Planar Area Location/Layout Problem in the Presence of Generalized Congested Regions with the Rectilinear Distance Metric

August 2003

Abstract

This paper considers the problem of placing a single rectangular generalized congested region (GCR) of given area but unknown dimensions in the presence of other rectangular GCRs, where the edges of the rectangles are parallel to the travel axes. GCRs are closed and bounded regions in $\mathbb{R}^2$ in which facility location is prohibited but travel through is allowed at an additional cost per unit distance. An interactive model is considered in which there is interaction not only between the server of the new GCR and the users (of the existing GCRs) but between the existing users themselves. Two versions of the problem are considered when: (i) the server of the new GCR is located on its boundary but its exact location is unknown, and (ii) the server is located inside the GCR at its centroid. The feasible region is divided into cells obtained by drawing a grid. We analyze the problem based on whether the new GCR’s placement intersects gridlines. When the new GCR does not intersect gridlines, we prove that the optimal location can be drawn from a finite set of candidate points. However when the new GCR intersects gridlines, we split the feasible region by Equal Travel Time Partitions (ETTP) such that the flows through gridlines can be uniquely classified as (i) travel through, or (ii) left bypass, or (iii) right bypass. The solution methodologies for all cases are shown to be polynomially bounded in the number of GCRs.

Keywords: Single Facility Location, Single Facility Layout, Generalized Congested Regions, Rectilinear Distance Metric.

1 Introduction

Facility layout and facility location are critical components in the overall problem of facilities design. Traditionally however, these problems have mostly been studied independently. The facility layout problem and the facility location problem have grown into different areas of interest even though they are closely related. Facility layout and facility location both deal with the location of new facilities in a region where there are existing facilities. In the facility location problem, facilities are typically infinitesimal. However, in facility layout, new constraints determining the area requirements and locational restrictions have to be considered because the facilities have finite sizes. In essence, the facility layout problem can be viewed as an area location/layout/placement problem. The relationships between the facilities appear in the objective function in the form of (i) a unit cost of interaction between facilities, (ii) the amount of interaction, and
(iii) the (rectilinear) distance between facilities. The finite size facility placement problem, introduced by Savas, Batta and Nagi [19] is the first attempt to bridge the gap between these two areas. The authors’ consideration of “area” in a location problem brought their approach closer to layout theory. Adapting the results of [19], Savas [18] proposes a non-traditional approach to the facility layout problem in which all facilities are rectangular and pose restrictions to rectilinear travel.

Restricted location problems are a class of location problems in which the minimum travel distance between two points in $\Re^2$ is increased by certain predescribed areas which obstruct travel. Barriers (prohibit travel through: for example, machines on a shop floor), forbidden regions (allow travel through at no extra cost: for example, aisles on a shop floor) and generalized congested regions (GCR) (allow travel through at a penalty: for example, assembly areas, finishing areas on the shop floor, through which travel is permitted but is slowed down due to congestion) are examples of restrictions. Restricted location problems have been studied by Katz and Cooper [9], Aneja and Parlar [1], Larson and Li [10], Larson and Sadiq [11], Batta, Ghose and Palekar [2], Nandikonda, Batta and Nagi [14], Dearing, Hamacher and Klamroth [7], Butt and Cavalier [5, 4] and others. Most of these works determine the optimal location of an infinitesimal new facility such that the (i) total weighted travel distance from the new facility to the existing facilities (minisum or median objective), or (ii) maximum weighted travel distance from the new facility to the existing facilities (minimax or center objective) in the presence of restricted regions is minimized.

In contrast, the objective of the facility layout problem is to minimize the total weighted distance between entities, in which the weights reflect material flow volumes or adjacency priorities. The output of the facility layout problem is a “block layout” which specifies the relative location of each department. Further work can be performed to obtain the “detailed layout” which specifies exact locations of departments, aisles, input-output points and the layout within each department. The facility layout problem is well studied in literature and practice. Excellent books on facilities layout exist, for example, the textbook by Francis, McGinnis and White [16]. In fact, this textbook also comprehensively covers facility location problems. Meller and Gau [12] is a detailed survey in which emerging trends in the facility layout problem in the decade of the nineties were studied. In the facility layout problem, the material handling between department pairs is typically assumed to take place along rectilinear paths between the department centroids. This centroid assumption often leads to inaccurate distance measurements. In an attempt to remove the limitations of the centroid-to-centroid distance measure, Bozer and Meller [3] developed the expected distance measure (EDIST), defined as the expected distance between any two points in each department. Castillo and Peters [6] and Norman, Smith and Arapoglu [15] have also tried to address this issue. Montreuil [13] presented a Mixed Integer Programming formulation that integrates the flow path design problem with the facility layout problem. A specialized case of Montreuil’s model was developed by Heragu and Kusiak [8]
where a department’s length, width and orientation are specified \textit{a priori}.

In this paper, we develop the preliminary results for a new approach to the facility layout problem. We cast the facility layout problem as a facility location problem, in which GCRs pose restrictions to travel. More specifically, we consider the problem of placing a new facility, which itself is a GCR, in the presence of existing facilities that are treated as GCRs to travel. The area of the new GCR to be placed is known but its exact dimensions are unknown. Assembly areas and finishing areas in a manufacturing facility could be considered to be GCRs through which travel is permitted but is penalized due to congestion slowdown. Also the exact dimensions of such areas may vary but their shapes can be approximated by rectangles. In a way, this paper may be viewed as an extension of the work by Savas \textit{et al.} [19] in which the new facility is arbitrarily shaped but has a fixed contour. Also the restriction in [19] comes in the form of barriers, as compared to GCRs in our work. As noted earlier, barriers and forbidden regions can be viewed as special cases of GCRs. Hence the area location/layout problem in the presence of GCRs subsumes the area location/layout problem in the presence of barriers and forbidden regions.

The remainder of this paper is organized as follows. In §2, we formally introduce and define the problem. In §3, we describe a grid construction procedure and present some definitions and results essential for our subsequent analysis. We classify the problem in §4 and present the solution methodology when the new GCR placement does not intersect any gridline in §4.2. In §5, we discuss the solution methodology when the new GCR intersects at least one gridline. The concept of “Equal Travel Time Partition” is introduced in this section. The complexity of the solution methodologies proposed in §4.2 and §5 is analyzed in §6. The paper is concluded by a final section in which the contribution of this work is summarized and directions for future research are outlined.

2 Problem Description

2.1 Problem Statement

There exists a finite number of rectangular \textit{generalized congested regions (GCRs)} with edges parallel to the travel axes in which location/placement is prohibited but travel is permitted at a possible extra cost. The additional cost per unit distance is called the congestion factor of the GCR and is denoted by $\alpha$, $0 \leq \alpha < \infty$. Thus, if $w$ is the cost of travel per unit distance between two points lying outside a GCR, then the cost of travel between the same points when lying inside the region would be $(1 + \alpha)w$. Each GCR has one or multiple \textit{users} (input/output (I/O) points) through which it communicates with other GCRs. The users are located inside the GCRs or on their boundaries.

A new rectangular GCR with its edges parallel to the travel axes is to be placed in the presence of existing GCRs. The congestion factor and area of the new GCR are known, but not its exact dimensions.
The new GCR communicates with the users of the existing GCRs through a single server. The server acts like an I/O point like other users. We consider the following two versions of the problem:

- when the server is located on the boundary of the new GCR but its exact location is unknown.
- when the server location is known a priori and is assumed to be at the centroid of the new GCR.

In both versions of the problem, there exist flows between:

- pairs of users of existing GCRs. This is termed as the user-user interaction.
- an existing user and the server of the new GCR. This is termed as the user-server interaction.

The planar area location/layout problem is to determine the exact dimensions (specified by the length) of a new rectangular GCR and its optimal location (specified by the location of its server and the location of its top left corner) such that the new GCR does not overlap with existing GCRs and the sum of the user-server and user-user interaction is minimized.

2.2 Definitions and Notations

We assume that each GCR is a rectangular region in \( \mathbb{R}^2 \), with finite area and a continuous closed boundary. Due to the rectangular shape, a GCR can be characterized by two horizontal and two vertical lines, obtained by passing tangential lines through each of its vertices. Let \( G_j \) (an open set) denote the set of points \((x, y) \in \mathbb{R}^2\) contained strictly within GCR \( j \). We also define \( G_j = G_j \cup \{\text{boundary of GCR } j\} \), a closed set. We let \( G = \bigcup_j G_j \) and \( \overline{G} = \bigcup_j \overline{G_j} \). Let \( B \) denote the set of points contained strictly within the new GCR and let \( \overline{B} = B \cup \{\text{boundary of the new GCR}\} \). The distinction between the inside and the boundary of a GCR is necessary to permit travel without congestion slowdown on the boundary but not on the inside. Hence the boundary of the new GCR can be written as \( \overline{B} - B \). Let \( E(G) \) define the smallest bounding rectangle that encloses all GCRs and users and whose sides are parallel to the \( x \) and \( y \) axes.

We will now define the feasible region for the planar area location/layout problem. To this end, let \( B(p) \) (an open set) denote the set of points that correspond to the new GCR when its “placement” is \( p \). We also define \( \overline{B}(p) = B(p) \cup \{\text{boundary of the new GCR}\} \), a closed set. Let \( E_k(\overline{B}), j = 1, 2, 3, 4 \) denote the vertices of the new GCR, starting from the bottom left corner and labeling in the counter clockwise direction. We now define the term “placement”. Since the new GCR is a finite-sized entity, the coordinates of its server alone cannot convey full information on its placement in \( \mathbb{R}^2 \). Hence we define \( p = [X, E_4(\overline{B}), l] \) to be the location-dimension vector of the new GCR. Here \( X = (x, y) \) and \( E_4(\overline{B}) \) represent the location of the server of the new GCR and its top left corner respectively. When the new GCR’s server is on its boundary, \( X \in \overline{B} - B \). Whereas, when the server of the new GCR is at its centroid, \( X \in B \). \( l \) represents the length of
the new GCR and is measured along the $x$ axis. Note that $l \times b = A$, where $A$ denotes the area of the new GCR and is a known parameter and $b$ is its width. It is appropriate to mention here that Savas et al. define a location-orientation vector to specify the placement of their arbitrarily shaped finite-sized new facility. Refer to [19] for further details. However, since in our problem, the new GCR is always oriented parallel to the travel axes but its exact dimensions are unknown, we specify its placement by the location-dimension vector. Note that when the server of the new GCR is located at its centroid, determination of the optimal location of $E_4(B)$ and the optimal value of $l$ suffice. However for the sake of uniformity in definition in both versions of the problem (unknown boundary server and centroid server), our location-dimension vector has three components. The feasible region is now defined as follows:

$$F = \{ [X, E_4(B), l] : B(p) \cap G = \emptyset \}.$$

There are two types of interactions in our problem. Firstly, there is interaction between user $i$ and server $X$, denoted by $u_i \geq 0$. Secondly, there is interaction between user $i$ to user $j$, denoted by $w_{ij} \geq 0$ (note: we do not assume $w_{ij} = w_{ji}$). The interaction between any pair of users takes place through a least cost path between the users. It is pertinent to note here that the concept of least cost paths was introduced by Butt and Cavalier [5]. Figure 1 of [5] depicts different scenarios of least cost paths between an origin and a destination in the presence of a convex polygonal GCR. The least cost path between two users may pass through some GCRs and may bypass some, depending on their congestion factors. Sarkar, Batta and Nagi [17] have proved the correctness of Butt and Cavalier’s result that the least cost path(s) between two users in the presence of rectangular GCRs coincides with segments of a grid that is obtained by drawing vertical and horizontal lines through the vertices of each GCR and the users. The grid construction procedure is discussed in detail in §3.1. Let $d_p(i, j)$ represent the length of such a feasible least cost path between two users $i$ and $j$ when the placement of the new GCR is $p$. The subscript $p$ signifies that the distance is a function of the placement. Similarly, $d_p(i, X)$ represents the length of a feasible least cost path between user $i$ and the server at $X$ when the facility placement is $p$.

Let $D$ denote the set of all users. For a given facility placement $p = [X, E_4(B), l]$, the total weighted travel distance between users and the facility (user-server interaction) is $J(p)$ and correspondingly between all users (user-user interaction) is $K(p)$.

$$J(p) + K(p) = \sum_{j \in D} u_j \cdot d_p(j, X) + \sum_{i \in D} \sum_{j \in D} w_{ij} \cdot d_p(i, j).$$

The planar area location/layout problem is to determine the exact dimensions of the new rectangular GCR and its optimal placement $p$ such that $J(p^*) + K(p^*) \leq J(p) + K(p)$, $\forall p \in F$.

For the unknown boundary server version of the problem, we have five continuous variables, the coordinates of $E_4(B)$ and $X$, and $l$. For the centroid server case, we have three continuous variables since...
the coordinates of $X$ are known once $E_4(\mathcal{B})$ and $l$ are specified. For either situation, it turns out that the objective function is non-convex and non-concave and hence the use of generalized gradient methods could lead to a local optimal solution. We are able, however, to obtain a global optimal solution by the following line of attack. We divide the feasible region associated with $E_4(\mathcal{B})$ into subregions where the objective function is concave and thus a finite set of candidate values for $E_4(\mathcal{B})$ and $X$ are obtained. For each such candidate set of values of $E_4(\mathcal{B})$ and $X$, we then optimize $l$. This allows us to conclude that the solution methodology is polynomial in the number of existing GCRs.

3 Preliminaries

3.1 Grid Construction and Cell Formation

In order to develop our analysis, we first describe the grid construction procedure that helps to identify the least cost path between two points in the presence of GCRs.

![Figure 1: Grid for rectangular GCRs](image)

Figure 1 depicts an example with four rectangular GCRs. Each GCR has users located inside or on its boundary. A grid is constructed by passing horizontal and vertical lines through the vertices of each GCR and its users. The lines intersect the other GCRs and pass through until they terminate at the smallest bounding rectangle $E(\mathcal{G})$. The resulting set of lines are called node traversal lines by Larson and Sadiq [11]. For simplicity in presentation, we shall refer to these lines as “gridlines” henceforth. This grid is an extension of the grid proposed by [11] for solving the $p$-median problem in the presence of barriers: because the gridlines of [11] terminate when they intersect barriers (as traveling through barriers is prohibited). We let $H$ and $V$ denote the set of horizontal and vertical gridlines respectively. Let $S = H \cup V$ denote the set of
all gridlines. The GCRs along with \( S \) divide the feasible region \( F \) into a number of regions, called cells, as illustrated in Figure 1. Each cell boundary is composed of solely gridlines or segments of GCR boundaries and gridlines. Since the GCRs are rectangular and due to the way the grid is constructed, all cells generated are also rectangular with their edges parallel to the travel axes. For a given cell \( C \), let us consider the points \((x_{\text{min}}, y_{\text{min}}), (x_{\text{min}}, y_{\text{max}}), (x_{\text{max}}, y_{\text{min}}), (x_{\text{max}}, y_{\text{max}})\), where \( x_{\text{min}}, y_{\text{min}}, x_{\text{max}}, y_{\text{max}} \) are the respective bounds on the \( x \) and \( y \) coordinates on any point in the cell. We shall henceforth refer to these points (or vertices) as the cell corners of \( C \) and denote them as \( E_k(C) \), \( k = 1, 2, 3, 4 \) starting from the bottom left corner and labeling them in the counter clockwise direction.

### 3.2 Some Results

For our future analysis, it is necessary to present a few definitions from Larson and Li [10] and some previous results suitably adapted for our work:

1. A rectilinear path between two points in \( \mathbb{R}^2 \) is a path whose length is not made longer due to the presence of GCRs. The path proceeds from one point to another in a sequence of connected steps alternating between horizontal and vertical, with the step direction indicating the direction of travel. For example, path 1 between the users of GCRs \( G_1 \) and \( G_2 \) is a rectilinear path, as illustrated in Figure 1.

2. A “stair-case” path between \((x_i, y_i)\) and \((x_j, y_j)\) is a rectilinear path having length \(|x_i - x_j| + |y_i - y_j|\). For example, paths 2 and 3 between the users of GCRs \( G_1 \) and \( G_4 \), as illustrated in Figure 1.

3. Two points are said to communicate if there exists at least one feasible stair-case path between them. For example, users 1 and 2 and 1 and 4 in Figure 1 communicate. However user pairs \((2, 3), (2, 4), (1, 3), (3, 6)\) and \((2, 6)\) do not communicate. Hence they are referred to as “non-communicating”.

4. Result 1 (Theorem 3.1 of Sarkar et al. [17]): At least one least cost path between two points in the presence of rectangular GCRs (with edges parallel to the travel axes) will coincide with segments of the grid obtained by following the procedure described in §3.1.

5. The least cost path between two communicating points would coincide with a stair-case path between the points. For example, path 1 is the least cost path between users 1 and 2. However, the least cost path between non-communicating users 2 and 3 will either enter GCR \( G_3 \) (along path 4) or will bypass it (along path 5 or 6). All the paths are as illustrated in Figure 1.

6. Among two rectilinear paths of equal length between two non-communicating points in the presence of GCRs, the path with lesser cumulative travel inside GCRs is the path of lesser cost. For example,
though the lengths of the rectilinear paths 7 and 8 between the users 2 and 4 are equal, path 8 costs less as it does not enter $G_4$. Paths 7 and 8 are depicted in Figure 1.

7. The following results of Butt and Cavalier [5] (who considered convex polygonal GCRs) have been proved to be correct for rectilinear GCRs by Sarkar et al. The correctness follows from Lemma 3.1 and Theorem 3.1 of [17]. The correctness of the results for convex polygonal GCR(s) is still an open problem.

(a) Result 2 (Theorem 1 of [5]): The optimal 1-median in a cell must coincide with a cell corner.

(b) Result 3 (Theorem 2 of [5]): There is at least one optimal solution to the rectilinear $p$-median problem where each new facility location coincides with a cell corner of the grid obtained by passing horizontal and vertical lines through the vertices of the GCRs and the user demand points.

We end this section with the following lemma. The corresponding result for barriers has been proved by Larson and Sadiq (Lemma 3 of [11]).

**Lemma 3.2.1.** A feasible rectilinear least cost path from a user outside a cell $C$ to an infinitesimal point located inside the cell passes through a cell corner $E_k(C), k = 1, 2, 3, 4$ of $C$.

**Proof:** Let us consider a cell $C$ as illustrated in Figure 2. Let $E_1(C) = (x_{\text{min}}, y_{\text{min}})$, $E_2(C) = (x_{\text{min}}, y_{\text{max}})$, $E_3(C) = (x_{\text{max}}, y_{\text{min}})$ and $E_4(C) = (x_{\text{max}}, y_{\text{max}})$ be its corners. Let the area outside cell $C$ be partitioned into areas $E$ (east), $W$ (west), $N$ (north), $S$ (south), $NE$, $NW$, $SE$ and $SW$.

![Figure 2: Proof of Lemma 3.2.1](image)

Let the least cost path from a point $X_1$ inside cell $C$ to a user $X_2 \in NE$ be as illustrated in Figure 2. The path be altered without penalty to pass through the cell corner $E_3(C)$. Similar arguments hold for any $X_2 \in NW, SE, SW$. 


Since any \( X_2 \in E, W, N, S \) will generate its own gridline thereby partitioning cell \( C \) into sub-cells, the lemma follows. ■

The result proved in Lemma 3.2.1 is central to our solution methodology, especially for the case when the new GCR placement does not intersect gridlines.

4 Determining Candidates for Optimal Placement of the new GCR

4.1 Problem Classification

In our problem, the area of the new GCR is a known parameter. If the area of the new GCR exceeds the area of a cell \( C \), then the new GCR cannot be fully contained in the cell. Hence it will intersect gridlines thereby interrupting the flows between user pairs. However if the area of the new GCR is less than the area of a cell, then it can be fully contained in the cell. In this scenario, the user-user interaction will remain unaffected. This observation motivates us to study the problem for two cases, when:

1. the new GCR does not intersect any gridline (in §4.2), and
2. the new GCR intersects at least one gridline in \( S \) (in §5).

4.2 Solution Methodology: The new GCR placement does not intersect any gridline

For a feasible placement \( p \), the area occupied by the interior of the new GCR may not interfere with any gridline, i.e., for all gridlines \( s_t \in S \), \( s_t \cap B = \emptyset \). This also implies that the area occupied by the new GCR is a subset of a cell \( C \), i.e., \( \overline{B} \subseteq C \).

Since we consider two versions of the area location/layout problem, we first provide the solution methodology for the problem when the new GCR’s placement (specified by the location of its top left corner), server location (on its boundary) and its exact dimensions are unknown.

4.2.1 Unknown Boundary Server Problem

We analyze this version of the problem with the aid of the following lemma.

**Lemma 4.2.1.** When \( \overline{B} \subseteq C \), the server \( X \) of the new GCR coincides with \( E_k(C) \), \( k = 1, 2, 3, 4 \).

**Proof:** Since \( \overline{B} \subseteq C \), the new GCR can be fully contained in cell \( C \). In such a case, the new GCR intersects no gridlines. Hence the user-user interaction \( K(p) \) will remain unchanged due to placement of the new GCR. Due to Lemma 3.2.1, the user-server interaction \( J(p) \) is the minimum of four linear functions and is therefore concave. Since \( K(p) \) remains unchanged, \( J(p) + K(p) \) is concave in cell \( C \).

The lemma follows. ■
Due to Lemma 4.2.1, it is necessary to evaluate \( J(p) = \sum_{j \in D} u_j d_p(j, X) \) for \( X = E_k(C) \), \( k = 1, 2, 3, 4 \) and locate \( X \) at the cell corner that minimizes \( J(p) \). The new GCR can have any dimensions \( l \) and \( b \) that satisfies the given area requirement. Note that the optimal server location would coincide with the optimal 1-median location of an infinitesimal facility in cell \( C \) due to result 2 mentioned earlier in §3.2.

4.2.2 Centroid Server Problem

In this version of the problem, the location of the new GCR’s server is known \emph{a priori} to be at its centroid. We now state and prove the following lemma that analyzes this case of the problem.

**Lemma 4.2.1.** When \( \overline{B} \subseteq C \), the optimal placement of the new GCR is such that one of its corners \( E_k(\overline{B}) \) coincides with a corner \( E_k(C) \), \( k = 1, 2, 3, 4 \), of cell \( C \).

**Proof:** When the new GCR does not intersect gridlines, the user-user interaction remains unaffected. For any dimension of the new GCR, the locus of the server \( X \) such that the new GCR does not intersect any gridline that define cell \( C \), is a rectangle \( R \). Due to Lemma 3.2.1, the least cost path from a user \( i \in D \) to the server \( X \in R \subseteq C \) is the minimum of four linear functions (each function evaluates the sum of the weighted rectilinear distances from the existing users to the server through cell corners \( E_k(C) \), \( k = 1, 2, 3, 4 \)) and is therefore concave. Hence the optimal server location must be a corner of rectangle \( R \).

Since the rectangle \( R \) is the locus of the server \( X \) such that the new GCR does not intersect the gridlines defining the cell, the lemma follows. \(\blacksquare\)

Due to Lemma 3.2.1, the users \( i \in D \) can be partitioned into sets \( D_k, k = 1, 2, 3, 4 \). \( i \in D_k \) indicates that the least cost path from a user \( i \) to \( X \) passes through corner \( E_k(C) \) of cell \( C \).

![Figure 3: Optimal dimensions of new GCR](image)

To determine the optimal dimensions of the new GCR, let us consider a situation, as illustrated in Figure 3, when a corner of the new GCR coincides with corner \( E_k(C) \) of cell \( C \). We know \( J(p) = \sum_{i \in D} u_i d_p(i, X) \), where \( d_p(i, X) = d_p(i, E_k(C)) + d_p(E_k(C), X) \). We drop the subscript \( p \) from further discussion. One of the possible values of \( d(E_k(C), X) \) is as follows. The least cost path from cell corners \( E_k(C) \) to the centroid server \( X \) is illustrated in Figure 3 by dotted lines.
where $u$'s are all known constants.

The area of the new GCR is $A = lb$ where $l$ and $b$ are shown in Figure 3. Substituting $b = \frac{A}{l}$ in (3), we have

$$J(p) = \psi + \sum_{i \in D_1} u_{i1} \left[ \frac{l}{2} + \frac{b}{2}(1 + \alpha) \right] + \sum_{i \in D_2} u_{i2} \left[ (x_{max} - \frac{l}{2}) + \frac{b}{2}(1 + \alpha) \right] + \sum_{i \in D_3} u_{i3} \left[ (y_{max} - b) + (x_{max} - \frac{l}{2}) + \frac{b}{2}(1 + \alpha) \right] + \sum_{i \in D_4} u_{i4} \left[ (y_{max} - b) + \frac{l}{2} + \frac{b}{2}(1 + \alpha) \right].$$

Differentiating $J(p)$ as in (4) with respect to $l$ and equating it to zero, we get the value of $l$ that minimizes $J(p)$ as follows:

$$l^* = \sqrt{\frac{A(1 + \alpha)(\sum_{i \in D_1} u_{i1} + \sum_{i \in D_2} u_{i2}) - A(1 - \alpha)(\sum_{i \in D_3} u_{i3} + \sum_{i \in D_4} u_{i4})}{\sum_{i \in D_1} u_{i1} - \sum_{i \in D_2} u_{i2} - \sum_{i \in D_3} u_{i3} + \sum_{i \in D_4} u_{i4}}}$$

where $A, \alpha, u_{ik}$'s are all known constants.
Since $l^\ast$ is the root of a quadratic equation, it can be real or imaginary. Moreover, even if $l^\ast$ is real, it may not lie in the domain of $l$, i.e., $l^\ast \notin [l_{\min}, l_{\max}]$, where $l_{\min}$ and $l_{\max}$ can be determined from cell dimensions and the area of the new GCR. Hence we analyze the function $J(p)$ further. Note that $J(p)$, as given by (4), is a function of $l$, of the form $\psi_1 l + \psi_2 l^2 + \psi_3$, where $\psi_1$, $\psi_2$ and $\psi_3$ are constants in terms of $A$, $\alpha$, $u_{ik}$’s which are all known parameters. This is true for any value of $d(E_k(C), X)$. We now consider the following cases.

1. $\psi_1 > 0$, $\psi_2 > 0$: in this case, $J(p)$ is convex. Hence we can differentiate $J(p)$ with respect to $l$ and set it to zero to obtain a unique minima $l^\ast$ for $l$. The following cases are now possible:

   (a) $l^\ast$ is real. We now consider the following sub-cases:
   i. $l^\ast \in [l_{\min}, l_{\max}]$: $J(p)$ is minimized by $l = l^\ast$.
   ii. $l^\ast \notin [l_{\min}, l_{\max}]$: $J(p)$ is minimized by either $l = l_{\min}$ or $l = l_{\max}$.

   (b) $l^\ast$ is imaginary: $J(p)$ is minimized by either $l = l_{\min}$ or $l = l_{\max}$.

2. $\psi_1 < 0$, $\psi_2 < 0$: in this case, $J(p)$ is concave. Hence $J(p)$ can be minimized by either $l = l_{\min}$ or $l = l_{\max}$.

3. $\psi_1 > 0$, $\psi_2 < 0$: in this case, $J(p)$ is a decreasing function in $l$. Hence $J(p)$ can be minimized by $l = l_{\max}$.

4. $\psi_1 < 0$, $\psi_2 > 0$: in this case, $J(p)$ is an increasing function in $l$. Hence $J(p)$ can be minimized by $l = l_{\min}$.

The previous analysis encompasses all possible cases. Since the area of the new GCR is known, $b$ can now be calculated. This value of $l$ is derived based on the assumed value of $d(E_k(C), X)$ earlier. Different values of $d(E_k(C), X)$ are possible depending on the least cost path that a user $i$ takes to travel from a cell corner $E_k(C)$ to the server $X$ of the new GCR. This would depend on the congestion factor of the new GCR. However similar calculations can be done for different values of $d(E_k(C), X)$ and the value of $l$ that minimizes $J(p)$ can be determined for all such values.

5 The new GCR placement intersects at least one gridline

5.1 Preliminaries

When the area of the new GCR exceeds the area of a cell, the new GCR cannot be contained fully in the cell. Hence it will intersect gridlines possibly disrupting the flows between users as the new GCR may
interfere with the least cost path between existing users. The interaction between user pairs is represented by weights \((w_{ij})\) associated with each user pair. Higher interaction between two users will affect \((i)\) the optimal placement of the new GCR, causing the placement to avoid interrupting the flow (gridline) between highly interacting user pairs, and, \((ii)\) the optimal dimension of the new GCR such that the new GCR’s edge cuts off the flow (gridline) between two highly interacting users minimally, if at all it does so. Another consequence of the new GCR intersecting gridlines is that the least cost path between two users may have to travel inside the new GCR or bypass it, increasing the cost of the path in either case. Hence it is critical to specify/identify the gridlines which the new GCR intersects for a particular placement \(p\).

5.1.1 Set \(Q(l)\)

To this end, consider an initial feasible placement of the new GCR \(p_{ini} = [X, E_4(B)_{ini}, l]\) of the new GCR, such that:

- the new GCR interferes with at least one gridline, i.e., given \(p_{ini} \in F\), there exists \(q \geq 1\) gridlines \(s_1, s_2, \ldots, s_q \in S\) such that \(s_t \cap B \neq \emptyset, \forall t \in \{1, 2, \ldots, q\}\), and
- no edge of the new GCR coincides with a gridline.

Now, let \(Q(l)\) denote the set of placements of the new GCR such that when \(E_4(B) \in Q(l)\), the new GCR will always intersect the same gridlines \(s_1, s_2, \ldots, s_q\), for a particular length \(l\) of the new GCR. More precisely, for a particular \(l\), set \(Q(l)\) denotes the set of locations of \(E_4(B)\) such that the new GCR will not intersect any gridline other than \(s_1, s_2, \ldots, s_q\), i.e., \(Q(l) = \{(x_{E_4(B)}, y_{E_4(B)}) : E_4(B) \in Q(l)\}\). Hence \(Q(l)\) represents the area in which \(E_4(B)\) can be located and can be constructed by moving the new GCR (having a particular length \(l\)) in all directions from the initial location \(E_4(B)_{ini}\) such that the new GCR intersects gridlines \(s_1, s_2, \ldots, s_q\). The idea is illustrated in Fig. 4 in which the new GCR having length \(l\) intersects vertical gridlines \(v_1, v_2, v_3, v_4\) and horizontal gridlines \(h_1, h_2, h_3\). The set \(Q(l)\) obtained for this length \(l\) of the new GCR is illustrated by the dotted rectangle. The boundary of any such set \(Q(l)\) originating from \(E_4(B)_{ini}\) will consist of segments comprising locations of \(E_4(B)\) such that the boundary of the new GCR coincides with some gridline.

We borrow this idea of set \(Q(l)\) from Savas, Batta and Nagi [19] who introduced the concept of set \(Qs\) when dealing with the problem of cutting gridlines. However their definition of set \(Q\) is different compared to the definition in this problem. Since they considered the placement of an arbitrarily shaped new facility with a fixed contour and known server location on the contour, Savas et al. defined set \(Q\) as the set of feasible placements of the server of the new facility such that the new facility intersects the same set of gridlines. We cannot do so, as the server location itself is an unknown in our problem. Before progressing...
further, we state and prove some properties of $Q(l)$ through the following lemmas which shall aid in our analysis of the problem.

**Lemma 5.1.1.** Set $Q(l)$'s are rectangular with their edges parallel to the travel axes.

**Proof:** Consider the example illustrated in Figure 4 in which the new GCR intersects vertical gridlines $v_1, v_2, v_3, v_4$ and horizontal gridlines $h_1, h_2, h_3$. From its initial position given by $E_4(B)_\text{ini}$, the edge $\overline{E_4(B)E_3(B)}$ of the new GCR can be moved parallel to the vertical travel axis in the $+y$ direction until $E_4(B)E_3(B)$ coincides with horizontal gridline $h_0$ (or until $E_1(B)E_2(B)$ coincides with horizontal gridline $h_3$). Conversely the edge $\overline{E_4(B)E_3(B)}$ can be moved parallel to the vertical travel axis in the $-y$ direction until $E_4(B)E_3(B)$ coincides with horizontal gridline $h_1$ (or until $E_1(B)E_2(B)$ coincides with horizontal gridline $h_4$). Analogous arguments can be made for edges $\overline{E_1(B)E_2(B)}, \overline{E_2(B)E_3(B)}$ and $\overline{E_1(B)E_4(B)}$ of the new GCR. Note that for edges $\overline{E_2(B)E_3(B)}$ and $\overline{E_1(B)E_4(B)}$, the movement of the new GCR is parallel to the horizontal travel axis in the $+x$ and $-x$ directions.

However, if gridlines $h_3$ and $v_4$ are generated due to an existing GCR, as illustrated in Figure 5, the new GCR cannot intersect (can only touch) gridlines $h_3$ and $v_4$ without rendering the placement of the new GCR infeasible.

If the new GCR intersects different gridlines, similar arguments hold. The lemma follows. 

Let us denote the corners of set $Q(l)$ as $E_k(Q(l)), k = 1, 2, 3, 4$, starting from the bottom left corner and labeling in the counter clockwise direction. An immediate consequence of Lemma 5.1.1 is that $E_k(Q(l))$ can be expressed as functions of $l$, the length of the new GCR, as shown in Table 1 for the example illustrated in Figure 4. Here $l_1, l_2, l_3, l_4, l_5, b_1$ are lengths and widths of cells as illustrated in Figure 4 and are constants.

**Lemma 5.1.2.** When the new GCR intersects vertical gridlines $h_1, \ldots, h_m \in H$ and horizontal gridlines $h_1, \ldots, h_n \in V$, $Q(l) \subseteq C$, where $C$ is a cell bounded by the gridlines $h_0, h_1, v_0$ and $v_1$. 

![Figure 4: Construction of set $Q(l)$](image1)

![Figure 5: Proof of Lemma 5.1.1](image2)
<table>
<thead>
<tr>
<th>$E_k(Q(l))$</th>
<th>x-coordinate</th>
<th>y-coordinate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 1$</td>
<td>$(l_1 + l_2 + l_3 + l_4) - l$</td>
<td>$\frac{4}{7}$</td>
</tr>
<tr>
<td>$k = 2$</td>
<td>$(l_1 + l_2 + l_3 + l_4 + l_5) - l$</td>
<td>$\frac{4}{7}$</td>
</tr>
<tr>
<td>$k = 3$</td>
<td>$(l_1 + l_2 + l_3 + l_4 + l_5) - l$</td>
<td>$\frac{4}{7} + b_4$</td>
</tr>
<tr>
<td>$k = 4$</td>
<td>$(l_1 + l_2 + l_3 + l_4) - l$</td>
<td>$\frac{4}{7} + b_4$</td>
</tr>
</tbody>
</table>

Table 1: $x$, $y$ coordinates of set $Q(l)$ corners as functions of $l$

**Proof:** From its initial position given by $E_4(B)_{ini}$, the edge $E_4(B)E_3(B)$ of the new GCR can be moved parallel to the vertical travel axis in the $+y$ direction until $E_4(B)E_3(B)$ coincides with horizontal gridline $h_0$ (or until $E_1(B)E_2(B)$ coincides with horizontal gridline $h_n$). Conversely the edge $E_4(B)E_3(B)$ can be moved parallel to the vertical travel axis in the $-y$ direction until $E_2(B)E_3(B)$ coincides with horizontal gridline $h_1$ (or until $E_1(B)E_2(B)$ coincides with horizontal gridline $h_{n+1}$). Analogous arguments can be made for edges $E_1(B)E_2(B)$, $E_2(B)E_3(B)$ and $E_1(B)E_4(B)$ of the new GCR. Note that for edges $E_2(B)E_3(B)$ and $E_1(B)E_4(B)$, the movement of the new GCR is parallel to the horizontal travel axis in the $+x$ and $-x$ directions.

The lemma follows. ■

**Lemma 5.1.3.** $Q(l)_{a} \cap Q(l)_{a'} = \emptyset$

**Proof:** Let sets $Q(l)_{a}$ and $Q(l)_{a'}$ result due to the new GCR intersecting gridlines $s_t \in S_a \subseteq S$ and $s_t \in S_{a'} \subseteq S$ respectively. The lemma follows due to Lemma 5.1.2 and the fact that $\overline{S_a} \neq \overline{S_{a'}} \vee \overline{S_a}, \overline{S_{a'}} \in S$. ■

We conclude this discussion by observing that the overall feasible region $F$ for new GCR placement is composed of:

1. cells whose areas exceed or equal the area of the new GCR, and
2. a number of regions $Q(l)$. The number of set $Q(l)$’s depends on the number of existing GCRs and their relative proximity.

### 5.2 Partitioning Set $Q(l)$

Unlike barriers, one may wish to pass through or bypass GCRs depending on $(i)$ the congestion factor of the GCR, $(ii)$ the dimension of the GCR, i.e., the distance traveled inside the GCR, and $(iii)$ the distance traveled to bypass the GCR. In our problem, when the new GCR intersects gridlines, the flows may choose to $(i)$ pass through, or $(ii)$ left bypass, or $(iii)$ right bypass the new GCR. Here “left bypass” and “right bypass” signify flows that bypass a GCR along its left or right edge respectively. We illustrate the concept of “left bypass” and “right bypass” conventions in Figure 6.
Given the previous motivation, it is critical to uniquely classify flows that are intercepted by the placement of the new GCR as (i) pass through, or (ii) left bypass, or (iii) right bypass. However this classification is complicated by the fact that as the new GCR’s placement \( p \in Q(l) \) changes, the classification of flows may not necessarily remain the same. Hence it is important to determine the range of movement of the new GCR in a set \( Q(l) \) such that the classification of a flow along an intercepted gridline remains unaltered, when the new GCR’s movement is bounded by the range. This can be achieved by partitioning a set \( Q(l) \) by Equal Travel Time Partitions (ETTP), as described in §5.3.

5.3 Equal Travel Time Partitions

With the previous background, we introduce the concept of Equal Travel Time Partitions (ETTP). The concept of ETTP is analogous to the concept of Equal Travel Time Lines (ETTL) introduced by Batta, Ghose and Palekar [2]. ETTLs help to uniquely assign users to cell corners when determining the shortest path from a user to an infinitesimal point in a cell in the presence of impenetrable barriers to travel. The idea of ETTLs is explained briefly with the aid of Figure 7 in which an ETTL \( E_1E_2 \) is generated in the
rectangular cell \( ABCD \) due to user 1. \( E_1E_2 \) splits the cell into sub-cells \( C_1 \) and \( C_2 \). The distance of any point on \( E_1E_2 \) to user 1 bypassing the barrier through cell corners \( A \) or \( D \) is equal. However the shortest path from user 1 to any point in sub-cell \( C_1 \) (but not on \( E_1E_2 \)) passes through cell corner \( A \), i.e., user 1 is uniquely assigned to \( A \). Whereas, for any point in sub-cell \( C_2 \) (but not on \( E_1E_2 \)), user 1 is uniquely assigned to cell corner \( D \).

![Figure 7: Concept of ETTL](image)

![Figure 8: Construction of ETTP](image)

With the previous background, we now establish a methodology to construct \( ETTP \)s for a particular length \( l \) of the new GCR. For the sake of illustration, let us redraw the example depicted in Figure 4 as Figure 8. In Figure 8, consider a feasible placement \( p \in \mathcal{Q}(l) \) of the new GCR, in which the edge \( E_4(B)E_3(B) \) of the new GCR coincides with the edge \( E_1(Q(l))E_2(Q(l)) \) of set \( \mathcal{Q}(l) \), such that the coordinate of \( E_4(B) \) is \((x', y')\). Let us consider a unit of flow between two users along the vertical gridline \( v_1 \) in the \( +y \) direction. If the flow passes through the GCR, then the distance traversed between the points \( y_1 \) and \( y_2 \) is \((\frac{4}{l})(1 + \alpha)\). Whereas, if the flow left bypasses the GCR along the path \( y_1E_1(B)E_4(B)y_2 \), the distance traversed is \( 2(l_1 - x') + \frac{4}{l} \). The time required to travel through or bypass would be equal if \((\frac{4}{l})(1 + \alpha) = 2(l_1 - x') + \frac{4}{l} \), i.e., if \( x' = l_1 - \frac{4\alpha}{2l} \), which is a function of \( l \). Note that \( A, l_1 \) and \( \alpha \) are constants. This value of \( x' \) will remain unaltered for any feasible placement \( p \in \mathcal{Q}(l) \) even if the \( y \) changes. Hence an \( ETTP \) is generated at \( x' = l_1 - \frac{4\alpha}{2l} \). The \( ETTP \) is a vertical line at \( x' = l_1 - \frac{4\alpha}{2l} \) whose end points have \( y \)-coordinates equal to the \( y \)-coordinates of \( E_1(Q(l)) \) (or \( E_2(Q(l)) \)) and \( E_4(Q(l)) \) (or \( E_3(Q(l)) \)), again functions of \( l \), as established earlier in §5.1.1. The \( ETTP \) is “feasible” if \( x' \) lies between the \( x \)-coordinates of \( E_1(Q(l)) \) and \( E_2(Q(l)) \), i.e., \((l_1 + l_2 + l_3 + l_4) - l \leq l_1 - \frac{4\alpha}{2l} \leq (l_1 + l_2 + l_3 + l_4 + l_5) - l \). The bounds on \( x' \) are obtained from Table 1. The significance of this \( ETTP \) is as follows: for a feasible placement \( p \in \mathcal{Q}(l) \), if the corner \( E_4(B) \) of the new GCR lies on the \( ETTP \), the least cost path between a pair of users may pass through or left bypass the new GCR. Considering the right bypass of a unit flow through \( v_0 \), another \( ETTP \) can be generated at \( x'' = \frac{4\alpha}{2l} + l_1 - l \). The bounds on \( x'' \) are same as the bounds on \( x' \). \( ETTP \)s for other vertical and horizontal gridlines can be similarly obtained. However the \( ETTP \)s generated by horizontal
gridlines will partition set \( Q(l) \) horizontally. Hence any gridline \( s \in S \) can generate at most two ETTPs in a set \( Q(l) \). It is pertinent to mention here that there are typically not as many ETTPs formed as may be portrayed by the previous analysis for a set \( Q(l) \).

The feasible ETTPs partition a set \( Q(l) \) along both vertical and horizontal directions. Hence a rectangular \( Q(l) \) is partitioned into smaller rectangles. Let us denote these smaller rectangles as \( R^Q(l) \). The boundary of a \( R^Q(l) \) consists of either (i) ETTPs, or (ii) ETTPs and edges of \( Q(l) \), or (iii) edges of \( Q(l) \). Note that \( R^Q(l) \subseteq Q(l) \). Since the ETTPs can be expressed as functions of \( l \), the coordinates of the corners of \( R^Q(l) \) can also be expressed as functions of \( l \). The significance of \( R^Q(l) \) is as follows: as the placement \( p \in R^Q(l) \subseteq Q(l) \) changes, the classification of flows along the gridlines (for which \( Q(l) \) is formed) as (i) pass through, or (ii) left bypass, or (iii) right bypass does not change. We now demonstrate the the concavity of the objective function when the facility placement is in \( R^Q(l) \). This will provide clues for determining candidate locations for optimal facility placement. We adapt the following lemma from Savas et al. [19] for our purposes:

**Lemma 5.3.1.** For a given \( l \), a candidate optimal solution of \( E_4(\overline{B}) \) lies at a corner of the set \( R^Q(l) \).

**Proof:** Assume the contrary, i.e., there is a solution \( E_4(\overline{B})^* \) which is strictly in the interior of set \( R^Q(l) \). Let \( X^* \) be the server placement associated with this solution. From Theorem 2 of Savas et al., for the given \( l \) and \( X^* \) values, we know that \( J(p) + K(p) \) (now a function of just \( E_4(\overline{B}) \)) is concave over the set \( R^Q(l) \). This implies that one of the corners of set \( R^Q(l) \) is at least as good as the solution \( E_4(\overline{B})^* \) even when the same \( X^* \) is used. The solution could only improve if we choose to optimize \( X \) once the corner was selected.

The lemma follows. ■

We note that Lemma 5.3.1 is possible to prove because the set \( R^Q(l) \) is not a function of \( X \). If it was a function of \( X \), then we would need to establish concavity over \((E_4(\overline{B}), X) \) space.

We also note that when the new GCR’s server is located at its centroid, i.e., \( X \in B \), fixing \( l \) automatically fixes the location of \( X \in B \). Hence Lemma 5.3.1 subsumes the corresponding lemma for the centroid server problem.

Due to Lemma 5.3.1, a candidate optimal for \( E_4(\overline{B}) \) will coincide with a corner of \( R^Q(l) \). Let us denote the corners of \( R^Q(l) \) as \( E_k(R^Q(l)) \), \( k = 1, 2, 3, 4 \), starting from the bottom left corner and labeling them in the counter clockwise direction.

The previous background applies for both the unknown boundary server and centroid server problems. Before proceeding further, we outline the steps of our solution methodology as follows:

1. Identify candidate points for optimal location of \( E_4(\overline{B}) \). This is achieved by Lemma 5.3.1.

2. Identify candidate points for optimal location of the server \( X \). Note that this step is essential for
the unknown boundary server version of the problem (as described later in §5.4.1). However it is redundant for the centroid server version of the problem.

3. For each candidate location of $E_4(\overline{B})$ and for each candidate server location, determine the optimal length of the new GCR (as described later in §5.4.2). For the centroid server problem, we determine the optimal length of the new GCR for each candidate location of $E_4(\overline{B})$. The server location automatically becomes fixed.

We now focus on the unknown boundary server problem and first determine the candidates for optimal location of the server $X$ on the boundary of the new GCR. This will aid in determining the optimal length of the new GCR.

5.4 Unknown Boundary Server Problem

5.4.1 Candidate Server Locations

We have established that the corners $E_k(R^{\mathcal{Q}(l)})$ of $R^{\mathcal{Q}(l)}$ are candidates for optimal placement of the new GCR. We now state and prove the following lemma which identifies the candidate points for optimal location of the server $X$ of the new GCR. Wang, Bhadury and Nagi [20] have proved an analogous result in the presence of impenetrable barriers to travel. Refer to Lemma 1 of [20].

Lemma 5.4.1. The candidate points for optimal server location are:

- corners $E_k(\overline{B})$, $k = 1, 2, 3, 4$, of the new GCR, and
- points of intersection of gridlines with edges of the new GCR.

Proof: For the sake of illustration, let us consider a placement of the new GCR such that it intersects gridlines $v_0, v_1, h_0, h_1$, as illustrated in Figure 9.

Lines 1, 2, 3, 4 are the traversal lines generated by the new GCR. The server $X$ must lie on any edge of the new GCR. Let $X$ lie on $E_2(\overline{B})E_3(\overline{B})$. For any user above line 4, the least cost path to $X \in E_2(\overline{B})E_3(\overline{B})$ must pass through corner $E_3(\overline{B})$. Similarly, for any user below line 2, the least cost path to $X \in E_2(\overline{B})E_3(\overline{B})$ must pass through corner $E_2(\overline{B})$. For any user in regions $P$ and $Q$, the least cost path to $X \in E_2(\overline{B})E_3(\overline{B})$ must pass through $E_2(\overline{B})$ or $E_3(\overline{B})$ or points $a$ or $b$. Clearly the user-server interaction can be minimized by locating the server at $E_2(\overline{B})$, $E_3(\overline{B})$ (points defined by the first bullet), $a$ or $b$ (points defined by the second bullet). For other edges similar arguments hold.

The lemma follows. ■

The candidate points for optimal server location in the case illustrated in Figure 9 are shown by black dots. Note that the coordinates of the candidate points for optimal server location can be expressed in terms of $l$. We now proceed to determine the optimal dimension (i.e., the length $l$) of the new GCR.
5.4.2 Optimal Dimension

Let the new GCR of length \( l \) intersect vertical gridlines \( v_1, v_2, \ldots, v_{z_1} \in V \) and horizontal gridlines \( h_1, h_2, \ldots, h_{z_2} \in H \). The labeling of vertical and horizontal gridlines is in increasing and decreasing order of their \( x \) and \( y \) coordinate values respectively, i.e., \( x_{v_m} > x_{v_{m'}}, \forall m > m' \) and \( x_{v_m} \neq x_{v_{m'}} \) and similarly \( y_{h_n} > y_{h_{n'}}, \forall n < n' \) and \( y_{h_n} \neq y_{h_{n'}} \). Since the GCR intersects vertical gridlines \( v_1, v_2, \ldots, v_{z_1} \), \( l > (x_{v_{z_1}} - x_{v_1}) \). Similarly, since it intersects horizontal gridlines \( h_1, h_2, \ldots, h_{z_2} \), \( l > \left( y_{h_{2}} - y_{h_1} \right) \). Hence \( l \) is bounded by \( (x_{v_{z_1}} - x_{v_1}) < l < \frac{A}{(y_{h_{2}} - y_{h_1})} \). Note that we have a strict inequality condition due to our assumption that the new GCR’s edge should not coincide with a gridline.

The corners of set \( Q(l) \) and \( R^Q(l) \) can be represented as functions of \( l \). Consider the server \( X = (x, y) \) to be located at \( E_4(B) \). Let \( E_4(B) \) coincide with a corner \( E_k(R^Q(l)) \) of \( R^Q(l) \subseteq Q(l) \). As mentioned earlier, in this placement, the classification of flows through the intercepted gridlines is unique as either \((i)\) travel through, or \((ii)\) left bypass, or \((iii)\) right bypass. Hence the user-user interaction \( K(p) \) can be expressed as a function of \( l \), as \( d_p(i, j) \) are functions of \( l, \forall i, j \in D \). Since \( X \) is located at a point whose coordinates are functions of \( l \), the user-server interaction \( J(p) \) can also be expressed as a function of \( l \). Hence \( J(p) + K(p) \) is a function of \( l \). We now study the form of \( J(p) + K(p) \) which will help in determining the value of \( l \) that minimizes \( J(p) + K(p) \).

Note that the length \( l \) of the new GCR is measured along the horizontal travel axis, whereas the width \( \frac{A}{t} \) is measured along the vertical travel direction. The least cost path between two users \( d^*_p(i, j) \) (or between a user and a corner \( E_k(R^Q(l)) \), i.e., \( d_p(i, X) \) consists of vertical and horizontal segments. The horizontal segment can be expressed in terms of \( l \), whereas the vertical segment can be expressed in terms of \( \frac{A}{t} \). Hence the objective function \( J(p) + K(p) \) is of the form \( \lambda_1 l + \frac{\lambda_2}{t} + \lambda_3 \), where \( \lambda_1, \lambda_2 \) and \( \lambda_3 \) are constants in terms of \( A, \alpha, u_i, w_{ij} \) and the coordinates of the existing users. The following cases are possible:

1. \( \lambda_1 > 0, \lambda_2 > 0 \): in this case, \( J(p) + K(p) \) is convex. Hence we can differentiate \( J(p) + K(p) \) with
respect to $l$ and set it to zero to obtain a unique minima $l^*$ for $l$. However if $l^*$ violates the bound

$$(x_{v_{z_1}} - x_{v_1}) < l < \frac{A}{(y_{h_1} - y_{h_2})},$$

$J(p) + K(p)$ can be minimized by either $l^* = (x_{v_{z_1}} - x_{v_1})$ or $l^* = \frac{A}{(y_{h_1} - y_{h_2})}$. It is appropriate to mention here that $l^*$ is the root of a quadratic equation, alike the case described earlier in §4.2.2. Hence it can be imaginary. In that case, $J(p) + K(p)$ can be minimized by either $l^* = (x_{v_{z_1}} - x_{v_1})$ or $l^* = \frac{A}{(y_{h_1} - y_{h_2})}$. 

2. $\lambda_1 < 0$, $\lambda_2 < 0$: in this case, $J(p) + K(p)$ is concave. Hence $J(p) + K(p)$ can be minimized by either $l^* = (x_{v_{z_1}} - x_{v_1})$ or $l^* = \frac{A}{(y_{h_1} - y_{h_2})}$. 

3. $\lambda_1 > 0$, $\lambda_2 < 0$: in this case, $J(p) + K(p)$ is a decreasing function in $l$. Hence $J(p) + K(p)$ can be minimized by $l^* = \frac{A}{(y_{h_1} - y_{h_2})}$.

4. $\lambda_1 < 0$, $\lambda_2 > 0$: in this case, $J(p) + K(p)$ is an increasing function in $l$. Hence $J(p) + K(p)$ can be minimized by $l^* = (x_{v_{z_1}} - x_{v_1})$.

Note that the optimal length $l^*$ obtained from above is for a particular candidate location $\hat{E}_4(B)$ of $E_4(B)$ and for a particular candidate location $\hat{X}$ of $X$. More precisely, $l^*$ minimizes $J(p) + K(p)$ for a candidate placement $p = [\hat{X}, \hat{E}_4(B), l]$. Repeating the same procedure for each candidate location of $E_4(B)$ and for each candidate server location will yield a local minima for a $Q(l)$. Repeating the procedure for all set $Q(l)$ results in the global minima.

5.5 Centroid Server Problem

The only difference in this version of the problem is the centroid location of the server, whose coordinates can again be expressed in terms of $l$. Also while expressing the user-server interaction as a function of $l$, the least cost path from a user to a server can be split up as the least cost path from a user to a corner $E_k(B)$ of the GCR and then from the corner to $X$, i.e., $d_p(i, X) = d_p(i, E_k(B)) + d_p(E_k(B), X)$. $E_k(B)$ and $X$ are both functions of $l$. Hence in this case too, $J(p)$ can be expressed as a function of $l$. Since $K(p)$ can also be expressed as a function of $l$, a similar procedure, as described in §5.4, also applies in this version of the problem.

6 Solution Complexity

Our solution methodology to the area location/layout problem is based on evaluating the user-server interaction $J(p)$ at the cell corners in §4.2.1 and §4.2.2. In §5.4 and §5.5, our analysis is based on the construction of ETTPs for set $Q(l)$s. The solution methodologies of §5.4 and §5.5 are outlined in the following steps more precisely:
1. Construct set $\mathcal{Q}(l)$s.

2. Construct ETTPs for each set $\mathcal{Q}(l)$ and partition $\mathcal{Q}(l)$ into rectangles $\mathcal{R}^\mathcal{Q}(l)$.

3. Uniquely classify flows in each $\mathcal{R}^\mathcal{Q}(l)$ as (i) flow through, or (ii) left bypass, or (iii) right bypass.

4. Identify potential candidate locations of $E_4(B)$ as functions of $l$.

5. Identify potential candidate locations for $X$ (in the unknown boundary server problem), again as functions of $l$.

6. For each candidate location $\widehat{E}_4(B)$ of $E_4(B)$ and each candidate location $\widehat{X}$ of the server $X$, formulate $J(p) + K(p)$ (in §5.4 and §5.5) in terms of $l$, where $p = [\widehat{X}, \widehat{E}_4(B), l]$

7. Differentiate $J(p) + K(p)$ with respect to $l$ and set it to zero.

8. Obtain the optimal dimension of the new GCR for a particular $\mathcal{R}^\mathcal{Q}(l)$.

9. Repeat steps 2 to 8 for each set $\mathcal{Q}(l)$.

Hence the number of cell corners, number of $\mathcal{Q}(l)$s, number of $\mathcal{R}^\mathcal{Q}(l)$s and number of candidate locations of $E_4(B)$ and $X$ are the factors that govern the complexity of the solution procedure.

Let us consider $N$ existing GCRs. Let $\beta$, a constant, be an upper bound on the number of users of each GCR. $N$ GCRs generate at most $N(\beta + 2)$ horizontal gridlines and $N(\beta + 2)$ vertical gridlines, i.e, $O(N)$ gridlines. Hence the maximal number of cell corners generated is $O(N^2)$.

When the new GCR intersects gridlines, the number of set $\mathcal{Q}(l)$s formed depends on the number of gridlines the GCR intersects. It is pertinent to note here that the set $\mathcal{Q}(l)$s are defined when the new GCR intersects successive horizontal and/or vertical gridlines. The new GCR can intersect $N(\beta + 2)$ vertical gridlines in $O(N^2)$ ways because intersecting: (i) one at a time is $N(\beta + 2)$ ways, (ii) two at a time is $N(\beta + 2) - 1$ ways because of the successive intersection requirement, and so on. Similarly, $N(\beta + 2)$ horizontal gridlines can be intersected in $O(N^2)$ ways, and the maximal number of set $\mathcal{Q}(l)$s is thus $O(N^4)$. In practice, the number of set $\mathcal{Q}(l)$s that need to be actually analyzed may be significantly lesser. This is because multiple least cost paths may exist between pairs of users. This is especially true for communicating user pairs between which multiple staircase paths exist. The new GCR placement may intersect one staircase path between two communicating users but the flow may bypass the GCR along another staircase path. Hence a set $\mathcal{Q}(l)$ that does not intersect all such paths at the same time should be excluded from analysis. If the new GCR placement intersects $N'$ vertical gridlines, the maximum number of ETTPs generated for a set $\mathcal{Q}(l)$ is $2N'$, since each gridline can generate at most two ETTPs. Similarly if the new GCR placement
intersects \(N''\) horizontal gridlines, the maximum number of \(ETTPs\) generated for a set \(Q(l)\) is \(2N'\). Hence a set \(Q(l)\) can be partitioned into \((2N'+1)(2N''+1)\) rectangles \(R^{Q(l)}\), i.e., the maximal number of \(R^{Q(l)}\)s is \(O(N^6)\).

Let us now determine the number of candidate locations of \(E_4(B)\) and \(X\). Since a \(Q(l)\) can have at most \(2N'\) vertical partitions and \(2N''\) horizontal partitions, the maximal number of potential candidate locations of \(E_4(B)\) is also \(O(N^6)\). Since the new GCR intersects \(N'\) vertical gridlines, the number of potential server locations due to intersection of gridlines with the new GCR is \(2N'\). Similarly, for horizontal gridlines, the number of potential server locations is \(2N''\). In addition, the vertices of the new GCR are also potential server locations. Hence the maximal number of potential server locations is \((2N' + 2N'' + 4)\), i.e., \(O(N)\) for each candidate location of \(E_4(B)\). So \(O(N^7)\) potential candidate server locations have to be evaluated in the worst case.

Summarizing, the complexity of steps 1, 2, 4 and 5, mentioned earlier in this section is \(O(N^4)\), \(O(N^6)\), \(O(N^6)\) and \(O(N^7)\) respectively. We conclude that the number of cell corners, set \(Q(l)\)s, rectangles \(R^{Q(l)}\)s and candidate locations of \(E_4(B)\) and \(X\) are all polynomially bounded in the number of GCRs. It is pertinent to note here that the construction of \(ETTPs\) for each intercepted gridline can also be performed in polynomial time.

### 7 Conclusions and Future Research

This work addresses the problem of placing a rectangular GCR of given area but unknown dimension in the plane in the presence of other rectangular GCRs. Hence the problem is named as the “area location/layout problem”. The existing GCRs are served by multiple users located in their interior or on their boundary. The new GCR however has one server. We study two versions of the problem: (i) when the server of the new GCR is located on its boundary but the location is unknown, and (ii) when the server location is known \textit{a priori} but is assumed to be inside the new GCR at its centroid. Two types of interactions are considered: (i) between an existing user and the server of the new GCR, and (ii) between pairs of existing users. The objective is to determine the optimal location of the new GCR, its dimensions (specified by its length) and server location of the new GCR thereby minimizing the sum of these two interactions.

The feasible region is partitioned into cells by constructing a grid. When the new GCR placement does not intersect gridlines, i.e., the area of a cell exceeds the area of the new GCR, we show that the optimal location of the new GCR can be drawn from the finite set of cell corners. When the new GCR intersects gridlines, user-user flows are intercepted. Hence we identify the gridlines that are cut off by the new GCR and introduce the concept of \textit{Equal Travel Time Partition} to partition the feasible region into areas where the user-user flows can be uniquely classified as (i) right bypass, or (ii) left bypass, or (iii) travel through.
Partitioning the flows as above helps in accurate calculation of distances between pairs of users. We express our objective function in terms of a single variable, the length \( l \) of the new GCR and obtain a unique minima for \( l \) to minimize our objective function.

For both versions of the problem, our solution methodologies are shown to be polynomially bounded in the number of existing GCRs. Note that since barriers (do not allow travel through, i.e., \( \alpha = \infty \)) and forbidden regions (allow travel through without penalty, i.e., \( \alpha = 0 \)) are special cases of GCRs, the solution methodologies presented for the area location/layout problem would apply in the presence barriers and forbidden regions as well.

The approach in this paper may be extended to study more general location/layout problems. For a single GCR, we need to study the multiple server case. Firstly, we may study the case in which the number of servers is known \textit{a priori}. Secondly, with fixed costs for opening servers and variable costs for using them, the problem to determine the optimal number of servers to be opened can be studied. This is similar to the traditional location-allocation problem, with the addition of spatial considerations. There are also other issues relevant to the shapes of GCRs to be located. An immediate extension would be to consider the area location/layout problem in the presence of convex polygonal GCRs.

**Acknowledgement**

This work was supported by the National Science Foundation, via grant DMI – xxxxxxx. This support is gratefully acknowledged.

**References**


