponential time algorithm as its worst case running time and size of the network generated grow exponentially as a function of ϕ. Note that while this is consistent with the inherently exponential nature of the network generation problem, the performance of the algorithm for small and medium sized problem instances ($\phi \leq 12$) is satisfactory. In conclusion, the authors conjecture that the problem of generating the network for the least cost path problem in the presence of congested regions is NP-hard. However a formal proof of NP-hardness (via the transformation of other network optimization problem(s) to the entry/exit point network generation problem) eludes the authors at this point and is hence outlined as a direction for future research.

7.4 Special Case: Known Entry-Exit

When all congested regions in $\Re^2$ are rectangles with their edges parallel to the travel axes, Sarkar *et al.* [16] established that the least cost path between a pair of origin-destination points is a subset of the grid structure proposed by Butt and Cavalier. We now consider a special case of our problem when the number of entry/exit points of a least cost path between a pair of origin-destination points (in the presence of **arbitrarily shaped** congested regions) and their exact locations are known *a priori*.

Consider a university campus located in an urban area. Speed limits inside a campus are typically less than city roads or highways. Additionally, one can enter and leave a campus at a fixed number of entry/exit points whose locations are distinguished. Such an area could be considered to be an arbitrarily shaped congested region with a fixed number of entry/exit points whose locations are predetermined. This is a special case of our problem. The objective is to determine the rectilinear least cost path between all pairs of origin-destination points for this special case. However, note that the discussion in this section applies to a more general shape of congested regions and is not limited to convex polygonal congested regions only.

Batta and Palekar [2] introduced a novel modeling framework for location problems that consisted of a mixture of planar and network components. In [2], the authors introduced the concept of mega nodes, which are arbitrarily shaped planar regions inside which any travel always follows the rectilinear distance metric. The authors also considered the existence of infinitesimal nodes of zero dimension. Demand was distributed at demand points which are (i) infinitesimal nodes, and (ii) a finite number of points spatially scattered in the mega nodes. Travel to/from the infinitesimal nodes to demand points located inside mega nodes occurs through "gates" located on the boundaries of mega nodes. Along the lines of Larson and Sadiq [10], [2] divided the region inside mega nodes into cells and proved that at least one shortest path from a demand point to a facility located in a cell C passes though the cell corner of C.

Note that mega nodes of [2] can be considered to be congested regions and their gates could correspond to the entry/exit points of a congested region. We assume that each such congested region is characterized by a finite number of points of tangency through which tangential lines can be passed along the rectilinear travel directions. We now describe our grid construction procedure, as outlined in the following steps and as illustrated in Figure 10.

1. Analogous to the grid construction procedure of [2], we pass horizontal and vertical lines through the known entry/exit point locations of the congested regions.

    (a) The lines drawn inside the same congested region are terminated when they intersect the boundary of the region.

    (b) The lines extending outside from a congested region are terminated when they intersect other congested regions, otherwise at the boundary of the smallest bounding rectangle.

2. Draw all possible tangency lines in the interior and exterior of the congested regions, along the directions of travel. Lines extending into the interior of a congested region are terminated as in step 1(a). Lines drawn in the exterior of a congested region are terminated as in step 1(b).

3. Draw horizontal and vertical lines through the origin-destination points which terminate when they intersect other congested regions, or at the smallest bounding rectangle.

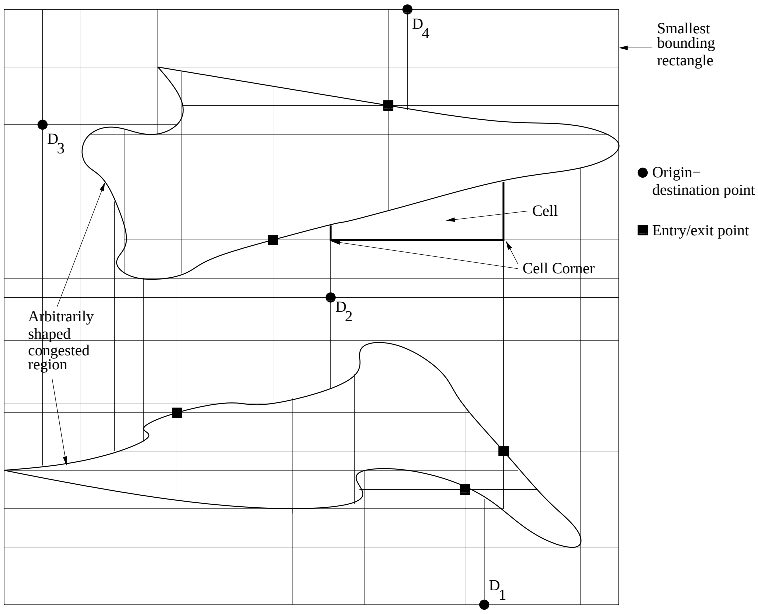

Figure 10: Special case: number of entry/exit points and their locations are known

A closed bounded region between intersection points on the grid and congested region boundaries (that lie outside congested regions but inside the bounding rectangle) are defined as cells. Points of intersection of gridlines with other gridlines or with congested region boundaries, that define a cell, are referred to as cell corners. We now state and prove the following lemma:

Lemma 7.4.1. *At least one least cost path between a pair of origin-destination points will coincide with segments of the grid, drawn as described in the previous steps.*

Proof: If the least cost path between a pair of origin-destination points bypasses all existing congested regions, the congested regions can be thought of as barriers to travel and the grid structure of [10] (which would consist of lines drawn in steps 2 and 3 previously but exclude the lines drawn in step 1) would apply. However, the grid structure of [10] is a subset of our grid structure. The lemma follows from Lemma 3 of [10].

If the least cost path between a pair of origin-destination points passes through the existing congested regions, the proof follows from Lemma 1 of [2]. ■

The rectilinear least cost path between a pair of origin-destination points can thus be determined by solving the shortest path problem on the network specified by the grid construction procedure. The cell corners, origin-destination points, entry/exit points, and points of intersection of gridlines that lie within congested regions comprise the nodes of the network. To formalize this, we note that N_1 congested regions generate at most CN_1 horizontal and CN_1 vertical lines, where C, a constant, signifies the upper bound on the number of tangency points of each congested region. Similarly, the entry/exit points generate at most N_2 horizontal and N_2 vertical lines; while the origin-destination points generate at most N_3 horizontal and N_3 vertical lines. Hence the maximal number of cells is $O({N_1}^2 + {N_3}^2)$, and the maximal number of points of intersections of gridlines that lie within congested regions is $O({N_1}^2 + {N_2}^2)$. Clearly the number of cells and points of intersections of gridlines that lie within congested regions is $O(n^2)$. Typically fewer cells and interior points of intersection will be generated due to the terminating conditions mentioned earlier in the steps of the grid construction procedure.

The length of an arc between a pair of nodes is given by either (a) the length of the horizontal or vertical segment between the nodes, or by (b) $(1+\alpha_i)$ times the length of the horizontal or vertical segment between the nodes, if they lie with congested region $\overline{G}_i$. Since the standard shortest path algorithm [7] is itself an $O(n^2)$ algorithm, a polynomial time $O(n^4)$ solution to the special case can thereby be obtained.

8 Conclusions and Future Research

This paper analyzes the problem of determining the least cost path between a pair of origin-destination points in the presence of convex polygonal congested regions. It has been established that the state-space for the problem could be exponential, i.e., there could possibly be $O(4^p)$ potential staircase least cost paths between a specified pair of origin-destination points in the presence of p congested regions. An upper bound for the number of entry/exit points for a single convex polygonal congested region has been obtained. A "memory-based probing algorithm" has been proposed for the problem. It can be viewed as a generalization of the algorithm proposed by Larson and Li [9] for the corresponding "barrier" problem. Computational experience for various problem instances has been reported, which reflect the inherent exponential nature of the problem.

When the network constructed by the "memory-based probing algorithm" becomes very large and consequently computation of the least cost path is prohibitively expensive, $\overline{G}$, the set of all congested regions, can be partitioned into smaller subsets. Analogous to the idea proposed by Lombard and Church [12] (who define a "gateway shortest path" to be a path between an origin and a destination that has to pass through a specified node, called a "gateway"), each such subset can be assumed to have one or multiple "gateways" (entry point(s) or exit point(s)) specified *a priori*, through which the least cost path between a given origin and a destination is constrained to pass. The motivation to partition the congested regions into smaller subsets is to reduce prohibitively high shortest path computation times for large scale networks. For each subset, the "local" least cost path can be determined between the entry gateway(s) and exit gateway(s). A heuristic least cost path can be obtained by comparing the lengths of combinations of the least cost paths in individual subsets of congested regions. The heuristic will yield an efficient path in terms of its length. This path will clearly be sub-optimal; however, its length will provide a valid upper bound to the length of P^*_{OD}.

We conclude this paper with the following directions for future research. Until now, the congestion factors (α) have been assumed to be constants. However, in regional, urban and facility settings alike, the congestion depends on the time of the day. Such situations can be modeled by dynamic congestion factors, where the parameter α is replaced by a function of time $\alpha(t)$. The probing algorithm and network formation is expected to be unchanged. The analysis will require time-dependent shortest path algorithms ([15], [21], [4]) on the network.

The dynamic routing of unmanned aerial vehicles (UAVs) is a problem frequently encountered in the context of military reconnaissance operations. Such operations are often performed in high risk environments due to the presence of hostile forces. A hostile region and its "risk level" (either dynamic or stochastic) could be surrogates for a congested region and its congestion factor. However, for this airborne version of the traditional vehicle routing problem, the Euclidean (or l_2) distance metric (rather than the rectilinear distances, as considered in this paper) would clearly provide a better (and more practical) measure of distances.

Acknowledgement

This work was supported by the National Science Foundation, via grant $DMI - 0300370$. This support is gratefully acknowledged. The first author also wishes to acknowledge graduate students Seethal Mishra and Abhay Joshi, Department of Industrial Engineering, SUNY at Buffalo for their help during implementation of the probing algorithm.

References

[1] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, Englewood Cliffs, NJ, 1993.

[2] R. Batta and U. Palekar. Mixed planar/network facility location problems. *Computers and Operations Research*, 15(1):61–67, 1988.

[3] S.E. Butt and T.M. Cavalier. Facility location in the presence of congested regions with the rectilinear distance metric. *Socio-Economic Planning Sciences*, 31(2):103–113, 1997.

[4] I. Chabini. Discrete dynamic shortest path problems in transportation applications: complexity and algorithms with optimal run time. *Transportation Research Record*, 1645:170–175, 1998.

[5] Y. Chou, E. Romeijn, and R. Smith. Approximating shortest paths in large scale networks with application to intelligent transportation systems. *INFORMS Journal on Computing*, 10(2):163–179, 1998.

[6] K.L. Clarkson, S. Kapoor, and P.M. Vaidya. Rectilinear shortest paths through polygonal obstacles in O$(n(log\ n)^2)$ time. In *Proceedings of the 3rd Annual ACM Symposium on Computational Geometry*, pages 251–257, 1987.

[7] E.W. Dijkstra. A note on two problems in connection with graphs. *Numerical Mathematics*, 1:269–271, 1959.

[8] S. Guha and I. Suzuki. Proximity problems for points on a rectilinear plane with rectangular obstacles. *Algorithmica*, 17(3):281–307, 1997.

[9] R.C. Larson and V.O.K. Li. Finding minimum rectilinear distance paths in the presence of barriers. *Networks*, 11:285–304, 1981.

[10] R.C. Larson and G. Sadiq. Facility locations with the Manhattan metric in the presence of barriers to travel. *Operations Research*, 31(4):652–669, July-August 1983.

[11] D. T. Lee, C. D. Yang, and C. K. Wong. Rectilinear paths among rectilinear obstacles. In *ISAAC: 3^{rd} International Symposium on Algorithms and Computation (formerly SIGAL International Symposium on Algorithms), Organized by Special Interest Group on Algorithms (SIGAL) of the Information Processing Society of Japan (IPSJ) and the Technical Group on Theoretical Foundation of Computing of the Institute of Electronics, Information and Communication Engineers (IEICE))*, 1992.

[12] K. Lombard and R.L. Church. The gateway shortest path problem: Generating alternative routes for a corridor location problem. *Geographical Systems*, 1:25–45, 1993.

[13] J.S.B. Mitchell. l_1 shortest paths among polygonal obstacles in the plane. *Algorithmica*, 8:55–88, 1992.

[14] J.S.B. Mitchell and C.H. Papadimitriou. The weighted region problem: Finding shortest paths through a weighted planar subdivision. *Journal of the Association for Computing Machinery*, 38(1):18–73, 1991.

[15] A. Orda and R. Rom. Shortest-path and minimum-delay algorithms in networks with time-dependent edge-length. *Journal of the ACM*, 37(3):607–625, July 1990.

[16] A. Sarkar, R. Batta, and R. Nagi. Commentary on 'Facility Location in the Presence of Congested Regions with the Rectilinear Distance Metric'. *Socio-Economic Planning Sciences*, 38(4):291–306, December 2004.

[17] A. Sarkar, R. Batta, and R. Nagi. Planar area location/layout problem in the presence of generalized congested regions with the rectilinear distance metric. *IIE Transactions*, 37:35–50, January 2005.

[18] S. Schuierer. An optimal data structure for shortest rectilinear path queries in a simple rectilinear polygon. *International Journal of Computational Geometry and Applications*, 6(2):205–226, 1996.

[19] C. Yang, D. T. Lee, and C. K. Wong. Rectilinear path problems among rectilinear obstacles revisited. *SIAM Journal on Computing*, 24(3):457–472, 1995.

[20] F.B. Zhan and C. Noon. Shortest path algorithms: An evaluation using real road networks. *Transportation Science*, 32(1):163–179, February 1998.

[21] A.K. Ziliaskopoulos and H.S. Mahmassani. A time-dependent shortest path algorithm for real-time intelligent vehicle/highway systems applications. *Transportation Research Record*, 1408:94–100, 1993.