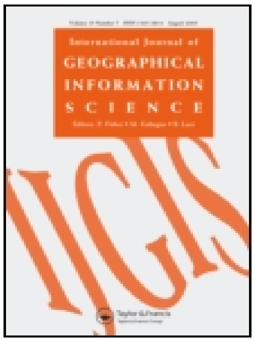

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Efficient measurement of continuous space shortest distance around barriers

Insu Hong & Alan T. Murray

To cite this article: Insu Hong & Alan T. Murray (2013) Efficient measurement of continuous space shortest distance around barriers, International Journal of Geographical Information Science, 27:12, 2302-2318, DOI: 10.1080/13658816.2013.788182

To link to this article: https://doi.org/10.1080/13658816.2013.788182

Published online: 22 Apr 2013.

International Journal of Geographical Information Science, 2013
Vol. 27, No. 12, 2302–2318, http://dx.doi.org/10.1080/13658816.2013.788182

Taylor & Francis
Taylor & Francis Group

Efficient measurement of continuous space shortest distance around barriers

Insu Hong* and Alan T. Murray

GeoDa Center for Geospatial Analysis and Computation, School of Geographical Science and Urban Planning, Arizona State University, Tempe, AZ, USA

(*Received 23 January 2013; final version received 15 March 2013*)

There are many different metrics used to estimate proximity between locations. These metrics are good in some situations and not so good in others, depending on permissible movement behavior. A complicating issue for general metrics to accurately reflect proximity is the presence of obstacles and barriers prohibiting certain directions of movement. This paper develops a continuous space-based technique for deriving a guaranteed shortest path between two locations that avoids barriers. The problem is formalized mathematically. A solution approach is presented that relies on geographic information system (GIS) functionality to exploit spatial knowledge, making it accessible for use in various kinds of spatial analyses. Results are presented to illustrate the effectiveness of the solution approach and demonstrate potential for general integration across a range of spatial analysis contexts. The contribution of the paper lies in the formal specification of the problem and an efficient GIS-based solution technique.

Keywords: optimization; convex hull; Euclidean distance

Introduction

Proximity and distance are perhaps cornerstone features of spatial analysis. The appropriate distance measure or metric is essential for best reflecting movement behavior, closeness, and general spatial relationships. One only need to examine any spatial analysis text to observe the significance of distance as it is central to almost all developed and applied methods, models, or approaches (e.g., Bailey and Gatrell 1995, Fischer and Getis 1997, Church and Murray 2009, Anselin and Rey 2010, O'Sullivan and Unwin 2010, Rogerson 2010, de Smith *et al*. 2012).

A key issue, of course, is how to measure distance. There are several widely used distance measures in spatial analysis: rectilinear, Euclidean, and network distance. Euclidean distance has continued to be a favorite because it is simple, intuitive, and easy to apply. If detailed road network characteristics are important, then network-based distance could be considered, especially in cases where actual movements or flows occur on a transportation system. However, network data sets are sometimes not available in certain situations. In developing countries or rural areas, network data could be incomplete, insufficient, or nonexistent (Yao *et al*. 2012). Moreover, there could be limitations in purchasing or acquiring such data. When this is the case, Euclidean distance is widely used as a surrogate

*Corresponding author. Email: insu.hong@asu.edu

for proximity, and in many cases, it has proven sufficient with high correlation observed between Euclidean and actual distance (Phibbs and Luft 1995, Fone *et al*. 2006, Jones *et al*. 2010, Carling *et al*. 2012, Cudnik *et al*. 2012). For rural areas, especially in developing countries, Euclidean distance is often the only possible choice because there is no existing road network data or actual movement is not restricted to roads (Stock 1983, Oppong and Hodgson 1994).

Of course, Euclidean distance is limited in many ways when used as a proxy for network or actual travel distance as no generic metric would be expected to be accurate or correct in all or even most cases. The reasons are that local nuances are likely varying and also obstacles that hinder directions of travel exist, such as rivers, mountains, coastlines, airports, military installations, and so on. Though there may be some exceptions, these obstacles generally do not allow travel through them. A generic metric like Euclidean distance would, therefore, be challenged to capture or approximate travel movement behavior with any degree of confidence as it ignores the existence of obstacles/barriers(Martin *et al*. 2002, Jordan *et al*. 2004).

Euclidean distance is important for many aspects of spatial analysis and planning. However, barriers are problematic and likely bias results in many ways, such as service areas of certain types of public facilities (Carling 2012). To overcome drawbacks of Euclidean distance, extensions of this metric can reflect spatial patterns more representative of actual travel behavior. Measuring Euclidean distance in the presence of obstacles has been referred to as the Euclidean shortest path (ESP) in computational geometry (Guibas and Hershberger 1989, Hershberger and Suri 1993, Mitchell 1999). It is therefore a recognized problem, with considerable attention focused on its solution. What is lacking to date is a formal mathematical specification of the problem. Further, implementation in a GIS environment provides many operational and efficiency benefits to a range of spatial analytical methods that must rely on metrics like Euclidean distance.

The aim of this article is to formalize and solve the ESP problem to support various spatial analytical methods within a GIS environment. A mathematical problem formulation is presented to account for the presence of barriers. A solution technique based on the notion of a convex hull is introduced and operationalized in a commercial GIS. Empirical results are presented for analysis in an urban region. This article ends with discussion and concluding comments.

Background

As mentioned previously, proximity and distance are central to virtually all spatial analytical methods. Bailey and Gatrell (1995), Fotheringham *et al*. (2000), de Smith *et al*. (2012), O'Sullivan and Unwin (2010), and Rogerson (2010) are popular texts illustrating the ubiquitous nature of proximity and distance in a range of spatial analysis approaches, including spatial autocorrelation, geographically weighted regression, point pattern analysis, geocomputation, spatial interpolation, and exploratory spatial data analysis. A review of Church and Murray (2009) suggests that most location models are dependent on proximity and distance in some way as well. For many reasons, the Euclidean metric is a popular and often relied upon method for deriving proximity and distance. Although there are many application domains that could be discussed, healthcare planning is one where there is much interest in the use and appropriateness of Euclidean distance in the analysis of accessibility (Phibbs and Luft 1995, Fone *et al*. 2006, Higgs 2009, Jones *et al*. 2010, Cudnik *et al*. 2012). Some have advocated the use of Euclidean distance in urban or rural settings (Phibbs and Luft 1995, Jones *et al*. 2010), while others have explored its appropriateness in various contexts (Martin *et al*. 1998, Martin *et al*. 2002, Jordan *et al*. 2004, Higgs 2009, Cudnik *et al*. 2012).

Unobstructed travel is often assumed in most continuous space location models. Obstacles/barriers therefore present a problem because travel between two locations is dependent upon spatial structure. This means that approaches or models that ignore obstacles when they do in fact exist unintentionally introduce errors or biases in the results, producing incorrect objective values and solutions that likely are not optimal. For example, in nonuniformly distributed rural area, Euclidean distance was found to cause sub-optimal service areas when compared with network distance and travel time (Carling *et al*. 2012). To address this, continuous space models must better represent a study region, and this requires travel obstacles to be explicitly considered (Klamroth 2001a, Bischoff and Klamroth 2007). Katz and Cooper (1981) were among the first to undertake such consideration in location modeling, introducing the barrier-restricted Weber problem. Subsequent work has followed to solve this problem (Aneja and Parlar 1994). Problem extension and solution has continued as well (Larson and Sadiq 1983, Batta *et al*. 1989, Klamroth 2001b). This location modeling work highlights the significance of obstacles/barriers. As GIS has various important roles in location modeling (Church 2002, Murray 2010), a GIS-based approach that exploits spatial knowledge about obstacles in a solution approach is very appealing. Further, it provides context for which shortest distance paths likely are useful in broader modeling efforts.

Research focused on ESPs emerged in computational geometry, with a number of solution techniques proposed that address obstacles/barriers. The most prominent approach is the visibility graph, which connects all mutually visible vertices in a given area (Lozano-Pérez and Wesley 1979). Extensive effort has been devoted to computational efficiency issues in the construction of the visibility graph (Welzl 1985, Asano *et al*. 1986, Ghosh and Mount 1991, Pocchiola and Vegter 1996, Kim *et al*. 2004, Zhang *et al*. 2005, Gao *et al*. 2011, Li *et al*. 2011). Another utilized solution approach for the ESP is the shortest path map, which discretizes an area based on the shortest paths from a source point (Hershberger and Suri 1993, Mitchell 1993, Hershberger and Suri 1999). For the case where there are complex polygon barriers, Voronoi diagram (Papadopoulou and Lee 1998) and funnel sequence approaches (Ghosh and Mount 1991) have been developed. Most of these approaches are not implemented in GIS environments. Further, existing approaches typically deal with the entire region, making the methods inefficient in various ways. Even though some research has examined the local visibility graph for more efficient search (Kim *et al*. 2004, Gao *et al*. 2011, Li *et al*. 2011), its efficiency and potential for integration in GIS is not particularly promising.

Euclidean distance is a relatively popular measure used in location and spatial analysis, but its appropriateness as a surrogate for actual travel distance/time is questionable when one considers obstacles/barriers that inhibit certain travel routes. Researchers have addressed various problem nuances and solution techniques. Further, a range of methods has been developed to solve the ESP. However, specialized computational geometry techniques for the ESP have limited efficiency when implemented for spatial analysis in a GIS environment. This is due to a lack of explicit mathematical specification of the problem, but also due to a lack of spatial perspective.

Problem formalization

As discussed previously, of interest in this article is mathematically formulating and solving the ESP problem in the context of spatial analysis. This problem involves identifying the shortest distance/pathway between two locations that avoids all obstacles/barriers that impede travel. If there were no obstacles between points A and B, then Euclidean distance represents the shortest length path. The presence of obstacles between A and B means that direct, straight line travel is not possible. The shortest distance/path between the two points that avoids crossing through any obstacle will necessarily involve a route comprised of one or more intermediate points. The issue then is to determine the necessary intermediate points. This is the so-called ESP. Although described in the literature, the ESP has not been explicitly formulated mathematically. Consider the following notation:

l = index of intermediate points
$(\tilde{x}_l, \tilde{y}_l)$ = coordinates of intermediate point l
p = number of intermediate points

The decisions associated with the shortest path between A and B involve finding the number of intermediate points, p, and their location through which the path is to be routed. Formally, this is

$$\mathit{Minimize} \sum_{l=1}^{p+1} \sqrt{(\tilde{x}_{l-1} - \tilde{x}_l)^2 + (\tilde{y}_{l-1} - \tilde{y}_l)^2} \tag{1}$$

where $(\tilde{x}_0, \tilde{y}_0) = (x_A, y_A)$, $\left(\tilde{x}_{p+1}, \tilde{y}_{p+1}\right) = (x_B, y_B)$ and (x_A, y_A) and (x_B, y_B) are the coordinates of points A and B, respectively. The straight line between two consecutive points, l and $l + 1$, is required to not cross any obstacle. Without doubt, this is not a trivial problem and not readily solvable as presented in Equation (1). Locating a number of points in continuous space while satisfying several conditions is a formidable task. One must locate intermediate points without knowing the exact number of points. As this is a continuous space problem, anywhere in the given study region is a potential location for intermediate points except the interior of obstacles. Further, the constraining conditions are challenging to impose, as all intermediate points have to be located such that connected line segments avoid intersecting all obstacles.

In a GIS environment, obstacles in a study region are represented as polygons. The formal specification of the obstacles is as follows:

k = index of obstacles (entire set K)
$\Omega_k = \left\{\left(\hat{x}_{k1}, \hat{y}_{k1}\right), \left(\hat{x}_{k2}, \hat{y}_{k2}\right), \left(\hat{x}_{k3}, \hat{y}_{k3}\right), \ldots \left(\hat{x}_{kn_k}, \hat{y}_{kn_k}\right)\right\}$
n_k = number of polygon vertices describing obstacle k

For a polygon, it is implicit that two consecutive vertices are connected by a straight line. Without loss of generality, assume a polygon obstacle, k, between points A and B. With this, we can define Φ as the set containing all obstacle vertices and the points A and B, $\Phi = \{(x, y) \in \Omega_k, (x_A, y_A), (x_B, y_B)\}$. The significance of this observation is that the feasible continuous space for intermediate points can be reduced. It has been proven that the intermediate points $(\tilde{x}_l, \tilde{y}_l)$ of the shortest path will consist of points in Φ, that is, $(\tilde{x}_l, \tilde{y}_l) \in \Phi$ (Viegas and Hansen 1985). The search for intermediate points, therefore, can be limited to Φ. Thus, the problem now not only has a finite number of potential locations to consider, but only includes relevant portions of the study region. The difficultly is accounting for line segments between members of Φ that do not intersect the interior of Ω_k for any obstacle k. That is, $i, j \in \Phi$ such that $\overline{ij} \cap \mathit{int}(\Omega_k) = \emptyset$ for any k, where int () is the interior region of an obstacle and $\bar{ij}$ is the line segment connecting vertex i directly to vertex j.

An additional consideration is the regional boundary and that it could inhibit travel. In general, it is assumed that travel outside the regional boundary is prohibited. Let $R = \{(\bar{x}_1, \bar{y}_1), (\bar{x}_2, \bar{y}_2), ..., (\bar{x}_m, \bar{y}_m)\}$ represent the boundary of the study region defined by m vertices. This would need to be accounted for in Φ as well. Consider the following additional notation:

A, B = point of origin/destination
Γ = set of impeding obstacles
Γ_R = set of impeding vertices in R
Φ = set of all impeding vertices and origin and destination points
N_j = set of vertices in Φ that can be connected to vertex j by an arc segment
i, j = index of vertices in Φ
α_{ij} = distance from vertex i to j

$$Z_{ij} = \begin{cases} 1 \text{ if arc from } i \text{ to } j \text{ is on the shortest path} \\ \quad 0 \text{ otherwise} \end{cases}$$

In general, $\Phi = \{(x, y) \in \Omega_k k \in \Gamma, (x, y) \in \Gamma_R, (x_A, y_A), (x_B, y_B)\}$. Also, particularly important in this notation is the set Γ, because it consists of only the obstacles that impede travel between given points A and B, not the entire set of obstacles in the study region. Impeding obstacles are not only those that directly inhibit a straight line segment, but also indirectly impeding obstacles that hinder possible pathways. Moreover, if there is any impeding portion of the regional boundary between points A and B or possible pathways, this is included in Γ_R. By applying a restricted search method, irrelevant obstacles and insignificant portions of the regional boundary do not waste computing effort. An important issue here is techniques for detecting direct and indirect impeding obstacles and relevant regional boundary vertices, Γ and Γ_R, that utilize spatial knowledge.

For deriving the shortest path, a graph of the vertices in Φ, G can be constructed by linking each vertex to members of the set N_j. The set N_j for each vertex in Φ consists of vertices that can be connected without intersecting the interior of obstacles or outside of the regional boundary. Thus, G is a network that represents all feasible path segments to travel from A to B, among which Viegas and Hansen (1985) proved the shortest distance will be found. As will be evident in the next section, there are many graphs G that are possible. We are interested in the most efficient graph G^*, where $G \subset G^*$. Such an efficient graph for the problem is shown in Figure 1a. The regional boundary contains 21 vertices and three obstacles (16 vertices describe the obstacles). The resulting graph contains 13 arcs connecting 12 vertices.

With the above pre-processing, notation, parameters, and decision variables, a formulation of the ESP can be structured based on the graph G (or G^*) for beginning and ending locations A and B. This amounts to a shortest path problem in a network:

$$\textit{Minimize} \sum_j \sum_{i \in N_j} \alpha_{ij} Z_{ij} \tag{2}$$

Subject to:

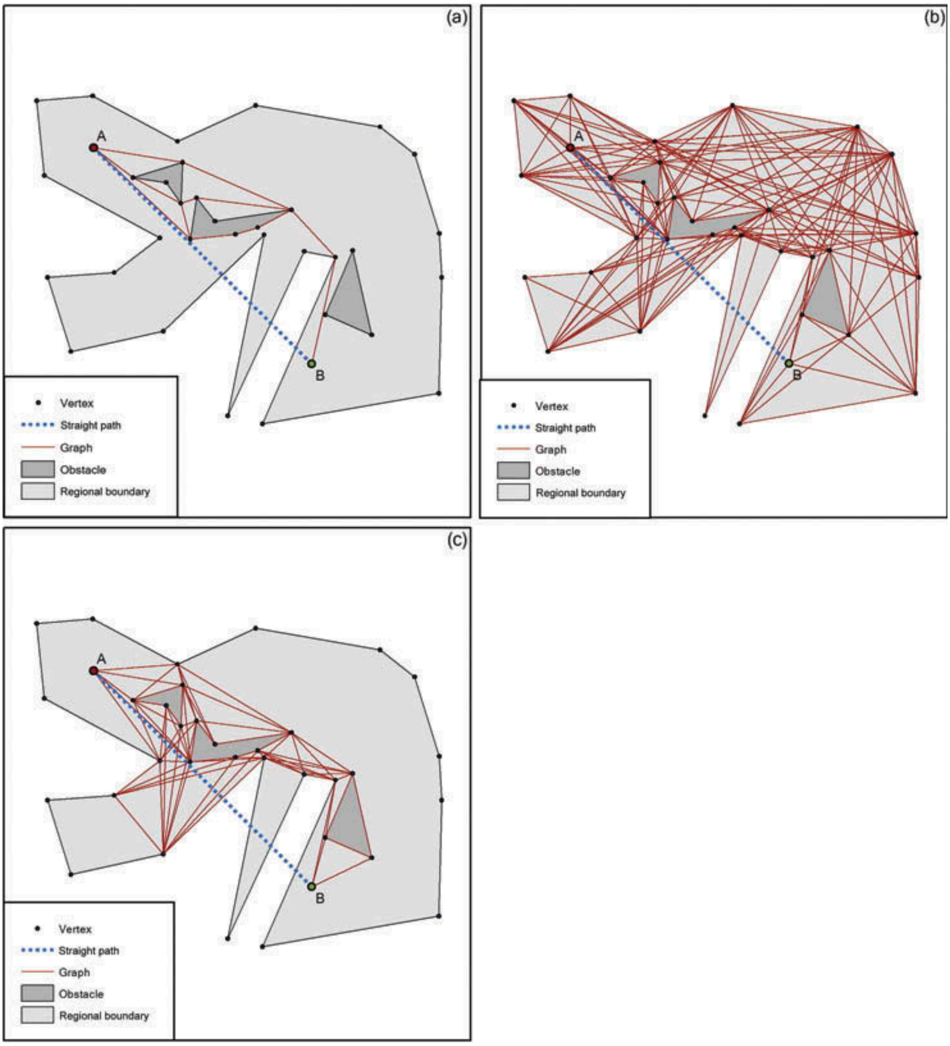

Figure 1. Line segments between obstacles and regional boundary vertices:(a) Graph G*; (b) Visibility graph; (c) Local visibility graph.

$$\sum_{j \in N_{\mathrm{A}}} Z_{\mathrm{A}j} = 1 \tag{3}$$

$$\sum_{j \in N_{\mathrm{B}}} Z_{i\mathrm{B}} = 1 \tag{4}$$

$$\sum_{i \in N_k} Z_{ik} - \sum_{j \in N_k} Z_{kj} = 0 \ \forall k, k \neq \mathrm{A}, \mathrm{B} \tag{5}$$

$$Z_{ij} = \{0, 1\} \ \forall i, j \tag{6}$$

The objective function (2) is to minimize the total length of line segments that connect the given points A and B. Constraint (3) and (4) stipulate flow from a point of origin, A, and to a point of destination, B. Constraint (5) ensures conservation of flow for each intermediate vertex except the point of origin and destination. Constraint (6) limits decision variables to be binary.

Potential solution approaches

As mentioned previously, several solution approaches for the ESP have been developed. Perhaps the most popular methods are visibility graph and local visibility graph. A fundamental concept of the visibility graph is linking mutually visible vertices to each other. A visibility graph is constructed by evaluating whether a line segment is possible between each pair of vertices in the study region and beginning–ending points (Lozano-Pérez and Wesley 1979). Of course, at issue is whether a straight line (line segment) between two points does not intersect any obstacle and remains in the study region. If so, such points are considered visible to each other, and the line segment is included as an arc of the visibility graph. With the visibility graph of nodes and arcs, VG, the shortest path can be determined using a shortest path algorithm.

An important question is how VG relates to the above described problem and its associated characteristics. In terms of resulting graphs, in general, VG $\approx G$. One major difference is that VG includes all vertices of the obstacles and the region boundary. Thus, $\Phi^{\mathrm{VG}} = \{(x, y) \in \Omega_k, k \in K, (x, y) \in R, (x_{\mathrm{A}}, y_{\mathrm{A}}), (x_{\mathrm{B}}, y_{\mathrm{B}})\}$, assuming only one origin –destination pair for comparison purposes. It should be noted, however, that VG typically includes all considered origin – destination pairs, whereas the above problem description has been simplified for a single origin – destination pair. The significance of the VG distinction is that $\left|\Phi^{\mathrm{VG}}\right| > |\Phi|$. That is, the number of vertices in the VG is notably larger than G^*, because G^* only requires impeding obstacles and relevant portions of the boundary. Although attention has been devoted to reducing the number of edges in VG (Rohnert 1986, Ghosh and Mount 1991, Pocchiola and Vegter 1996), all approaches effectively utilize most vertices in the study region. What is new, unique, and different in this article is the recognition that all obstacles and the entire regional boundary need not be considered when evaluating a given origin – destination pair. Spatial knowledge can be exploited to select only relevant obstacles and portions of the regional boundary.

The comparative difference is highlighted in Figure 1b. Recall that there are 37 vertices in total. Figure 1 suggests that for an origin–destination pair, VG is much larger than G^*.This is due to the substantially larger number of identified arcs. In this case, VG has 226 arcs (Figure 1b) while G^*has only 13 arcs (Figure 1a).

Another potential solution approach for the ESP is the local visibility graph, and it differs in noteworthy ways from the visibility graph. A graph is generated in a similar fashion as the visibility graph, but the local visibility graph attempts to filter obstacles (see Zhang *et al*. 2005). By utilizing several spatial queries for a given origin/destination pair (Zhang *et al*. 2005, Gao *et al*. 2011, Li *et al*. 2011), the local visibility graph tries to exploit proximity-based information. Once relevant/impeding obstacles are identified, a local visibility graph (LVG) is generated by evaluating visibility between pairs of vertices. A significant concern, however, is how to detect relevant/impeding obstacles. Zhang *et al*. (2005) suggest a circle-based search method for spatial query that can be applied for the ESP. A search circle is generated centered on the midpoint of two given points; its diameter is the Euclidean distance between the two points. Any obstacles intersecting the circle are considered relevant, forming the set Γ^{LVG}. Thus,

$\Gamma^{\mathrm{LVG}} \subseteq K$. Most of the local visibility graph literature does not consider issues of intersection with the regional boundary. However, the set Γ_R^{LVG} of vertices in R within the search circle can be assumed to remain consistent with the above discussion. Thus, $\Phi^{\mathrm{LVG}} = \{(x,y) \in \Omega_k, k \in \Gamma^{\mathrm{LVG}}, (x,y) \in \Gamma_R^{\mathrm{LVG}}, (x_{\mathrm{A}}, y_{\mathrm{A}}), (x_{\mathrm{B}}, y_{\mathrm{B}})\}$. The local visibility graph has a benefit of reduced size compared to the visibility graph, because $|\Phi^{\mathrm{LVG}}| \leq |\Phi^{\mathrm{VG}}|$. However, there are several drawbacks to the local visibility graph. First, the local visibility graph possibly violates a fundamental constraint of ESP by failing to detect indirectly impeding obstacles. Such a case can happen if the size of the obstacle is larger than the search circle. Second, the reduction in graph size is not necessarily substantial. Proximity-based filtering methods possibly select irrelevant obstacles for Γ^{LVG}. The likelihood increases as the distance between the two points increases. In extreme cases, it is possible that the search circle covers most of the study region if the origin and destination are on opposing sides of the region. All local visibility graph methods have similar limitations in detecting impeding obstacles (Zhang *et al*. 2005, Gao *et al*. 2011, Li *et al*. 2011). Furthermore, the local visibility graph is equally limited in detecting impeding regional boundary vertices. In fact, there is not any discussion/recognition in most local visibility graph methods about regional boundary issues.

Returning to the example shown in Figure 1, the comparative difference for the LVG can be seen. Figure 1c depicts the LVG in this case with 94 arcs. This is fewer than the 226 arcs needed in the VG (Figure 1b) but more than the 13 arcs for G^* (Figure 1a).

While the visibility graph and the local visibility graph are popular for solving the ESP, limited spatial knowledge is utilized. Further, in the latter case, significant problems could be encountered that would produce invalid results if the LVG is relied upon. It is possible to exploit spatial knowledge, and GIS offers much potential for this and operational benefits for general usage and application in spatial analysis.

Deriving an efficient graph

In previous sections, notation was defined associated with an efficient graph, G^*, through which the optimal ESP can be found. Figure 1 supports efficiency inferences for the example problem, as G^* is substantially smaller in size than the VG and LVG. In this section, details on how to efficiently find Γ and Γ_R, and derive G^* using GIS functionality are provided. That is, we would like the smallest and most efficient graph possible, G^*. The most important consideration for efficient solution of the ESP is detecting direct and indirect impeding obstacles in the set Γ. Obstacles that impede a straight line segment can be easily found. However, more complex spatial knowledge is required to find indirect impeding obstacles. Also important is addressing impeding vertices on the regional boundary.

The convex hull is an important concept in computational geometry. It can be an effective way to exploit spatial knowledge for filtering obstacles and regional boundary vertices. Assume there is a finite point set N. A convex hull is the intersection of all convex sets containing N or the smallest and unique convex polygon that contains all points (see De Berg *et al*. 2008). By definition, the length of the boundary of a convex hull is the minimum possible. Further, we assume here that A and B are on the convex hull's boundary.

Theorem 1: *The optimal ESP between two points separated by a single contiguous obstacle will be on the convex hull boundary.*

Proof: Consider two points, A and B, and obstacle k that impedes straight line travel between A and B, like Figure 2. Without loss of generality, the regional boundary is ignored as it is sufficiently at distant from these spatial objects and does not impact travel in any

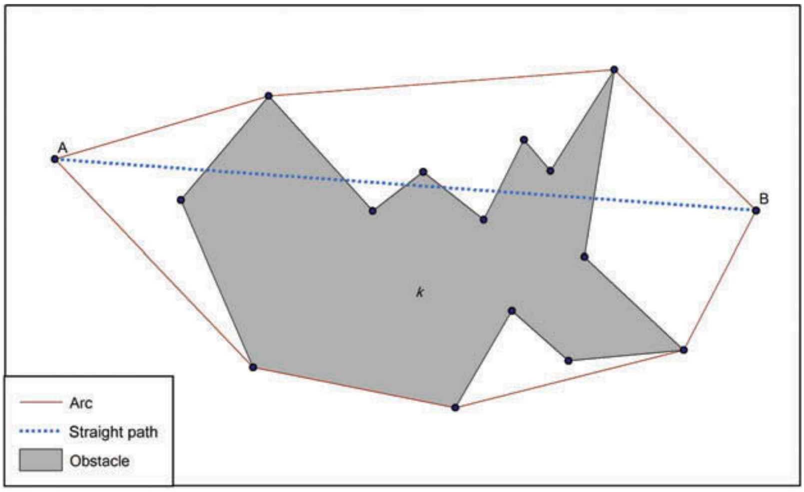

Figure 2. Convex hull for two points and an obstacle.

way. Suppose there exists a point outside the convex hull through which a shorter path exists, and A and B are on the hull. If this were true, the convex hull would not be the minimum length possible. This would contradict the definition of a convex hull, so is not possible. Alternatively, suppose there exists a point inside the convex hull through which a shorter path exists. Such a situation would necessarily create a nonconvex path around the obstacle. By the triangle inequality, this would increase the distance traveled to get from A to B compared to the shortest distance on the convex hull. This too is a contradiction, so any point travelled through on the ESP must be along the convex hull when A and B are on the convex hull.

The significance of Theorem 1 is that an algorithm is possible for selectively identifying vertices and arcs to include in a graph through which the optimal ESP may be found. However, it should be noted that here we consider cases where only the origin and destination points are on the convex hull's boundary. Recall that $\Omega_k = \left\{\left(\hat{x}_{k1}, \hat{y}_{k1}\right), \left(\hat{x}_{k2}, \hat{y}_{k2}\right), \ldots, \left(\hat{x}_{kn_k}, \hat{y}_{kn_k}\right)\right\}$ reflects the vertices of obstacle k and $\Phi = \{(x, y) \in \Omega_k, (x_A, y_A), (x_B, y_B)\}$. Viegas and Hansen (1985) proved that the optimal ESP is comprised of arcs in VG obtained from Φ. However, it is clear from Theorem 1 that a much smaller graph, G^*, is possible, one that includes the optimal ESP. Again, Figure 1 demonstrated empirically the comparative reduction in graph size possible.

As the boundary of the convex hull contains the shortest path around an obstacle, it can be utilized for detecting indirectly impeding obstacles to be included in Γ. Let there be several obstacles between points A and B, as shown in Figure 3a. In this case, only k_1 and k_2 impede the straight path. However, the convex hull for each direct impeding obstacle intersects one or more other obstacles (Figure 3b). If hull line segments intersecting other obstacles are replaced with associated convex hulls, a combined set of hulls results (Figure 3c and d). This then provides an approach for detecting direct and indirect impeding obstacles to be included in the set Γ (and Γ_R).

Based on convex hulls, an efficient algorithm for solving the ESP is proposed, referred to here as *convexpath*. Convexpath derives a graph G^* enabling the shortest path/distance

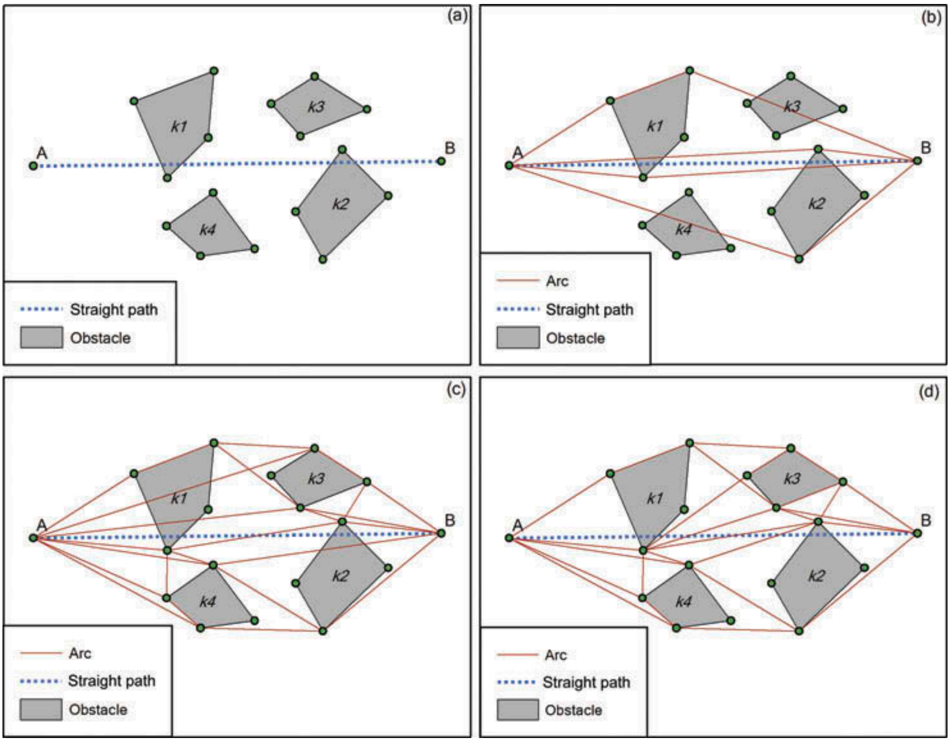

Figure 3. Combined set of multiple convex hulls: (a) Detection of direct impeding obstacles; (b) Separate convex hulls for each direct impeding obstacle; (c) Combined set of convex hulls for direct impeding obstacles; (d) Combined set of convex hulls for direct and indirect impeding obstacles.

between two points to be found by utilizing a series of the convex hulls for impeding obstacles/vertices. The steps of the convexpath approach are detailed in Figure 4. Convexpath evaluates an origin–destination pair using a straight line between them. If there are any obstacles intersecting the straight line, they are considered as directly impeding obstacles and included in set Γ. A convex hull of origin–destination points and each obstacle in Γ is generated in an iterative fashion. If any arc in the convex hulls intersects with another obstacle, that arc is substituted by an additional convex hull. If the obstacle is not in Γ, it is considered indirect impeding and included in Γ. Noncrossing line segments from the origin and destination to the vertices of the segment convex hull are added. This continues until there are no more impeding obstacles. If the initial straight line or arcs of the convex hulls intersect with the regional boundary, a boundary-induced obstacle results. Boundary-induced obstacles are considered impeding vertices Γ_R. The vertices in Γ and Γ_R define the resulting graph G^*, and with this the shortest path/distance is calculated using Dijkstra's shortest path algorithm. The entire process is depicted as pseudo-code in Figure 5, and all operations in convexpath use standard GIS functions.

Theorem 2: *The optimal ESP between two points with multiple obstacles inhibiting travel is contained in graph G^*.*

Proof: Lozano-Perez and Wesley (1979) suggested (and Viegas and Hansen 1985 give proof) that the shortest path from A to B goes through one or more vertices of obstacles

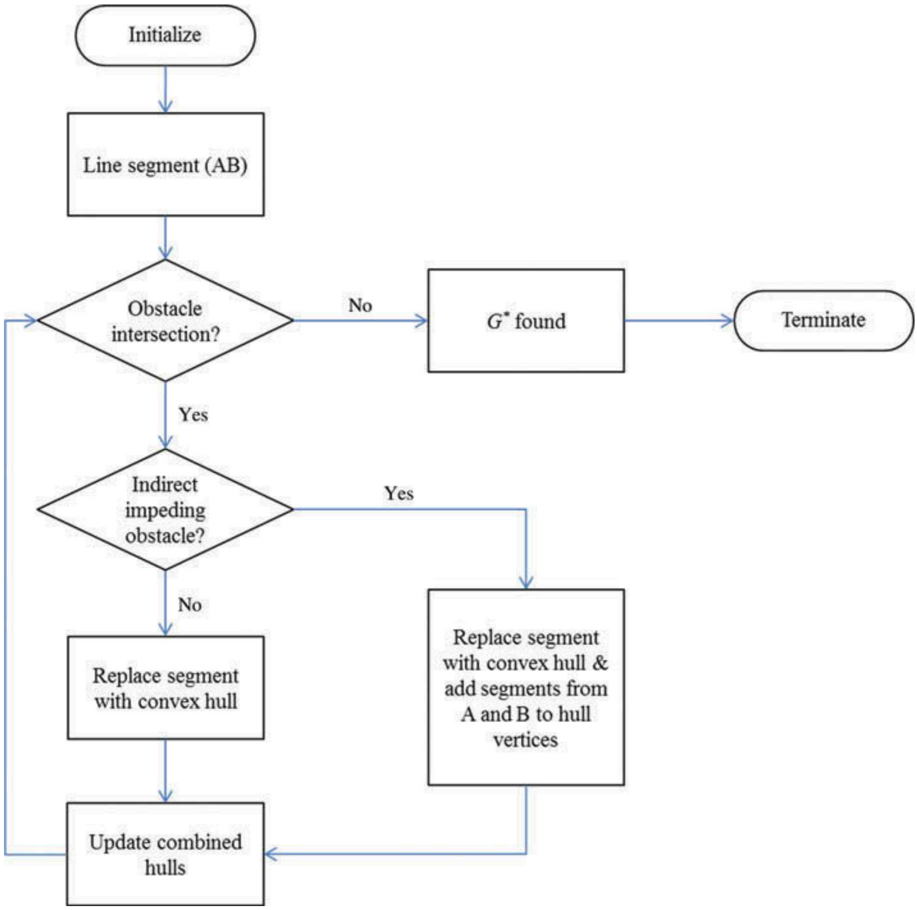

Figure 4. Flowchart of G* generation

$k \in K$. Thus, the search may be limited to $\Phi = \{(x, y) \in \Omega_k k \in K, (x_A, y_A), (x_B, y_B)\}$. What remains to prove is that no vertices or arcs on the ESP have been omitted in G^*. Suppose vertex v is on the optimal shortest path, but $v \notin G^*$. Given A, B, and obstacle k_1, from Theorem 1 there would be two potential shortest pathways on the convex hull for {A, B, k_1} if A to obstacle k_1 is not impeded by another obstacle. Call the last vertex on obstacle k_1 for each pathway α and γ. Note that $\alpha, \gamma \in G^*$ by convex hull algorithm. If $\overline{\alpha B}$ and $\overline{\gamma B}$ do not intersect another obstacle, that is, $\overline{\alpha B} \cap int(k') = \emptyset$ and $\overline{\gamma B} \cap int\left(k'\right) = \emptyset$ for $k' \in K$, then v cannot be on the shortest path because of the triangular inequality as $\overline{\alpha v} + \overline{vB} \geq \overline{\alpha B}$ and $\overline{\gamma v} + \overline{vB} \geq \overline{\gamma B}$. Alternatively, if $\overline{\alpha B} \cap int\left(k'\right) \neq \emptyset$ or $\overline{\gamma B} \cap int\left(k'\right) \neq \emptyset$ for another obstacle k'', then by Theorem 1 the shortest distance would be on the convex hull boundary for the intermediate vertex B and the obstacle (e.g., $\left\{\alpha, B, k''\right\}$ and/or $\left\{\gamma, B, k''\right\}$). Both cases contradict that vertex v, $v \in G^*$, could be on the ESP as the convexpath algorithm accounts for all possible shortest path options through the combined convex hulls.

There are many significant aspects of the convexpath approach. First, it is finite in terms of the number of operations required. It will terminate after a finite number of iterations. Second, the resulting graph, G^*, contains the optimal ESP. Finally, the convex hull–based

```
function convexpath
generate straight line segment for origin-destination points
if straight line crosses obstacle(s) then
  generate convex hull for each intersecting obstacle
  while segments cross obstacles do
      select segment impeded by obstacle
      generate convex hull for intersecting obstacle and segment and end points
      replace segment with convex hull
      if indirect obstacle then
        generate non-crossing segments from origin and destination to convex hull
 identified vertices and segments are nodes and arcs in graph G*
 calculate shortest path in G* using shortest path algorithm
 return shortest distance
else
 return Euclidean distance
```

Figure 5. Pseudo code for the convexpath approach.

approach is very efficient. As only some of the vertices are utilized for the convex hulls, the size of the $G^*$is smaller than G, as $G^* \subseteq G$.

The convexpath method exploits spatial knowledge to solve the ESP in a GIS environment. It finds Γ and Γ_R efficiently and precisely and generates a minimal-sized graph G^* for use in shortest path calculation.

Application results

To demonstrate the operational efficiency of the convexpath approach, a portion of the Tampa, Florida, region is considered. The region is part of school districting work involving the author in an effort to reduce bus transportation costs. The interest here is finding the optimal ESP between an origin and a destination. Six convex obstacles are in the interior of the regional boundary. The shortest distance between an origin and a destination is sought that avoids obstacles. An origin and destination, the regional boundary, and obstacles are depicted in Figure 6. The convexpath, visibility graph and local visibility graph approaches are implemented using C#.NET and ArcObjects 10. The analysis is carried on an Intel i5 personal computer (2.80 GHz) with 12 GB memory.

For assessment, a number of different origins and destinations are considered. In total, the analysis examines 2853different origin–destination instances of the ESP. The results for the three methods are compared in terms of computing time and graph size (the number of vertices and arcs).

To illustrate differences between the three methods, the shortest path/distance for the problem is presented in Figure 6 based on the convexpath approach. In this case, the convexpath approach identified the graph, G^*, as16 vertices and 20 arcs in size. In contrast, the visibility graph approach found VG with1010 vertices and 33104 arcs and the local visibility graph approach determined LVG to be 117 vertices and 919 arcs in size. In percentage terms, VG is over 6000% and LVG is over 600% larger than G^*. Such differences in graph size have direct implications for computation. The total computing time required

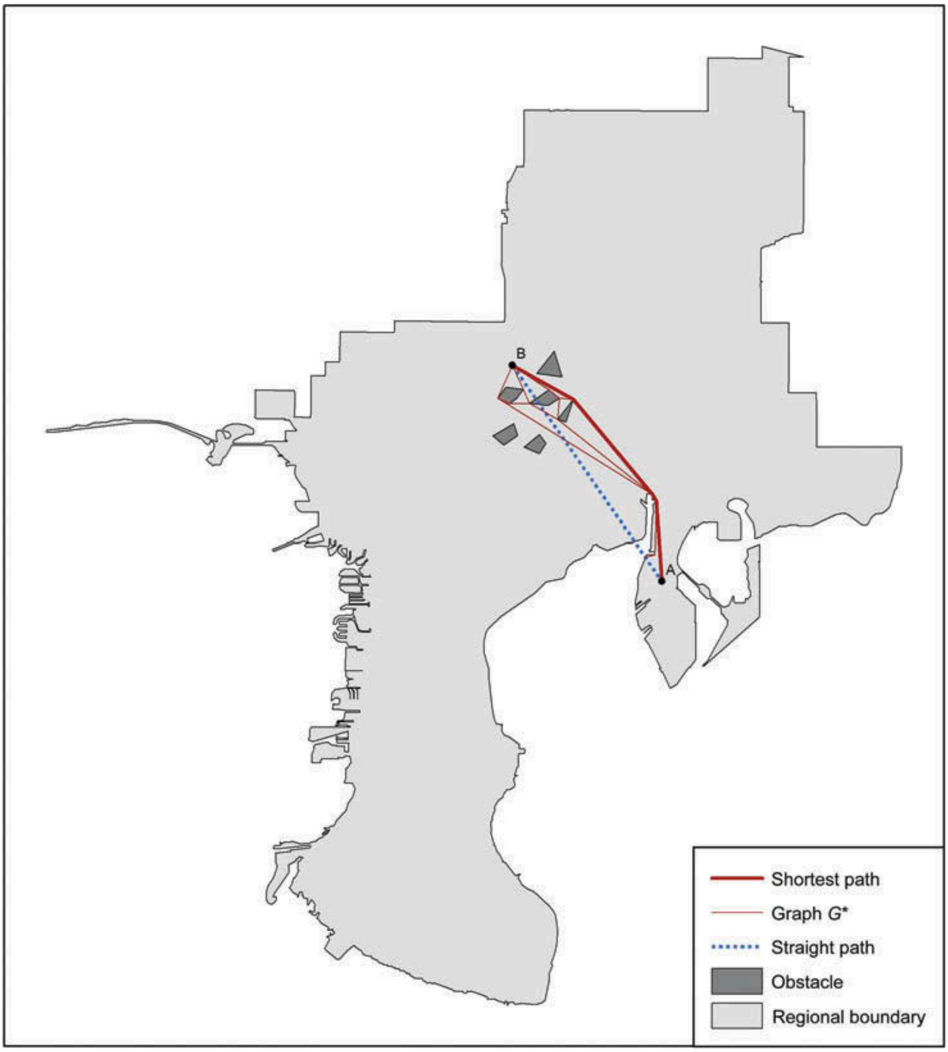

Figure 6. Application area, graph G^* and shortest path.

for the convexpath approach was 4.03 seconds, including graph identification and shortest path solution. In contrast, the visibility graph approach requires 6231.79 seconds and the local visibility graph approach requires 61 seconds. Again, in percentage terms, this translates to over 150,000% total computing time for the visibility graph approach and over 1400% total computing time for the local visibility graph approach when compared to the convexpath approach.

Similar findings were observed for the other origin–destination pairs. Statistically, the 2853 different cases are summarized based on observed minimum, mean, and maximum values. This is reported for number of vertices, number of arcs, and total computing time. Figure 7 gives these findings for each method. In terms of graph size and computing time, the convexpath approach appears to perform well in most cases. The only exception is that there is a minimum case (Figure 7a) where graph size is the same for the convexpath and

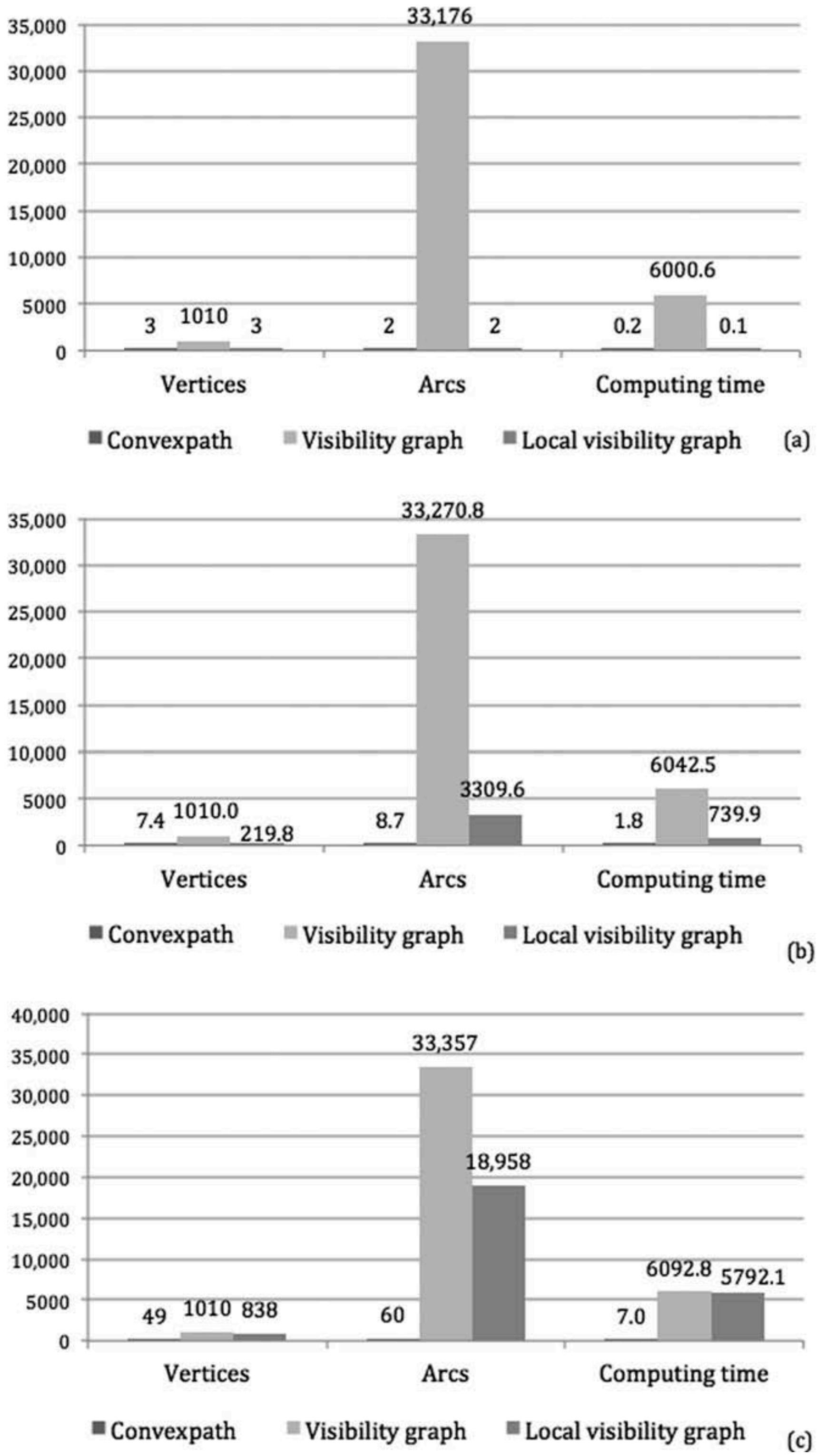

Figure 7. Application results: (a) Minimum; (b) Mean; (c) Maximum.

local visibility graph approaches. Otherwise, the visibility graph always requires a larger graph and more time to process. For example, Figure 7b illustrates that the visibility graph always had substantially larger graphs (33,270.8 arcs on average) and needed 6042.5 seconds of total computing time on average. Therefore, total computing time is 1.8 seconds on average for the convex path approach and 739.9 seconds for the local visibility graph approach. One final point is that the local visibility graph approach is unable to correctly identify the shortest distance in 19 of the 2853 cases. This is due to the previously noted limitations associated with the proximity search circle.

Discussion and conclusions

The comparative results highlight substantial improvements in computational capabilities possible using the convexpath approach for solving the ESP optimally. The reason for this can be attributed to graph size, as the convexpath exploits spatial knowledge for intelligently detecting direct and indirect impeding obstacles and impeding regional boundary vertices, Γ and Γ_R. As a result, the graph through which a shortest path is found is significantly smaller than is possible using the visibility graph or local visibility graph approaches. By exploiting appropriate spatial knowledge in a GIS environment, the convexpath approach is able to solve the ESP efficiently and effectively.

It was noted in this article that the visibility graph approach typically solves the ESP for a set of origins and destinations simultaneously. In the application considered here, it would be possible to apply the visibility graph approach to all origin–destination combinations (2853 in total) at the same time. Doing so would require approximately 10,000 seconds, suggesting that it spends about 3.47 seconds per case. Taking this into account and summing the convexpath solution time for each of the 2853 instances would total approximately 5000 seconds. This is still about half of the total computation time compared to the visibility graph approach. For the local visibility graph approach, total computing time is approximately 2,083,585 seconds, not particularly comparable to either of the other methods and would not be a practical solution approach for large-sized instance of the ESP. Moreover, the local visibility graph approach may not be able to actually find the optimal shortest path/distance in certain situations.

The problem of finding the shortest path between two points in the presence of obstacles is referred to as the ESP problem. Numerous solution approaches have been developed for this problem. However, the ESP has not been formally defined to date. To address this, a mathematical formalization of the ESP as a spatial optimization problem was presented in this article. A new solution approach for the ESP was developed, named convexpath. By utilizing convex hulls, the convexpath approach exploits appropriate spatial knowledge for identifying direct and indirect impeding obstacles and impeding vertices in the region boundary. The convexpath approach constructs a minimal-sized graph and, as a result, is very computationally efficient both to identify and solve for a shortest path. The application results highlighted this effectiveness.

References

Aneja, Y. and Parlar, M., 1994. Technical note—Algorithms for weber facility location in the presence of forbidden regions and/or barriers to travel. *Transportation Science*, 28 (1), 70–76.

Anselin, L. and Rey, S.J., 2010. Perspectives on spatial data analysis. *In:* L. Anselin and S Ray ed., *Perspectives on spatial data analysis*, 1–20. New York: Springer.

Asano, T., *et al*., 1986. Visibility of disjoint polygons. *Algorithmica*, 1 (1), 49–63.

Bailey, T.C. and Gatrell, A.C., 1995. *Interactive spatial data analysis*. Harlow, Essex: Longman Scientific & Technical Essex.

Batta, R., Ghose, A., and Palekar, U.S., 1989. Locating facilities on the Manhattan metric with arbitrarily shaped barriers and convex forbidden regions. *Transportation Science*, 23 (1), 26–36.

Bischoff, M. and Klamroth, K., 2007. An efficient solution method for Weber problems with barriers based on genetic algorithms. *European Journal of Operational Research*, 177 (1), 22–41.

Carling, K., Han, M., and Håkansson, J., 2012. Does Euclidean distance work well when the p-median model is applied in rural areas? *Annals of Operations Research*, 201 (1), 83–97.

Church, R.L., 2002. Geographical information systems and location science. *Computers & Operations Research*, 29 (6), 541–562.

Church, R.L. and Murray, A.T., 2009. *Business site selection, location analysis, and GIS*. Hoboken, NJ: Wiley.

Cudnik, M.T., *et al*., 2012. Surrogate markers of transport distance for out-of-hospital cardiac arrest patients. *Prehospital Emergency Care*, 16 (2), 266–272.

De Berg, M., Cheong, O., and Van Kreveld, M., 2008. *Computational geometry: algorithms and applications*. New York: Springer-Verlag.

de Smith, M., Longley, P., and Goodchild, M., 2012. *Geospatial analysis - A comprehensive guide*. 4th ed. Available from: http://www.spatialanalysisonline.com/

Fischer, M.M. and Getis, A., 1997. *Recent developments in spatial analysis: spatial statistics, behavioural modelling, and computational intelligence*. New York: Springer Verlag.

Fone, D.L., Christie, S., and Lester, N., 2006. Comparison of perceived and modelled geographical access to accident and emergency departments: a cross-sectional analysis from the Caerphilly Health and Social Needs Study. *International Journal of Health Geographics*, 5 (1), 16.

Fotheringham, A.S., Brunsdon, C., and Charlton, M., 2000. *Quantitative geography: perspectives on spatial data analysis*. Thousand Oaks, CA: Sage Publications Ltd.

Gao, Y., *et al*., 2011. On efficient obstructed reverse nearest neighbor query processing. *In*: *Proceedings of the 19th ACM SIGSPATIAL international conference on advances in Geographic Information Systems*, 1–14 November, Chicago, 191–200. New York: ACM.

Ghosh, S.K. and Mount, D.M., 1991. An output sensitive algorithm for computing visibility graphs. *SIAM Journal on Computing*, 20 (5), 888–910.

Guibas, L.J. and Hershberger, J., 1989. Optimal shortest path queries in a simple polygon. *Journal of Computer and System Sciences*, 39 (2), 126–152.

Hershberger, J. and Suri, S. 1993. Efficient computation of Euclidean shortest paths in the plane. *In*: *Foundations of computer science*, *1993*. *Proceedings*, *34th Annual Symposium*, 3–5 November, Palo Alto. IEEE, 508–517.

Hershberger, J. and Suri, S., 1999. An optimal algorithm for Euclidean shortest paths in the plane. *SIAM Journal on Computing*, 28 (6), 2215–2256.

Higgs, G., 2009. The role of GIS for health utilization studies: literature review. *Health Services and Outcomes Research Methodology*, 9 (2), 84–99.

Jones, S.G., *et al*., 2010. Spatial implications associated with using Euclidean distance measurements and geographic centroid imputation in health care research. *Health Services Research*, 45 (1), 316–327.

Jordan, H., *et al*., 2004. Distance, rurality and the need for care: access to health services in South West England. *International Journal of Health Geographics*, 3 (1), 21.

Katz, I.N. and Cooper, L., 1981. Facility location in the presence of forbidden regions, I: Formulation and the case of Euclidean distance with one forbidden circle. *European Journal of Operational Research*, 6 (2), 166–173.

Kim, D.S., *et al*., 2004. Shortest paths for disc obstacles. *Computational Science and Its Applications–ICCSA 2004*, 3045, 62–70.

Klamroth, K., 2001a. Planar Weber location problems with line barriers. *Optimization*, 49 (5–6), 517–527.

Klamroth, K., 2001b. A reduction result for location problems with polyhedral barriers. *European Journal of Operational Research*, 130 (3), 486–497.

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

Li, Z., Gao, Y., and Lu, Y. 2011. Continuous obstructed range queries in spatio-temporal databases. *In*: *System Science*, *Engineering Design and Manufacturing Informatization (ICSEM)*, *2011 International Conference*, 22–23 October, Guiyang, China. IEEE, 267–270.

Lozano-Pérez, T. and Wesley, M.A., 1979. An algorithm for planning collision-free paths among polyhedral obstacles. *Communications of the ACM*, 22 (10), 560–570.

Martin, D., *et al*., 1998. Geographical aspects of the uptake of renal replacement therapy in England. *International Journal of Population Geography*, 4 (3), 227–242.

Martin, D., *et al*., 2002. Increasing the sophistication of access measurement in a rural healthcare study. *Health & Place*, 8 (1), 3–13.

Mitchell, J.S.B., 1993. *Shortest paths among obstacles in the plane*. In: *SCG '93 Proceedings of the ninth annual symposium on Computational geometry*, 18–21 May, San Diego. New York: ACM, 308–317.

Mitchell, J.S.B., 1999. Geometric shortest paths and network optimization. *In*: J.R. Sack and J. Urrutia, ed. *Handbook of computational geometry*, 633–702. New York: Elsevier.

Murray, A.T., 2010. Advances in location modeling: GIS linkages and contributions. *Journal of Geographical Systems*, 12 (3), 335–354.

Oppong, J.R. and Hodgson, M.J., 1994. Spatial accessibility to health care facilities in Suhum District, Ghana. *The Professional Geographer*, 46 (2), 199–209.

O'Sullivan, D. and Unwin, D., 2010. *Geographic information analysis*. 2nd ed. Hoboken, NJ: Wiley.

Papadopoulou, E. and Lee, D., 1998. A new approach for the geodesic Voronoi diagram of points in a simple polygon and other restricted polygonal domains. *Algorithmica*, 20 (4), 319–352.

Phibbs, C.S. and Luft, H.S., 1995. Correlation of travel time on roads versus straight line distance. *Medical Care Research and Review*, 52 (4), 532–542.

Pocchiola, M. and Vegter, G., 1996. Minimal tangent visibility graphs. *Computational Geometry*, 6 (5), 303–314.

Rogerson, P.A., 2010. *Statistical methods for geography: a student's guide*. 2nd ed. Thousand Oaks, CA: Sage Publications.

Rohnert, H., 1986. Shortest paths in the plane with convex polygonal obstacles. *Information Processing Letters*, 23 (2), 71–76.

Stock, R., 1983. Distance and the utilization of health facilities in rural Nigeria. *Social Science & Medicine*, 17 (9), 563–570.

Viegas, J. and Hansen, P., 1985. Finding shortest path in the plane in the presence of barriers to travel (for any l_p-norm). *European Journal of Operational Research*, 20 (3), 373–381.

Welzl, E., 1985. Constructing the visibility graph for n-line segments in O (n2) time. *Information Processing Letters*, 20 (4), 167–171.

Yao, J., *et al*., 2012. Geographic influences on sexual and reproductive health service utilization in rural Mozambique. *Applied Geography*, 32 (2), 601–607.

Zhang, J., *et al*., 2005. Query processing in spatial databases containing obstacles. *International Journal of Geographical Information Science*, 19 (10), 1091–1111.